

پیش همارائی، آری یانه؟

نوشته الین اسونونیوس^۱

ترجمه علی مزینانی^۲

چکیده: هنگامی که در سال ۱۹۵۱ امکان بالقوه استفاده از کامپیوتر در بازیابی اطلاعات مطرح شد، «مورتیمرتوب»^۳ این عقیده را مطرح کرد که یک ترکیب^۴ (نحو) ساده ماشینی را می‌توان به‌طور مؤثر جایگزین ترکیب نسبتاً پیچیده‌ای کرد که در آن زمان برای ساخت سرعنوانهای موضوعی به‌کار می‌رفت. فرق بین زبانهای پیش‌همارا و پس‌همارا برای ساخت سرعنوانهای موضوعی در این است که اولی از ترکیب معمولی و رایج برای سرعنوانی موضوعی استفاده می‌کند در حالی که زبان دومی، برای تهیه سرعنوانهای موضوعی از ترکیب ماشینی با استفاده از عملکردهای بولی^۵ بهره می‌جوید. در خلال انقلاب کامپیوتری، بازیابی اطلاعات دچار تحول شده و استفاده از زبانهای پس‌همارا عمومیت یافت؛ اما در عین حال پدیده‌ای متضاد در حال شکل‌گیری بود یعنی توسعه زبانهای^۶ نمایه‌سازی رشته‌ای^۷ زبانهای نمایه‌سازی رشته‌ای مثل «پرسی» پیش‌همارایی را تا آنجایی که از لحاظ منطقی امکان پذیر باشد به‌کار می‌گیرد، در این حالت لازم است که سرعنوانهای موضوعی را به گزاره‌های موضوعی^۸ که معادل و بیانگر موضوعهای اسناد هستند تعمیم داد. از آنجایی که از نیمه قرن حاضر تجزیه و تحلیل موضوعی بیشتر مدنظر قرار گرفته است. این مقاله، ابتدا به ریشه‌های تاریخی آن پرداخته و سپس مشکلات و محسّنات پیش و پس‌همارایی در فهرستهای پیوسته کنونی را مورد تجزیه و تحلیل قرار می‌دهد.

1. Svenonius, Elaine. "Pre-Coordination or not?" *IFLA Satellite Meeting*, Lisbon: 17-18 August 1993.

۲. دانشجوی کارشناسی ارشد مدیریت و اطلاع رسانی در دانشگاه علوم پزشکی ایران.

۳. مورتیمرتوب (Mortimer Taube) پایه‌گذار نمایه‌سازی نکواژهای هماراست. - م.

4. syntax

5. Boolean operators

6. string index languages

۷. رشته یا استرینگ = در ساختار داده‌ها، ترتیبی از عناصر خطی هستند، مثل کاراکترهای الفبایی عددی. در اینجا مقصود از رشته عبارت است از تعدادی اصطلاح یا لغت که برطبق قوانین ترکیب زبان ساختگی به یکدیگر ارتباط داده شده‌اند. - م.

8. subject statements

فصلنامه کتاب. دوره چهارم، شماره چهارم،

زمستان ۱۳۷۲

سؤالی که عنوان این مقاله قرار گرفته برای نخستین بار، در سال ۱۹۵۱، در مقاله‌ای که به قلم «مورتیمرتوب» و همکارش «ای اف تامسون» (۱) نوشته شده بود، مطرح گردید. این مقاله هفت صفحه‌ای که به صورت فتوکپی تکثیر شده بود، هیچ‌گاه منتشر نشد. این مقاله در سمپوزیوم «مساعدهای مکانیکی برای دکومانتاسیون اسناد شیمی» که توسط بخش متون شیمی انجمن شیمیدانان امریکا برگزار شده بود، خوانده شد. اکنون بعد از چهل و دو سال، شاهد تغییرات عمده‌ای در زمینه تکنولوژی و نظریه‌ی بازیابی اطلاعات هستیم. در آن زمان نیز تفاوت‌های بین بازیابی موضوعی پیوسته و دستی مدنظر بوده و بینشهایی در این زمینه وجود داشت. در این مقاله سعی خواهیم کرد تا مختصراً تاریخچه‌ی دستیابی موضوعی را مورد بررسی قرار دهیم زیرا در ارتباط با موضوع اصلی ما، یعنی پیش‌همارایی است. سپس با استفاده از بیست معیار ارزشیابی، اثر بخشی نسبی هریک از زبانهای نمایه‌سازی پیش و پس‌همارا را مقایسه خواهیم کرد.

«مورتیمرتوب» از نخستین کسانی است که امکانات بالقوه به‌کارگیری کامپیوتر در بازیابی اطلاعات را درک کرد. وی مفهوم نمایه‌سازی همارا و معرف آن یعنی «نظام تک‌واژه‌ای»^۱ را ارائه کرد. در آن زمان نوعی از زبان نمایه‌سازی همارا به صورت نظامهای دستی و برگه‌های لبه سوراخ‌دار وجود داشت که بسیار اولیه بود و تنها عملکرد «و» آن‌هم به طور مکرر در آن به‌کار می‌رفت. در چنین وضعیتی تنها توب بود که درک کرد با استفاده از استراتژیهای تجسس بولی در نظامهای مکانیزه می‌توان به قدرت کامل زبان نمایه‌سازی همارا دست یافت.

توب، نمایه‌سازی همارا را با توجه به مختصات ریاضی تعریف کرد، بدین معنا که وی نظامی دو بعدی را ارائه داد که از دو محور مختصات

X و Y تشکیل شده‌اند و فرض کرد که این مختصات به جای اینکه بیانگر مقادیر ریاضی باشند نشانگر طبقات مختلف اطلاعات هستند. (۲) مثلاً سن و سطح تحصیلات یک فرد می‌تواند نشانگر چنین طبقه‌بندی اطلاعات باشد، مقیاس‌بندی آنها ضروری است. مثلاً مقیاس سال را می‌توان برای اندازه‌گیری سن و تحصیلات به کار برد و آنها را بر روی محور مختصات X و Y نشان داد. برای ارائه دیگر اطلاعات می‌توان مختصات دیگری را اضافه کرد. معه‌ذا توب خاطر نشان ساخت که ایجاد نظامی از مختصات که دارای چندین بعد و نمایانگر کل دانش باشد هرگز امکان پذیر نیست. توب اعتقاد داشت برای مؤثر بودن نمایه‌سازی همارا، تعریفی که در ارتباط با حوزه خاصی از دانش باشد الزامی است. علاوه بر این امکان تحلیل چنین حوزه خاصی از علم به صورت طبقات مختلف بایستی امکان‌پذیر باشد بنحوی که هر ترکیبی از این طبقات معقول نماید.

نمایه‌سازی همارا یا همارایی کتابشناختی^{۱۰} (اصطلاحی که توب به کار می‌برد) را در چنین شرایطی بایستی نظام پس‌همارا نامید. از نقطه نظر توب و دیگر نظریه پردازان اولیه بازایی اطلاعات، پس‌همارایی به وضوح، مناسب یک نظام ماشینی بود، به نحوی که می‌توانست خیلی سریع جایگزین روش استاندارد پیش‌همارایی در کتابخانه گردد، اما چنین نشد. پیش از ادامه تاریخچه لازم است فرق بین نمایه‌سازی پس‌همارا و پیش‌همارا را مشخص کنیم.

زبانهای نمایه‌سازی پس‌همارا و پیش‌همارا در سه جنبه تفاوت دارند:

— چه کسی همارایی را انجام می‌دهد؟

— در چه زمانی همارایی انجام می‌شود؟

— اصطلاحات چگونه همارا می‌شوند؟

در نظام پیش‌همارا، فهرستنویس یا نمایه ساز واژه‌ها را به صورتی زنجیروار به یکدیگر ارتباط می‌دهد تا سرعنوانهای موضوعی یا رشته‌های

(استرینگ‌های) موضوعی شکل گیرند. در نظام پس‌همارا استفاده‌کننده عمل همارایی را انجام می‌دهد که نتیجه آن سرعنوان موضوعی نیست بلکه درخواستهای موضوعی است. (۳) در زبان پیش‌همارا، متصل کردن رشته‌های اصطلاحات به‌صورت زنجیر پیش از شروع جستجو انجام می‌شود در حالی که در نظامهای پس‌همارا اصطلاحات در هنگام جستجو به یکدیگر ارتباط داده می‌شوند.

در یک زبان پیش‌همارا، اصطلاحات براساس قوانین نحو و ترکیب که برای سرعنوان موضوعی یا زبان نمایه‌سازی تعریف شده‌اند، به یکدیگر متصل می‌شوند. از نقطه‌نظر یک نمایه‌ساز، زبان پس‌همارا دارای ترکیب کم یا فاقد ترکیب است و واژگان آن متشکل از اصطلاحات واحد است. از نقطه نظر استفاده‌کننده، در جستجویی که در آن از زبان پس‌همارا استفاده می‌شود تلفیق اصطلاحات امری ضروری است که در مفهومی گسترده‌تر می‌توان گفت که زبان پس‌همارا مستلزم بهره‌جویی از ترکیب بولی است. گاهی اوقات ترکیب بولی توسط عملگرهای اضافی همچون عملگرهای^{۱۱} مجاورتی تقویت می‌شود. در یک زبان پیش‌همارا ارتباط بین اصطلاحات یک رشته، حالت دستوری^{۱۲} دارد. در یک زبان پس‌همارا ارتباط بین اصطلاحات در پرسش مورد جستجو حالت منطقی دارد. برای درک بهتر و ساده‌تر اختلاف بین زبانهای پیش‌همارا و پس‌همارا بایستی به‌طور خلاصه چنین گفت: واژگان زبان پس‌همارا متشکل از اصطلاحات است (در حالت ایده‌آل آن از واژه‌های منفرد تشکیل شده) در حالی که واژگان زبان پیش‌همارا علاوه بر اصطلاحات از یک سری رشته اصطلاحات نیز تشکیل شده است. در مجموع می‌توان گفت که زبان پس‌همارا یک زبان نمایه‌سازی اصطلاحی است، در حالی که زبان پیش‌همارا یک زبان نمایه‌سازی رشته‌ای است. (۴)

اکنون به تاریخچه برمی‌گردیم. در سال ۱۹۵۲ بنگاه اطلاع‌رسانی فنی

نیروهای مسلح آمریکا (ASTIA) قراردادی با «توب» و همکارانش منعقد کرد تا نظام تک‌واژه‌های آنها را سروسامان داده و با توجه به دیگر زبانهای نمایه‌سازی آن را مورد ارزیابی قرار دهند. (۵) از طریق آزمونهای متعدد نظام تک‌واژه‌های نیروهای مسلح با چندین نظام رده‌بندی و سرعنوان موضوعی و روشهای مختلف نمایه‌سازی مقایسه گردید. در یکی از این آزمونها، سرعنوانهای موضوعی بخش اطلاع‌رسانی فنی کتابخانه کنگره به‌صورت تک‌واژه‌ها درآورده شد و دو نظام دیگر از نظر مؤثر بودن مورد بررسی قرار گرفت. این اقدام اولین مقایسه عملی یک نظام پیش‌همارا و پس‌همارا با یکدیگر بود. در آزمون مشابهی سرعنوانهای موضوعی اطلاع‌رسانی فنی نیروهای مسلح آمریکا به نظام تک‌واژه‌ها تبدیل گردید. توب، شانزده معیار برای ارزیابی زبانهای مختلف نمایه‌سازی مطرح کرد که از طریق آنها مؤثر بودن این زبانها مورد بررسی قرار گرفت. (۶) (در بخش دوم مقاله که زبانهای پیش‌همارا و پس‌همارا مورد مقایسه قرار می‌گیرد، این معیارها هم بررسی خواهند شد.) نباید فراموش کرد که این معیارهای مفید تحت‌الشعاع واقعه مهمی قرار گرفتند که تغییر عمده‌ای در ارزیابی زبان نمایه‌سازی به‌وجود آورد.

این واقعه مهم، تجربه «کرانفیلد»^{۱۳} بود که در اواخر دهه پنجاه روی داد. (۷) هدف تجربه «کرانفیلد» بررسی کارایی نسبی زبانهای مختلف نمایه‌سازی بود. اهمیت این تجربه در این است که برای اولین بار زبانهای نمایه‌سازی در مقیاسی بزرگ و به صورتی عملی مورد ارزیابی قرار گرفت. این کارهای «توب» بود که باعث شد تا تجربه کرانفیلد به‌وجود آید. توب در بررسی امتیازات نسبی هریک از زبانهای پیش‌همارا و پس‌همارا دریافت که برای اندازه‌گیری عوامل تشکیل دهنده یک جستجوی موفق روشی وجود ندارد. (۸) در تجربه کرانفیلد روش اندازه‌گیری دوگانه، یعنی اندازه‌گیری «مانعیت» و «جامعیت» توسعه یافت و این امر باعث شد تا در

تاریخ روشهای دستیابی موضوعی، نقش برجسته‌ای برای تجربه کرانیفیلد در نظر گرفته شود.

جالب است بدانیم که در تجربه کرانیفیلد نظام تک‌واژه‌ای یکی از آن نظامهایی بود که مورد ارزیابی قرار گرفت. نظام تک‌واژه‌ای با نظامهای زیر مقایسه گردید: با یک زبان سرعنوانهای موضوعی الفبایی، یک زبان چهاریزه‌ای و رده‌بندی دهنده‌ی جهانی. در این تجربه مشخص شد که هر نظام به تنهایی خوب عمل می‌کند و این نتیجه قابل قبول همه نبود (زیرا طرح تجربه مناسب نبود).

«کالوین موئرز»^{۱۴} که نظریه پردازی کارآمد در زمینه‌ی بازایی اطلاعات بود، نظام تک‌واژه‌ای را مورد ارزیابی قرار داد. «موئرز» به دلیل ارائه اصطلاح «توصیف‌گر»^{۱۵} در ادبیات بازایی، فردی کاملاً شناخته شده‌است. وی نظام تک‌واژه‌ای توب را به دلیل اینکه واژه‌ها را از متن، عنوان یا چکیده‌ای که توسط نویسنده تهیه می‌شود، استخراج می‌کند و مستقیماً به صورت اصطلاحات نمایه‌سازی به کار می‌برد، مورد انتقاد قرار داد. (۹) روشی که وی ارائه کرده بود آنقدر از سادگی و سهولت برخوردار بود که باعث شده بود تا عیوب جدی نظام تک‌واژه‌ای آشکار نگردد. یکی از عیوب نظام تک‌واژه‌ای این بود که براساس واژه‌ها پایه‌گذاری شده بود، در حالی که نظام توصیف‌گر «موئرز» براساس مفهوم تهیه شده بود. جدیدترین عیب نظام تک‌واژه‌ای که موئرز آن را پیدا کرد عبارت است از: «نظام تک‌واژه‌ای، چون بیشتر براساس واژه‌ها پایه‌گذاری شده تا مفهوم، در کنترل واژه‌های مترادف ناتوان است».

در خلال دهه ۶۰ و ۷۰ به دلیل انقلاب کامپیوتری، بازایی اطلاعات دچار تحول شده و درست در همین زمان هردو نوع زبانهای پس‌همارا (هم با کنترل واژگانی و هم بدون آن) عمومیت یافتند. در همین زمان پدیده‌ای متضاد در حال شکل‌گیری بود، یعنی توسعه نوع جدیدی از

زبانهای نمایه‌سازی رشته‌ای. این زبانها همگی از ترکیب ماشینی شده بهره می‌جستند. بسیاری از این زبانها مثلاً «پرسی» پیش‌همارایی را تا آنجایی که از لحاظ منطقی امکان‌پذیر باشد بکار می‌گیرد، در این حالت لازم است که سرعنوانهای موضوعی یا رشته‌های واژه‌ها به‌نحوی تهیه شده باشند که معادل و بیانگر موضوعهای مدارکی باشند که قرار است نمایه‌سازی شوند. (۱۰) قبل از بحث درباره خصوصیت «معادل یا بیانگر بودن» موضوعهای مدارک لازم است تا به‌طور مختصر، دیگر خصوصیات زبانهای نمایه‌سازی رشته‌ای نوین را مورد بررسی قرار دهیم زیرا در ارتباط با موضوع این مقاله است. همان‌طور که خواهیم دید سؤالی که عنوان مقاله قرار گرفته (پیش‌همارایی، آری یا نه؟) سؤالی ساده نیست و پاسخ دادن به آن بستگی به آشنایی با زبانهای پیش‌همارا دارد.

همان‌طور که قبلاً ذکر کردیم بسیاری از زبانهای نمایه‌سازی رشته‌ای جدیدتر، از ترکیب ماشینی شده استفاده می‌کنند. در چنین زبانهایی کامپیوتر (نه نمایه‌ساز) عناصر را به یکدیگر متصل می‌کند تا رشته‌های سرعنوان موضوعی را به‌وجود آورد. وظیفه نمایه‌ساز تنها وارد کردن عناصر اختصاصی یا اصطلاحات به درون کامپیوتری است که قرار است رشته موضوعی را به‌وجود آورد، همزمان با این عمل، نمایه‌ساز، چهریزه‌ها^{۱۶} یا نقش نماهای^{۱۷} متعلق به اصطلاحات را نیز وارد کامپیوتر می‌کند. سپس کامپیوتر با استفاده از یک ترکیب آلوگوریتمی که با توجه به این نقشها و چهریزه‌ها برایش تعریف شده، مجموعه‌ای از رشته‌ها را به‌وجود می‌آورد. این رشته‌ها، سرشناسه‌های نمایه هستند. معمولاً هر رشته‌ای در کل مجموعه نوعی حرکت چرخشی^{۱۸} یا (گردشی) از دیگر رشته‌هاست. برای مثال در زبان «پرسی» یک نمایه‌ساز می‌تواند اصطلاحات زیر را ارائه و کدبندی کند:

16. facets

18. permutation

17. roles indicator

- آموزش (به عنوان یک عمل و کار رمزگذاری^{۱۹} شده)
 - متصدیان ماشین (به عنوان سیستم کلیدی رمزگذاری شده)
 - کانادا (به عنوان محیط و منطقه جغرافیایی رمزگذاری شده)
- پس از این مرحله کامپیوتر سه سرشناسه به شرح زیر را تهیه می‌کند:

کانادا

متصدیان ماشین. آموزش

متصدیان ماشین. کانادا
آموزش

آموزش. متصدیان ماشین. کانادا

از دیگر خصوصیات زبانهای نمایه‌سازی رشته‌ای جدید این است که تا حد زیادی ساختگی هستند و فرق بین زبان نمایه‌سازی ساختگی^{۲۰} و شمارشی^{۲۱} در این است که در زبان ساختگی، رشته‌های مجاز واژه‌ها یا اصطلاحات را می‌توان از عناصر واژگان اصلی زبان استخراج و براساس قوانین ترکیب آن زبان، این رشته‌ها را تهیه کرد. سرعنوانهای موضوعی کتابخانه‌کنگره تا حد زیادی یک زبان نمایه‌سازی شمارشی است و «پرسی» یک زبان نمایه‌سازی ساختگی است. در عین حال باید خاطر نشان ساخت که هردو زبان از زبانهای پیش‌همارا محسوب می‌شوند.

در بررسی بیشتر امتیازات و محسنات زبانهای پیش‌همارا و پس‌همارا اختلافات بین زبانهای نمایه‌سازی ساختگی و شمارشی بیشتر مشخص می‌گردد. تفکیک و تفاوت این دو زبان امری تازه نیست و حدود سال

19. coded

21. enumerative index language

20. synthetic index language

۱۹۱۱ توسط جولوس کایزر^{۲۲} واضح نمایه‌سازی سیستماتیک نیز ارائه شده بود. (۱۱) معه‌ذا اهمیت تفکیک دوزبان ساختگی و شمارشی تا هنگامی که رانگاناتان آن را مورد بررسی قرار داد کاملاً درک نشده بود. داستان از این قرار بود که رانگاناتان هنگام کار با طرح رده‌بندی شمارشی دچار زحمت شد و دریافت که چنین طرح رده‌بندی قادر به انطباق خود با تغییرات دانش نیست. در این ایام، وی در فروشگاهی چشمش به اسباب‌بازی «مکانو»^{۲۳} افتاد که تجربه بزرگ و بینشی عمیق برای وی به دنبال داشت. (۱۲) رانگاناتان با دیدن این اسباب‌بازی دریافت که با تعدادی نوارهای فلزی شکاف‌دار، چند چرخ کوچک، چند میله فلزی، چند پیچ و مهره و چند تکه سیم و نخ، و ترکیب آنها به روشهای گوناگون، به سادگی می‌تواند وسایلی با اشکال، اندازه‌ها و اهداف متفاوت ساخت. وی با خود چنین گفت که اگر بتوان با چند پیچ و مهره و غیره به چنین ساختارهایی دست یافت چرا با زبان نتوان این عمل را انجام داد. وی دریافت که از طریق مجموعه‌ای از واژه‌های ساده و حروف ربط می‌توان به اندیشه‌های پیچیده و ترکیباتی در زبان دست یافت. بدین طریق رده‌بندی کولن که یک رده‌بندی تحلیلی - ساختگی است متولد شد. زبانهای ساختگی (خواه برای رده‌بندی باشند یا برای نمایه‌سازی) به سمت کوچکتر شدن حرکت می‌کنند و بالقوه از قدرت بیان بیشتری برخوردار هستند تا زبانهای شمارشی. رده‌بندی کولن از نظر اندازه، درحکم بخش کوچکی از رده‌بندی کنگره به حساب می‌آید که تا حد زیادی رده‌بندی شمارشی است. در همان زمان و از آنجایی که رده‌بندی کولن کل دانش را به بخشهای هسته‌ای و اصلی تجزیه می‌کرد از صراحت بیشتری، برای ارائه دقیق معنا، برخوردار بود و قادر بود تا از طریق واژگان اصلی و قوانین ترکیبی تمام

22. Julius Kaiser

۲۳. مکانو (meccano) اسباب‌بازی کودکان است، مرکب از تعدادی از قطعات کوچک که کودک می‌تواند با آنها اشکال مختلفی را بسازد. - م.

موضوعهای گذشته، حال و آینده را پیش‌بینی کند. اکنون با توجه به آنچه که گفته شد درمی‌یابیم که تفکیک زبانهای ساختگی و شمارشی و بررسی امتیازات هریک برای مقایسه زبانهای نمایه‌سازی پیش‌همارا و پس‌همارا از اهمیت بسزایی برخوردار است. اشتباه توب در این بود که در ارزیابیهای اولیه خود این تفکیک را قایل نشد و زبانهای پس‌همارا را معادل زبانهای ساختگی در نظر گرفت.

استفاده از متن یکی دیگر از خصوصیات مشترک بسیاری از زبانهای نمایه‌سازی رشته‌ای قدیمی و جدید است. در سال ۱۹۵۹ هانس پیتروهن^{۲۴} واضع نمایه‌سازی کوئیک^{۲۵} ارزش به‌کارگیری متن در نمایه‌سازی را آشکار ساخت. (۱۳) در نمایه‌سازی به روش کوئیک کلید واژه‌های عنوانها به وسیله کامپیوتر انتخاب می‌شوند. کامپیوتر برای شناسایی واژه‌های خاص که به‌وسیله خط فاصله از یکدیگر جدا شده‌اند مشکلی ندارد، بلکه مشکل در به‌کارگیری واژه‌های واحد به عنوان اصطلاحات نمایه‌ای است. زیرا در این حالت واژگانی که دارای معانی مبهم و نامعلوم و یا متعدد هستند ایجاد اشکال می‌کنند. برای رفع این مشکل لوهن کلید واژه‌های داخل متن را در مدنظر می‌گیرد که در عنوانها به‌کار گرفته شده‌اند.

فرج^{۲۶} با ارائه این نظریه که واژگان به تنهایی معنایی ندارند و تنها در بافت جمله است که معنای واقعی خود را پیدا می‌کنند، صاحب اعتباری در این زمینه شده است. (۱۴) اگر چنین نظریه‌ای صحیح و مورد قبول باشد، پس نمایه‌های آخر کتاب که از واژه‌هایی تنها ساخته شده، کاربردی نخواهد داشت، زیرا در چنین وضعیتی واژه‌ها به صورت منفرد به‌کار برده می‌شوند و طبق نظریه بالا بی‌معنی خواهند بود. در واقع تفاوت واژه‌ها به عوامل زیر بستگی دارد:

اولاً، به درجه معنایی که برای آنها به‌طور ثابت در نظر گرفته شده

24. Hans Peter Luhn

26. Frege

25. Keyword in Context (KWIC)

است.

ثانیاً، به درجه متغیر بودن معنای آنها.

ثالثاً، بستگی به متنی دارد که واژه‌ها در آنها به کار گرفته می‌شود. معمولاً در استانداردهای تهیه نمایه‌ها و تزاروسها پارهای از واژه‌ها را مستثنی می‌کند، زیرا آنها به تنهایی قادر به رساندن مفهوم و معنای مورد نظر نیستند. صفات و اسامی عام جزء این دسته از واژه‌ها هستند. برای مثال دو لغت «عمومی»^{۲۷} یا «آزادی»^{۲۸} را می‌توان ذکر کرد. استفاده کننده‌ای که در جستجوی اطلاعاتی در ارتباط با واژه «آزادی» است ممکن است آزادی سیاسی، آزادی کلام، آزادی مطبوعات، آزادی اراده و غیره را در نظر داشته باشد. می‌بینیم که در چنین حالتی امکانات متعدد و بیشماری وجود دارد و لغت تنها دریافت جمله است که معنای اصلی خود را آشکار می‌سازد. در یک زبان طبیعی، مثل انگلیسی، با قراردادن واژه‌ها دریافت جمله ابهام در معنی واژه از بین می‌رود. در یک زبان نمایه‌سازی هم‌ارز چنین عملی با قراردادن لغت در رشته‌ای از واژه‌های موضوعی امکان‌پذیر است.

همان‌طور که قبلاً ذکر کردیم معادل بودن رشته واژه‌ها با مضمون متن خصوصیت بسیاری از زبانهای نمایه‌سازی رشته‌ای است. معادل بودن شکلی از اخص بودن است. کاتر^{۲۹} برای نخستین بار مفهوم اخص بودن را ارائه کرد که در واقع به معنای مدخل موضوعی مستقیم است که نقطه مقابل مدخل غیرمستقیم است و از طریق فهرستهای رده‌ای به وجود می‌آید (۱۵). اخص بودن می‌تواند مفاهیم دیگری نیز داشته باشد. به هر حال این سؤال یعنی «اخص بودن چگونه اخص تلقی می‌شود؟»^{۳۰} سالها مطرح بوده و پاسخهای متفاوتی به آن داده شده است. یکی از پاسخهایی که توسط جولی

27. general

28. freedom

29. Cutter

30. "How specific is specific"

پتی^{۳۱} ارائه شده چنین است: چون وظیفه اخص بودن تنظیم اندازه فایل است (۱۶)، یک سرعنوان یا رشته را وقتی می‌توان به اندازه کافی اخص دانست که سرعنوانهای زیادی به آن نسبت داده نشود، با چنین تعریفی هدف اخص بودن را می‌توان چنین بیان کرد:

— تنظیم پهنا و دامنه^{۳۲} نمایه‌سازی.

— فایل‌های بزرگ را قابل تورق^{۳۳} یا چشم انداختن نماید.

پاسخ دیگر به سؤال «اخص بودن چگونه اخص تلقی می‌شود» «معادل بودن سرعنوان با مضمون متن»^{۳۴} است که مفهومی بریتانیایی برای اصطلاح اخص بودن است. (۱۷ و ۱۸) سرعنوانهای معادل (یا به تعبیر امریکایی آن سرعنوانهای اخص) سرعنوانهایی درست مطابق اندازه هستند، بدین معنا که تمامی موضوعهای اسناد موردنظر را به صورتی مختصر و فشرده ارائه می‌کنند. در واقع هدف از تعریف اخص بودن بدین شکل، ارائه چکیده‌ای کوتاه^{۳۵} از یک سند است تا بتوان بدین طریق اسناد مرتبط با یکدیگر را با سهولت انتخاب کرد.

پیش از پرداختن به موضوع اصلی مقاله، نکته مهم تاریخی دیگری که اخیراً بموقع پیوسته قابل ذکر است. در بهار سال ۱۹۹۱ با تشویق و حمایت مالی «شورای منابع کتابخانه» کنفرانسی برای بررسی عملکرد تقسیمات فرعی سرعنوانهای موضوعی کتابخانه کنگره برگزار گردید. (۱۹) درك عموم افراد این بود که گردآوری و ساخت رشته‌های واژه‌ای موضوعی پیش‌همارا در کتابخانه کنگره امر شاق، پیچیده و پرهزینه است. همین مطلب انگیزه‌ای برای بررسی مجدد تقسیمات فرعی گردید.

یکی از موضوعات مورد بحث در این کنفرانس ارائه به‌کارگیری و توسعه تقسیمات فرعی شناور^{۳۶} توسط کتابخانه کنگره بود. زبان

31. Julire Pettec

32. indexing breadth

33. browsable

34. Coextensive

35. miniabstract

36. free-floating subdivision

سرعنوانهای موضوعی کتابخانه کنگره تا حدودی ساختگی است، زیرا از تقسیمات فرعی شناور استفاده می‌کند و تا حدودی شمارشی در نظر گرفته می‌شود، زیرا تقسیمات فرعی به کار برده شده در آن به وسیله کتابخانه کنگره تهیه و ویرایش شده و برای هر کدام شماره مورد نظر اختصاص داده می‌شود. دفاع از ادامه به کارگیری تقسیمات فرعی شناور شامل دو جنبه می‌شد: اولاً دفاع از یک زبان نمایه‌سازی ساختگی، ثانیاً دفاع از یک زبان نمایه‌سازی پیش‌همارا. دفاع از جنبه پیش‌همارایی این زبان به دلایل زیر مطرح شد:

۱- به استفاده‌کنندگان کمک می‌کند تا نحوه محدود کردن دامنه جستجو را دریابند.

۲- انجام هدف دقیق و اصلی فهرست.

۳- ساخت و ارائه فایل‌های بزرگ قابل تورق.

۴- تهیه نمایه‌ای برای رده‌بندی کتابخانه کنگره. (۲۰)

در این کنفرانس شاید هیچ‌کس با جنبه پیش‌همارایی زبان سرعنوانهای موضوعی کتابخانه کنگره مخالفتی نداشت. موضوعهای مورد علاقه خیلی اختصاصی بود، مثلاً این سؤال مطرح بود که آیا تقسیمات فرعی جغرافیایی و شکلی را بایستی در رشته‌های پیش‌همارا به کار گرفت؟ یا اینکه آنها را بیرون کشیده و درجایی دیگر مثل پیشینه «مارک» قرار داد. مشکل دیگری که با آن مواجه بودند اصطلاحات جغرافیایی جدا از هم بود، زیرا ارتباط معقولی بین این اصطلاحات وجود ندارد و این خود باعث می‌شود تا تعداد ریزشهای کاذب^{۳۷} افزایش یافته و از دقت بازیابی کاسته شود. در این زمینه توصیه شد تا امکان تهیه یک کد فرعی جداگانه برای اصطلاحات شکلی مورد بررسی قرار گیرد. (۲۱)

موضوع دیگری که در کنفرانس مورد بحث قرار گرفت این بود که آیا ترتیب تقسیمات فرعی در یک سرعنوان موضوعی بایستی به شکل

معمول آن یعنی تقسیمات فرعی موضوعی، جغرافیایی، زمانی و شکلی باشد یا ترتیب دیگری را می‌توان به کار برد. افراد شرکت کننده در کنفرانس عقیده داشتند که به کارگیری ترتیب معمول برای ساخت یک رشته واژگان موضوعی باعث می‌شود تا فهرستنویسی موضوعی صحیح‌تر و ارزانتر انجام شود. به دلیل کم بودن وقت کنفرانس، پیشنهاد فوق بیشتر مورد بحث قرار نگرفت و بررسی آن به انجمن کتابداران امریکا محول گردید. (۲۲)

توصیه دیگری که در این کنفرانس ارائه شد این بود که آیا لازم است تا علاوه بر تهیه ترکیب‌هایی از «سرعنوانهای موضوعی، تقسیمات فرعی موضوعی» یک فایل مستند موضوعی ملی تهیه گردد که حاوی پیشینه‌های سرعنوانهای موضوعی نیز باشد. ادله مخالفان این بود که تهیه چنین پیشینه‌هایی نیازمند ایجاد فایل‌های خیلی بزرگ است. مثالی که آنها ارائه کردند این بود که در صورتی که در این فایل به جای سرعنوانهای موضوعی از سرعنوانهای اسامی استفاده شود، کاری مشابه لازم است. در اینجا موضوع را بیشتر مورد بررسی قرار می‌دهیم. نظام سرعنوانهای موضوعی کتابخانه کنگره دارای ۱۱۴۰۰ سرعنوان اسامی است و مجاز هستیم که زیر هر سرعنوان اسم، ده تقسیم فرعی ایجاد کنیم. با ترکیب این سرعنوانهای اسامی و تقسیمات فرعی چیزی حدود ۲/۰۵ میلیون سرعنوان موضوعی پیش‌همارا به وجود می‌آید که از چنین الگویی پیروی می‌کند: «اسم - تقسیم فرعی». (۲۳) برای تهیه چنین فایل مستند و بزرگی استدلال این است که این فایل فهرستنویسان را قادر می‌سازد تا سرعنوانها را از فایل مستند بیرون بکشند و به پیشینه کتابشناختی که در دست دارند، الحاق نمایند. با انجام چنین عملی احتمالاً دقت در تخصیص سرعنوان موضوعی افزایش می‌یابد. اخیراً با پیاده کردن رشته‌های واژگان «موضوع - سرعنوان/ موضوع - تقسیم فرعی» از پیشینه‌های کتابشناختی توانست‌اند ۱۰۰۰۰ پیشینه مستند موضوعی تهیه کنند. این عمل برای تهیه هسته یک

فایل مستند رشته‌ای در او. سی. ال. سی.^{۳۸} انجام گرفت. (۲۴)

۲

اکنون با کسب زمینه تاریخی پیش‌همارایی، به موضوع و سؤال اصلی باز می‌گردیم، یعنی «پیش‌همارایی، آری یا خیر؟» این موضوع تنها در جایی قابل طرح است که دارای یک نظام کامل بازیابی پیوسته باشد و می‌توان از طریق این نظام کلید واژه‌های رشته‌های واژگان موضوعی را مورد جستجو قرار داد. همچنین فرض بر این است که امتیازات واژگان کنترل شده مورد سؤال قرار نمی‌گیرد. در ادامه بحث شانزده معیار پیشنهادی «توب» برای ارزیابی پیش‌همارایی مورد بحث قرار می‌گیرد، در ضمن برای بررسی عمیق‌تر من نیز چهار معیار به آن اضافه می‌کنم که عبارتند از: مانعیت، جامعیت (فراخوان)^{۳۹}، قابلیت تورق و چشم انداختن^{۴۰} و بالاخره توانایی ارائه تعریف با توجه به متن و زمینه موضوع که تابعیت در متن^{۴۱} باشد.

۲ - ۱. ترکیب (نحو) منطقی

«توب» معتقد بود که زبانهای نمایه‌سازی پس‌همارا به دلیل اینکه از یک ترکیب منطقی استفاده می‌کنند نسبت به زبانهای نمایه‌سازی پیش‌همارا ارجحیت دارند. در یک زبان پس‌همارا موضوعهای چند مضمونی در صورتبندیهای^{۴۲} سؤالی ارائه می‌شوند نه در زبان نمایه. در صورتبندیهای سؤالی، ترکیب بکار برده شده براساس منطق «بولی» تهیه می‌شود و معمولاً (نه همیشه) محدود به برآیند منطقی و تقاطع اصطلاحات است. در صورتی که ترکیب اکثر زبانهای پیش‌همارا از نوع دستوری است و

38. Online Computer Library Center (OCLC)

39. recall

40. browsability

41. contextuality

42. formulation

طبیعتاً پیچیده‌تر از ترکیب زبانهای پس‌هماراست. برای مثال زبان سرعنوانهای موضوعی کتابخانه کنگره یک زبان پیش‌هماراست که به قوانین ترکیبی زیادی نیازمند است. در اینجا بهتر بودن ترکیب منطقی یا دستوری مدنظر نیست، بلکه مسئله اصلی این است که کدام یک از این ترکیبها می‌تواند در دسترسی موضوعی، بهتر و مؤثرتر عمل کند. در حالی که پاسخ دادن به این سؤال خیلی مشکل است و به نظر می‌رسد که لازمه یک بازیابی خوب در ساختگی بودن زبان آن نهفته است، بالتبع یک ترکیب ساده «بولی» می‌تواند راهگشا باشد. در این باره در بخش «مانعیت» و «اخص بودن» بیشتر بحث خواهیم کرد.

۲ - ۲. سادگی^{۴۳}

یک زبان پس‌همارای بدون ترکیب یا با ترکیب کم، از یک زبان پیش‌همارا که بایستی دارای کمی ترکیب و یا حتی ترکیب پیچیده‌ای باشد، ساده‌تر است. واژگان زبان پس‌همارا ساده‌تر است، زیرا تنها از اصطلاحات تشکیل می‌شود در صورتی که واژگان زبان پیش‌همارا از اصطلاحات و رشته‌ها تشکیل شده است.

۲ - ۳. سرعت در نمایه‌سازی یا فهرستنویسی

در کل چون در زبان نمایه‌سازی پس‌همارا سرعنوانهایی برای ساخته شدن وجود ندارد و از قوانین ترکیب نیز خبری نیست تا بتوان آنها را آموخت، بنابراین عمل نمایه‌سازی و فهرستنویسی این زبان سریعتر از زبانهای پیش‌هماراست. معه‌ذا این تفاوت سرعت اهمیت زیادی ندارد. این امر را می‌توان با یک مثال روشن ساخت. در زبانهای نمایه‌سازی رشته‌ای مثل «پرسی» که در ترکیب به‌صورت الگوریتمهای کامپیوتری وارد نظام می‌شود، زمانی که برای ساختن و تهیه سرشناسه‌ها صرف می‌شود خیلی

اهمیت ندارد زیرا این کامپیوتر است که عمل ساختن و تهیه را انجام می‌دهد نه فهرست‌نویس. (۲۵)

در اکثر چنین نظام‌هایی لازم است تا فهرست‌نویس اصطلاحات را به‌طریقی گدگذاری کند تا نشانگر چهریزه‌ها، عملگرها یا قواعد دستوری زبان نمایه‌سازی باشد. واضح است که انجام چنین کاری احتیاج به کمی وقت دارد که در واقع زمان نسبتاً کوتاهی است.

۲ - ۴. اندازه

زبانهای پس‌همارا دارای واژگانه‌های کوچکتری نسبت به زبانهای پیش‌همارا هستند، زیرا تمامی ارتباطات اصطلاحات، خارج از واژگانه‌ها ساخته می‌شود. در واقع واژگانه‌های زبان پس‌همارا تنها از اصطلاحات ساخته شده و در آنها رشته‌ها وجود ندارد. «توب» و همکارانش در تجارب بنگاه اطلاع‌رسانی فنی نیروهای مسلح امریکا دو زبان استاندارد موضوعی را به نظام‌های تک‌واژهای تبدیل کردند. وی ادعا کرد که می‌توان نظام‌های نمایه‌سازی استاندارد را ۵۰ تا ۹۰ درصد از نظر حجم کاهش داد. (۲۶) یک واژگان کوچک دارای محسنات زیر است: با صرفه است، در شروع، کار کمتری را می‌طلبد و بالاخره نگهداری و کار با آن آسانتر است.

هنگام کنترل اسنادها، مشکلات کار با واژگانی که علاوه بر اصطلاحات دارای رشته‌های واژگان نیز هست، مشخص می‌گردد. تهیه پیشینه‌های مستند نه تنها برای اصطلاحات واژگان اصلی لازم است (مثلاً برای بلوکهای ساخته شده با استفاده از یک زبان نمایه‌سازی همارا)، بلکه برای هر رشته‌ای که از طریق ارتباط دادن اصطلاحات و برطبق قوانین ترکیب زبان مورد نظر به وجود می‌آید، تهیه یک پیشینه مستند نیز لازم است، با توجه به این مورد بدون شک در می‌یابیم که تعداد پیشینه‌های مستند لازم برای یک زبان پیش‌همارا خیلی بیشتر از یک زبان

پس‌هماراست. از طرف دیگر تصمیم‌گیری درباره دو مورد زیر بستگی به اقتصادی بودن اندازه واژگان زبان مورد نظر دارد:

اولاً، تصمیم‌گیری درباره اینکه زبان پیش‌همارا بایستی تا حد زیادی ساختگی باشد.

ثانیاً، تصمیم‌گیری درباره اینکه پیشینه‌های مستند تنها برای اصطلاحات ساخته شود نه رشته‌ها.

۲-۵. میزان رشد

در اینجا منظور، میزان رشد اندازه واژگان است. هرچه سرعت رشد یک واژگان آهسته‌تر باشد نگهداری آن آسانتر و با صرفه‌تر است. سرعت افزایش اندازه واژگان اصلی، نسبت به ایجاد رشته‌های جدید ضرورتاً آهسته‌تر است زیرا این رشته‌ها خود از ترکیب اصطلاحات داخل واژگان بوجود می‌آید. بنابراین، در حالی که میزان رشد واژگان یک زبان پس‌همارا ممکن است تقریباً مساوی میزان رشد هسته واژگان یک زبان پیش‌همارا باشد، لیکن در مقایسه با کل واژگان یک زبان پیش‌همارا، سرعت رشد واژگان پس‌همارا خیلی آهسته‌تر از پیش‌هماراست، زیرا واژگان پیش‌همارا علاوه بر اصطلاحات حاوی رشته‌ها نیز می‌باشد.

۲-۶. میزان منسوخ شدن^{۴۴}

همانند معیار قبلی یعنی میزان رشد، میزان منسوخ شدن نیز تابعی از اندازه واژگان است. واژگانی که خیلی ساده و تنها از اصطلاحات تشکیل شده (برداشت رانگانان^{۴۵}) از نوارهای فلزی شکافدار بازی مکانو) خیلی دیرتر منسوخ می‌شود تا واژگانی که علاوه بر اصطلاحات از رشته‌ها نیز تشکیل شده است. این رشته‌ها بر طبق قوانین ترکیب زبان نمایه‌سازی بوجود آمده‌اند.

۴۴. به بخش تاریخچه مراجعه شود.

۲ - ۷. مناسب بودن برای اشاعه در هم کردی^{۴۶}

با فرض اینکه زبانهای نمایه‌سازی پیش‌همارا در زمینه فهرستهای پیوسته مورد مقایسه قرار می‌گیرند، دلیلی وجود ندارد که بررسی کنیم کدام‌یک از این زبانها از نظر اشاعه فهرستنویسی و نمایه‌سازی بهتر و سریعتر عمل می‌کنند. زیرا در هر دو مورد تعداد مساوی پیشینه‌های کتابشناختی اشاعه می‌یابند.

معهدا هنگامی که اشاعه ابزار یک زبان نمایه‌سازی (اصطلاحنامه، فایل مستند، دستنامه نمایه‌سازی) مورد بررسی قرار می‌گیرد، آنگاه می‌توان از نظر سهولت اشاعه بین دو زبان تفاوتی قائل شد. به دلیل کوچک بودن واژگان و نداشتن ترکیب در زبان نمایه‌سازی پس‌همارا ابزارهای آن مثل اصطلاحنامه، فایل مستند و دستنامه نمایه‌سازی فشرده‌تر هستند و بدین طریق اشاعه آنها سهل‌تر است.

باز هم خاطرنشان می‌سازیم که در مواردی این تفاوتها زیاد اهمیت ندارد. مثلاً یک زبان پیش‌همارا را در نظر بگیرید که اصطلاحنامه و فایل مستند آن تنها از اصطلاحات پایه و اصلی واژگان به‌وجود آمده و ترکیب آن ماشینی شده است، در این حالت نیز ابزارهای آن کوچک و اشاعه آن آسان است. بررسی مناسب بودن اشاعه درهم کردی دو زبان نمایه‌سازی پیش‌همارا و پس‌همارا، با مقایسه تفاوت‌های «شمارشی - ساختگی» این دو زبان عملی‌تر است.

۲ - ۸. هزینه

با توجه به اینکه سرعت نمایه‌سازی و فهرستنویسی یک زبان پس‌همارا سریعتر است، بنابراین استفاده از این زبان برای سازمانهای فهرستنویسی ارزانتر تمام می‌شود. در واقع به دلیل اینکه زبان نمایه‌سازی پیش‌همارا تا حد زیادی شمارشی است و دارای قوانین ترکیب پیچیده است، تهیه فهرست

و نمایه با استفاده از این زبان گران تمام خواهد شد. از طرف دیگر اگر زبان نمایه‌سازی پیش‌همارا تا حد زیادی ساختگی باشد و از یک ترکیب ساده ماشینی استفاده نماید باز هم صرفه‌جویی مالی زیاد نخواهد بود.

۲ - ۹. قابلیت عمومی شدن^{۴۷}

منظور «توب» از قابلیت عمومی شدن حوزه عملکرد و مقبولیت^{۴۸} یک زبان نمایه‌سازی است. (۲۷) «توب» معتقد بود که نظام تکواژه‌ای پس‌همارای او دارای مقبولیت بیشتری است تا یک نظام پس‌همارای دیگر. دلیل وی این بود که نظام تکواژه‌ای می‌تواند از طریق همارایی تکواژه‌ها، موضوعهایی را که هنوز کشف نشده‌اند با نظام منطبق سازد. در اینجا باز شاهد هستیم که «توب» بر طبق عادت، محسنات یک زبان پس‌همارا را بار زبانهای پیش‌همارای ساختگی یکسان در نظر گرفته و دچار مغلطه شده است. هنگامی که «رانگانانان» رده‌بندی کولن را به عنوان یک نظام بویا مطرح و ادعا کرد که دارای مقبولیت زیاد جهانی است دچار اشتباه «توب» شده بود. در واقع می‌توان گفت که یک زبان پیش‌همارا با ترکیب صریح‌تر، مقبولیت بیشتری دارد تا یک زبان پس‌همارا که تنها روابط بین اصطلاحات را به صورت عملکردهای بولی بیان می‌کند.

۲ - ۱۰. بیطرفی (خنثی بودن)^{۴۹}

منظور «توب» از بیطرفی، توانایی یک زبان نمایه‌سازی است که بتواند بدون ایجاد رقابت و تضاد در دیگر زبانهای نمایه‌سازی حل شده یا کامل گردد. (۲۸) به عنوان مثال «توب» دو زبان نمایه‌سازی خنثی را مورد استناد قرار داد که اولی از سرعنوانهای غیر مقلوب و دومی از سرعنوانهای مقلوب^{۵۰}

47. universality

48. hospitality

49. neutrality

50. inverted headings

بهره می‌جست. وی چنین استدلال کرد که یک زبان پس‌همارای بدون ترکیب، لزوماً از هرزبان پیش‌همارا خنثی‌تر است. معهذا باید اذعان داشت که در چنین استدلالی زبانهای پیش‌همارای ساختگی در نظر گرفته نشده‌اند. در چنین زبانهایی واژگان اصلی (پایه) می‌تواند ساده‌تر و کوچکتر از واژگان هر زبان پس‌همارا باشد. اگر قرار باشد که خنثی بودن را براساس سادگی یا کامل بودن واژگان پایه یک زبان نمایه‌سازی تعریف کنیم، آنگاه از این نظر تفاوتی بین زبانهای پیش‌همارا و پس‌همارا وجود ندارد.

۲ - ۱۱. اخص بودن

«توب» در این رابطه عقیده زیر را ارائه کرد: «اخص بودن که از طریق تقاطع، همارایی یا استنتاج منطقی اصطلاحات عام به‌دست می‌آید، با اخص بودن که از طریق ترکیب اصطلاحات و برطبق ترکیب زبان نمایه‌سازی به‌وجود می‌آید، در اصل متفاوت هستند. اخص بودن از نوع دوم را می‌توان در سرعنوان یا رشته‌واره‌ها مشاهده کرد.» وی معتقد بود که اخص بودن سرعنوانهایی که در زبان سرعنوان موضوعی کتابخانه کنگره قابل دسترسی است، بارزتر است. وی چنین استدلال کرد که موضوعهای علمی مجموعه‌ای پیچیده از چند ایده کلی هستند و بیان آنها به‌صورت رشته‌های بولی خیلی ساده‌تر است تا اینکه آنها را به‌صورت رشته‌های زبان سرعنوان موضوعی کتابخانه کنگره نشان دهیم. زیرا در زبان سرعنوان موضوعی کتابخانه کنگره تقسیمات فرعی تنها نشانگر اختصاصات و جنبه‌های یک ایده کلی و واحد هستند. (۲۹) عدم توانایی زبانهای پس‌همارا در اخص کردن روابط پیچیده‌تر و غیرقابل بیان بودن این پیچیدگیها به‌صورت عملگرهای بولی، یکی از عمده‌ترین انتقاداتی است که به زبانهای پس‌همارا وارد شده است. همان‌طور که مشخص شد این ایده کاملاً برخلاف برداشت توب است. در یک جستجوی بولی که در آن از عملگر «و» برای ارتباط اصطلاحات استفاده شده، هنگام بازیابی دو اصطلاح «فلسفه» و «تاریخ»، علاوه بر اسنادهایی با عنوان

«فلسفه تاریخ»، اسنادهای «تاریخ فلسفه» نیز بازیابی می‌شوند.^{۵۱} این درحالی است که اکثریت زبانهای پیش‌همارا قادر به تفکیک چنین اسنادها و در نتیجه تفکیک اسناد هستند. این توانایی زبانهای پیش‌همارا به دلیل پیچیدگی ترکیب (نحو) آنهاست. مثلاً زبان «پرسی» دارای آن چنان ترکیب (نحو) صریحی است که می‌تواند اسنادی با عنوانهای زیر را از یکدیگر تفکیک کند: «گزیده شدن بچه‌ها توسط سگها»، «گزیده شدن سگها توسط بچه‌ها».

۲ - ۱۲. قابلیت تورق و چشم انداختن

«توب» قابلیت تورق را جزء معیارهای ارزشیابی زبانهای نمایه‌سازی قرار نداد، شاید به دلیل اینکه نظام تکیواژهای او از چنین خاصیتی برخوردار نیست. همان‌طور که قبلاً تذکر داده شد، اصل اخص بودن به این دلیل در نمایه‌سازی مطرح شده، تا از انتساب موضوعهای متعدد به یک سرعنوان جلوگیری شود. در تاریخ دوم ماه مه ۱۹۹۳ با استفاده از سیستم اوربون در دانشگاه لس‌آنجلس در کالیفرنیا^{۵۲} جستجویی درباره سرعنوانهای موضوعی که دارای واژه «هنر» هستند انجام گرفت، در نتیجه چیزی حدود ۴۰۰۰۰ پیشینه کتابشناختی بازیابی شد. کمتر استفاده‌کننده‌ای قادر است، برای یافتن پاسخ خود به ترتیب این ۴۰۰۰۰ پیشینه را مورد بررسی قرار دهد. بنابراین لازم است تا از طریقی مناسب فایل را تقسیم‌بندی (پارتیشن‌بندی) کنیم. در چنین وضعیتی مقوله هنر بایستی هرچه بیشتر اخص گردد. اکثر جستجوگران از این شکایت دارند که در هنگام جستجوی موضوعی، اطلاعات بسیار زیادی بازیابی می‌شود و آنها نمی‌دانند که چگونه دامنه جستجو را محدود کنند. این وضعیت حتی هنگامی که اندازه بازیابی از طریق پیش‌همارایی بولی کوچکتر شده، باز هم صادق است. سرعنوانهای پیش‌همارایی شده که موضوعهای بزرگی همچون

۵۱. منظور ریزش کاذب - بخصوص در نظام تکیواژهای - است. - م.

«هنر» را تقسیم‌بندی کرده‌اند، دارای قابلیت تورق مناسبی هستند و راههای جستجوی مؤثر را به استفاده‌کنندگان ارائه می‌کنند.

۲- ۱۳. مانعیت

توانایی یک زبان نمایه‌سازی در عدم انتخاب اسناد غیر مرتبط، در بازیابی مانعیت نامیده می‌شود. در واقع درصد اسناد مرتبط با موضوع که بازیابی شده‌اند نشانگر مانعیت زبان نمایه‌سازی است. به دلایلی که در معیار اخص بودن بحث شد، زبانهای پیش‌همارا از نظر مانعیت در بازیابی فهرستها از زبانهای پس‌همارا بهتر هستند. این دلایل عبارتند از: اولاً، به دلیل کلی بودن روابط بولی، زبانهای پس‌همارا دارای ریزشهای کاذب زیادتری هستند. نظامهایی مثل نمایه‌سازی رابطهای فاراندین^{۵۳} و نقش‌نماها^{۵۴} و اتصالاتی^{۵۵} که دانشگاه وسترن ریزرو^{۵۶} تهیه کرده‌اند، در واقع برای جبران کاستیهای نظامهای پس‌همارا به‌وجود آمده‌اند.

ثانیاً، جستجوی پیشینه‌های کتابشناختی که دارای اصطلاحات عام و کلی هستند، حتی با ترکیب بولی، اغلب منجر به بازیابی تعداد زیادی اسناد می‌شود. استراتژی جستجوی بهتر به این شرح است:

— جستجوی کلید واژهای را بر روی رشته‌های موضوعی پیش‌همارا (در فایل مستند) آنقدر ادامه می‌دهیم تا آنها بیابیم.

— رشته‌ها را آنقدر تورق می‌کنیم تا موضوع اخص را بیابیم.

— درخواست می‌کنیم تا تمامی اسنادی که تحت این موضوع اخص نمایه‌سازی شده‌اند ارائه شود.

با به کارگیری چنین استراتژی، می‌توان به مانعیت زیادی دست یافت، خصوص هنگامی که نمایه‌سازی معادل^{۵۷} (هم‌افزون) به کار برده شده باشد.

53. Farrandane's relational indexing

54. roles indicators

55. links

56. Western Reserve University

57. Coextensive indexing

مرحله دوم این استراتژی (تورق رشته‌های موضوعی)، در یک نظام کاملاً پس‌همارا قابل استفاده نیست و بدین طریق این زبان مانعیت خود را از دست می‌دهد.

۲- ۱۴. جامعیت

یکی از اهداف کاتر از فهرستنویسی، گرد آوردن و کنار هم قرار دادن تمامی اسنادی است که دارای موضوع معینی هستند. (۳۰) جامعیت مرحله عملی این کار است. در واقع یک زبان نمایه‌سازی هنگامی دارای جامعیت مناسب است که بتواند تمامی اسناد مرتبط با موضوع سؤال را بازیابی نماید. همان‌طور که قبلاً ذکر شد، کرانفیلد در تجربه اول خود با توجه به عامل جامعیت، زبانهای پیش‌همارا و پس‌همارا را مقایسه کرد، ولی نتایج معینی به دست نیاورد. در واقع حتی اگر کرانفیلد به نتایجی هم می‌رسید، این نتایج در حال حاضر کاربردی نداشت. زیرا در آن زمان، بازیابی، عملی یک مرحله‌ای بود. در حالی که استراتژی بازیابی در حال حاضر چند مرحله‌ای است. یکی از یافته‌های کرانفیلد این بود که هرچه مانعیت نظامی افزایش یابد از قدرت جامعیت آن کاسته می‌شود و برعکس. اما مجدداً این فرض که به «فرض مبادله‌ای»^{۵۸} معروف شده بود با زمینه‌ای که بازیابی را یک مرحله‌ای در نظر می‌گرفت فرمول‌بندی و ارائه شد. با وجود فهرستهای پیوسته و مدرن کنونی که امکان اتخاذ استراتژیهای متعدد بازیابی را مهیا می‌سازد و همگی آنها مستلزم یک مرحله میانی یعنی تورق رشته‌های موضوعی هستند، معلوم نیست که این فرض بتواند جایگاهی در بازیابی نوین داشته باشد. از آنجایی که زبانهای پیش‌همارا سهم زیادی در کسب دقت در بازیابی داشتند، لزومی ندارد تا کاهش جامعیت را به دلیل افزایش دقت مورد بررسی قرار دهیم. در واقع آنچه که باعث می‌شود یک زبان از نظر جامعیت از دیگری برتر باشد ساختار ترکیبی زبانهای پیش‌همارا و پس‌همارا نیست، بلکه ساختار معنا

۵۸. trade off hypothesis، در اینجا منظور، مبادله «مانعیت» در برابر «جامعیت» است. - م.

شناسی^{۵۹} آنها مدنظر است (یادآور می‌شویم که تفاوت ساختار ترکیبی، مهمترین وجه تمایز زبانهای پیش‌همارا و پس‌هماراست). باوجود یک فهرست پیوسته که از جستجوی بولی رشته‌های پیش‌همارا پشتیبانی می‌کند، دلیلی ندارد که ادعا کنیم زبانهای پس‌همارا از نظر جامعیت بهتر از زبانهای پیش‌همارا هستند و برعکس.

۲ - ۱۵. تعداد نقاط بازیابی

اگر در زبانهای پس‌همارا ابزارهای کار محدود به فهرستهای دستی باشد، آنگاه تعداد متوسط نقاط بازیابی یک سند از اهمیت زیادی برخوردار است. زیرا در فهرستهای دستی تنها لغت مقدم^{۶۰} در یک توصیف‌گر یا رشته موضوعی را می‌توان مستقیماً مورد جستجو قرار داد. یکی از امتیازات هر سیستمی که دارای فهرستهای پیوسته است تعداد نقاط بازیابی است. زیرا در چنین نظامی جستجوی کلید واژه‌های رشته‌های موضوعی امکان‌پذیر است. اصطلاحات درون رشته‌های موضوعی زبانهای پیش‌همارا تنها در یک فیلد با یکدیگر مرتبط هستند، در حالی که در یک زبان پس‌همارا اصطلاحات به‌صورت منفرد و جداگانه با فیلد مربوطه ارتباط دارند.

۲ - ۱۶. سرعت جستجو

یکی از راههای سنجش سرعت جستجو، اندازه‌گیری زمان لازم برای یافتن یک تعداد عنوانهای مرتبط با یکدیگر است. در هنگام جستجو با یک نظام بولی که در آن از زبان پس‌همارا استفاده شده، آنچه که بازیابی می‌شود مجموعه‌ای از استندهاست. عنوانهای داخل این مجموعه باید مورد بررسی قرار گیرد تا آنهایی که مرتبط هستند پیدا شوند. هرچه این مجموعه بزرگتر باشد زمان بازیابی افزایش می‌یابد. مثلاً برای بازیابی مقوله هنر که چیزی بالغ بر ۴۰۰۰۰ مورد بازیابی می‌شود زمان لازم خیلی زیاد است. در جستجو با نظامی

که از زبان پیش‌همارا استفاده می‌کند، آنچه که بازیابی می‌شود بستگی به استراتژی جستجو دارد. اگر این استراتژی مرحله میانی جستجو یعنی تورق رشته‌های موضوعی را شامل شود و این مرحله بتواند با موفقیت اندازه بازیابی را کاهش دهد، آنگاه تعداد عنوانهایی که مورد بررسی قرار می‌گیرد خیلی کم خواهد شد. با استفاده از زبان پیش‌همارا در چنین استراتژی‌هایی، مرحله آخر جستجو یعنی بررسی عنوانهای بازیابی شده کوتاه می‌شود که این خود یکی از محسنات زبان پیش‌همارا است. (۳۰)

۲ - ۱۷. القاکنندگی^{۶۱}

«توب» قبول داشت که نظامهای نمایه‌سازی که از عبارات بلند و تقسیمهای فرعی موضوعی استفاده می‌کنند، از نظر القاکنندگی بر نظام تک‌واژه‌ای او برتری دارند. یکی از مهمترین استدلالهایی که در پشتیبانی از زبانهای پیش‌همارا در کنفرانس تقسیمات فرعی رده‌بندی کنگره مطرح شد این بود که این زبانها به استفاده کننده کمک می‌کند تا هنگامی که عنوانهای زیادی بازیابی می‌شود بتواند دامنه جستجو را تنگ کند و به عناوین مورد نظر دست یابد. (۳۱) استفاده کنندگان اغلب نمی‌دانند که چگونه اندازه بازیابی را کاهش دهند. ارائه نظر گاههای مختلف یا پرسپکتیوهای متفاوت از اصطلاحات مورد جستجو در ادبیات مربوطه، توسط زبانهای پیش‌همارا، یکی از دلایل بروز این مشکل است. به دلیل اینکه رشته‌های پیش‌همارا تداعی معنایی را می‌توانند القا کنند، بنابراین لازم است تا استفاده کنندگان از رشته‌های واژه‌ها، تصویری ذهنی از اصطلاحات داشته باشند تا بتوانند معنی انتزاعی آن را دریابند. بنابراین رشته‌های پیش‌همارا می‌توانند مواردی را برای استفاده کننده اخص و ارائه کنند که قبلاً قادر به انجام آن نبوده است.

۲ - ۱۸. قابلیت ارائه مضمون اصطلاحات با توجه به متن (تابعیت در

یکی از وظایف هر واژگان کنترل شده‌ای تفکیک واژه‌های هم تلفظ^{۶۳} از یکدیگر است. اگر واژگانی نتواند چنین عملی را انجام دهد از دقت جستجو کاسته می‌شود، زیرا در بازیابی، معانی مختلف اصطلاح بازیابی می‌شود و اصل اخص بودن از بین می‌رود. راه حل رایج این مشکل، محدود کردن معنای چنین واژه‌هایی از طریق بیانگرها^{۶۴} است. در زبانهای نمایه‌سازی پس‌همارا با اضافه کردن اصطلاحات اضافی از طریق حرف ربط «و»، می‌توان تا حدودی معنای اخص را به دست آورد. همان‌طور که قبلاً نیز اشاره شد، بسیاری از واژه‌ها تا حدودی چندآوایی^{۶۵} هستند. (بعضی از فلاسفه معتقدند همه واژه‌ها چنین وضعیتی دارند.) در فهرستهای بزرگ تعداد چنین واژه‌هایی کم نیست و نمی‌توان از طریق بیانگرهای معترضه یا تقاطعهای بولی به‌طور کامل ابهام معنایی این واژه‌ها را از بین برد، در نتیجه از دقت جستجو کاسته می‌شود. در چنین مواردی رشته‌های پیش‌همارا چاره‌ساز است زیرا اصطلاحات را با توجه به متن ارائه می‌کند. از این نظر زبانهای پیش‌همارا نسبت به پس‌همارا برتری دارند.

۱۹-۲. قابلیت آشنایی^{۶۶}

«توب» هنگام ارائه نمایه‌سازی تک‌واژه‌ای تصور می‌کرد که به دلیل جدید بودن این نظام مخالفت‌هایی در به کارگیری آن، بروز خواهد کرد. در حال حاضر شاید بتوان چنین استدلال کرد که یک زبان پیش‌همارا با ترکیب ساختگی آن قابلیت آشنایی کمتری دارد تا یک واژگان توصیفگر و ساده یک زبان پس‌همارا. معه‌ذا یک واژگان کنترل شده که استفاده کننده با آن آشنایی بیشتری دارد خیلی بهتر عمل می‌کند تا یک واژگان بزرگ که از

62. contextuality

63. homonyms

64. parenthetical qualifiers

65. multivocal

66. familiarity

طریق روابط معنایی به وجود آمده است. (در اینجا فرض ما این است که هر دو زبان پیش‌همارا و پس‌همارا دارای چنین واژگانی هستند.) به عبارت دیگر قابلیت آشنایی یک زبان نمایه‌سازی بیشتر به روابط معنایی بستگی دارد تا ترکیب آن زبان. لازم به یادآوری است که از عوامل مهم تمایز زبانهای نمایه‌سازی پیش‌همارا و پس‌همارا ترکیب آن است.

۲-۲۰. سازگاری با محیطی ماشینی

همان‌طور که در ابتدای مقاله اشاره شد زبان نمایه‌سازی پس‌همارا با به‌کارگیری کامپیوتر توسعه یافت. جایگزینی ترکیب ساده و منطقی بولی به جای ترکیب پیچیده سرعنوانهای موضوعی که گاهی اوقات ناسازگار بود، از طریق به‌کارگیری کامپیوتر امکان پذیر شد. در محیطهای دستی، استفاده از زبانهای پیش‌همارا روشی استاندارد بوده و هست. اما آیا این بدان معنی نیست که زبانهای پیش‌همارا مناسب محیطهای ماشینی نیستند؟ در اینجا، پاسخ بایستی منفی باشد. در واقع و به روشهای مختلف زبانهای پیش‌همارا در محیطی ماشینی‌شده بهتر از محیطهای دستی عمل می‌کنند. زیرا در محیطی ماشینی می‌توان هر اصطلاحی در رشته موضوعی را از طریق جستجوی کلیدواژه‌ای مورد بازیابی قرار داد و در نتیجه به ارجاعهای متقابل کمتری نیاز است. در واقع بعضی از نظامهای نمایه‌سازی پیش‌همارا مثل «پرسی» با ایجاد ارجاعهای متقابل بین رشته‌ها، یک سند خاص را زیر گردشهای مختلف یک رشته، نمایه‌سازی می‌کنند. همچنین این نوع نظامها به خاطر قابلیت‌هایی که ماشین فراهم کرده در وقت صرفه جویی می‌کنند، زیرا به محض ورود اصطلاحات انتخابی برای نمایه‌سازی، این کامپیوتر است که سرعنوانهای پیش‌همارا را بر اساس ترکیب ماشینی به یکدیگر متصل (درواقع می‌بافد) و ارائه می‌کند.

به دلیل اینکه در ارزیابی دو زبان مورد نظر این بیست معیار از ارزش یکسانی برخوردار نیستند، بنابراین نمی‌توان تنها با استناد به آنها برتری یک زبان را تشخیص داد. از دیدگاه یک استفاده کننده، جنبه‌هایی که دقت ارزیابی یک زبان پیش‌همارا را تقویت می‌کنند مهمتر هستند، مثل اخص بودن، الفاکنندگی و بیانگری. زبانهای پیش‌همارا به خاطر جستجوی کلیدواژه‌های رشته‌های پیش‌همارا، بدون اینکه قدرت بازیافت را قربانی کنند، از دقت بیشتری برخوردار هستند. از نقطه نظر یک فهرست‌نویس یا نمایه ساز جنبه‌های صرفه‌جویی مالی زبانهای پس‌همارا مدنظر هستند، مثل اندازه، رشد نسبتاً کند واژگان و بالاخره نداشتن ترکیبی برای اتصال دادن اصطلاحات پیش از جستجو که خود نیاز به ساخت سرعنوانها را منتفی می‌سازد. هرچه اندازه پایگاه اطلاعاتی بزرگتر شود، هزینه‌ها و نیاز به اخص بودن سرعنوانها افزایش می‌یابد.

در مجموع شاید بتوان گفت که امتیازات پیش‌همارایی بیشتر است، ولی دسترسی به این امتیازات مستلزم هزینه‌هایی خاص است. تصمیم‌گیری در این مورد به عهده سازمانی است که نمایه‌سازی می‌کند تا ببیند این کار به صرفه است یا نه. زبانهای مختلف نمایه‌سازی پیش‌همارا از نظر هزینه‌بری خیلی فرق می‌کنند. زبانهای شمارشی و آنهایی که از ترکیب پیچیده‌ای استفاده می‌کنند خیلی هزینه‌بر هستند تا آنهایی که از یک ترکیب ساختگی یا ماشینی ساده استفاده می‌کنند. اگر تصمیم گرفته شد که از زبانهای پیش‌همارا استفاده شود (یا استفاده از آن ادامه یابد)، آنگاه این سؤال مهم پیش می‌آید که «تا چه حد و به چه طریقی اصطلاحات بایستی همارا شوند؟» یا مثلاً آیا لازم است که دقت ارزیابی زبان پیش‌همارا تا آن حد افزایش یابد که اصطلاحات جغرافیایی و تقسیمات شکلی موجود در رشته‌های موضوعی را نیز شامل گردد؟ یا اینکه می‌توان با کمی صرفه‌جویی مالی آنها را به طور جداگانه موضوع بندی کنیم؟ و یا اینکه چه متنی می‌تواند بهترین اخص بودن،

القایذیری و یا بیانگری را ارانه کند؟ نظریه پردازان موضوعی در حال حاضر مواجه با مشکل ترین، جالب ترین و پیچیده ترین مسائل بازایی اطلاعات هستند که سؤالهای فوق الذکر بیانگر بخشی از آنها هستند.

پانویسها:

1. Taub, Mortimer and Associates. *Studies in Coordinate Indexing* (Washington, D.C.: Documentation Incorporated, 1953-1957), v.3, pp.86-87.

2. *Ibid.*, v.1, p.26.

۳. چنین گفته شده که وی شرطی باید قائل می شد. با وجود اینکه «توب» عناصر واژگان زبان نمایسازی خود را «تک واژه» نامید، خیلی زود درک کرد که نمی توان تنها با استفاده از تک واژه ها زبان کاملی ساخت. اصطلاحات ترکیبی خاص مثل «صفت به علاوه اسم» بایستی به گونه ای مطرح شوند که مفاهیم را ارانه کنند. بنابراین وی مجبور شد تا بعضی از اصطلاحات ترکیبی را قبول کرده و وارد سیستم تکواژهای نماید.

۴. چند شرط لازم است: (۱) همان طور که در پانویشت شماره ۳ اشاره شد یک واژگان زبان نمایسازی پس همارا اگر بخواهد مفاهیم خاصی را بیان کند لازم است از اصطلاحات ترکیبی استفاده کند. چنین واژگانی دارای ترکیب غیرپیشرفته ای است که معمولاً از قوانینی تشکیل می شود که می توان با آنها اصطلاحات ترکیبی شرطی را ارانه کرد. (۲) تعریفی که در اینجا از یک زبان نمایسازی زنجیرهای ارانه شده با تعریفی که در متون به کار رفته است فرق می کند. برای مثال نگاه کنید به

Craven, Timothy. *String Indexing*. (Orlando, Florida: Academic Press, p.4).

(که طبق آن رشته بر اساس قوانین ترکیبی صریح و منظم و توسط یک تولید کننده رشته نمایای بوجود می آید). (۳) جهت سادگی فرض بر این است که بیان پیچیده یک زبان نمایای رشته، به دلیل ساخته شدن این رشته ها از انواع مختلف رابطهای واژگان است، این فرض از نظر تنوری بی معنی است. زیرا رشته های یک زبان نمایای همانند جمله های یک زبان طبیعی از عناصر واژگان ساخته نشده اند بلکه از عناصر خارج از واژگان بوجود آمده اند. معذراً این فرض در سطح گسترده ای مورد قبول واقع شده و باعث ساده تر شدن درک این موضوع می شود.

5. Taub, *op.cit.*, v. 2, p. 44.

6. *Ibid.*, v.1. pp. 101-107.

7. Cleverdon, C. W. *Report on the First Stage of an Investigation into the Comparative Efficiency of Indexing Systems* (Cranfield, England: College of Aeronautics, 1960).

8. Taub, *Ibid.*, v.3, p.21.

9. Mooers, Calvin, "The Indexing Language of an Information Retrieval System," *Information Retrieval Today: Papers Presented at an Institute Conducted by the Library School and the Center for Continuation Study, University of Minnesota, Sept. 19-22., 1962*. Edited by Wesley Simonton (Minneapolis, Minn: The Center, 1963, pp. 21-36. Reprinted in *Theory of Subject Analysis: A Sourcebook*. Edited by Lois Mai Chan, Phyllis A. Richmond and Elaine Svenonius. (Littleton, Colorado: Libraries Unlimited, 1985), pp. 247-261.

۱۰. در حالی که نظام COMPASS جایگزین نظام PRECIS در کتابخانه ملی بریتانیا شده، PRECIS دارای حبه‌های پیچیده‌ای است که COMPASS دارای آن نیست. PRECIS زبانی پیش‌هما را، ساختگی و دارای ترکیبی ماشینی شده است.

11. Kaiser, Julius. *Systematic Indexing* (London: J. Gibson, 1911).
12. Ranganathan, S. R. *The Colon Classification* (New Brunswick, N.J.: Graduate School of Library Service -- The State University 1965), p.15.
13. Luhn, Hans Peter. "Keyword-In-Context Index for Technical Literature (KWIC INDEX)" *Tech Report RC-127* (Yorktown Heights: IBM, 1959). Reprinted in *Theory of Subject Analysis: A Sourcebook, op. cit.*, pp. 170-180.

۱۴. این عقیده برای طرح جدید نبود، زیرا برای مثال «حی بنات» قلا آن را ارائه کرده بود ولی طرح بر آن پافشاری کرده و آن را تعمیم داد. بر اساس همین نظریه «وینشتاین» نظریه انقلابی «معنای» یک واژه را توسعه داد.

15. Cutter, Charles. *Rules for a Dictionary Catalog* (4th ed., rewritten; Washington, D.C.: Government Printing Office, 1904), p.67.
16. Pettec, Julia. *Subject Headings: The History and Theory of the Alphabetical Subject Approach to Books* (New York: H.W. Wilson, 1946, p.81.
17. Angell, Dick, "Library of Congress Subject Headings -- Review and Forecast." in: *Subject Retrieval in the Seventies: New Directions: Proceedings of an International Symposium held at the Center of Adult Education University of Maryland, College Park, May 14 to 15, 1971*. Edited by Hans Wellich and Thomas D. Wilson (Westport, Conn.: Greenwood Publishing, 1972), p. 149.
18. Coates, Eric J. *Subject Catalogues: Headings and Structure* (London: The Library Association, 1988), pp. 33 ff.
19. *The Future of Subdivisions in the Library of Congress Subject Headings System: Report From The Subject Subdivisions Conference Sponsored by The Library of Congress, May 9-12, 1991*. (Washington, D. C. Library of Congress, 1992).

20. *Ibid.*, pp. 36-73.
21. *Ibid.*, p. 9.
22. Thomas, Sarah E. *Letter to participants in The Subdivisions Conference, March 9, 1993.*
23. *The Future of Subdivisions in the Library of Congress Subject Headings System*, op. cit., p. 137.
24. Thomas, op. cit.
25. بعضی از زبانهای نمایسازی رشتهای بصورت گزاره‌های زبان طبیعی در نظر گرفته شده و وارد دستگاه می‌شوند. مثل زبان NEPIIS.
26. *Ibid.*, v.1, p. 37.
27. Taube, op. cit., v.1, p. 103.
28. *Ibid.*, v.1, p. 104.
29. *Ibid.*, v.1, p. 35.
30. Cutter, op. cit., p. 12.
31. استدلال مذکور، استدلالی نظری است. بررسی رابطه بین تعداد مراحل بازیابی و زمان بازیابی مورد نظر برای بررسی مجموعه‌ای از ارجاعات که بازیابی شده، موضوعی تجربی است و بستگی به اندازه پایگاه اطلاعاتی، عرض و عمق نمایسازی دارد.