

داده آرایی در فهرست‌های رایانه‌ای

نویسنده: افشین موسوی چلک
مدیر اطلاع رسانی
کتابخانه مجلس شورای اسلامی
mousavimajlislib.com

خط فارسی به دلیل ویژگی آن در نرم افزارهای رایانه‌ای دارای مسائل پیچیده‌تری نسبت به زبان لاتین است. این ویژگی در نظام دستی تا حدی قابل کنترل است و به صورت جدی مشکل آفرین نیست، اما در فهرست‌های رایانه‌ای، رسم الخط فارسی، مشکلات فراوانی را به وجود می‌آورد که در ذیل به آنها پرداخته می‌شود:

- حروفی هستند که به حروف پیش از خود می‌چسبند ولی به حروف پس از خود نمی‌چسبند، مانند: ا، ر، ز، د، ذ، ژ و هاء غیر ملفوظ.

- بعضی کلمات به دو صورت جدا یا سرهم نوشته می‌شوند، مانند «ایران‌شناسی» و «ایران‌شناسی».

- از مشکلات رسم الخط فارسی به کارگیری «همزه» در جایگاه‌ها و به صورت‌های مختلف می‌باشد؛ مانند: «الف مقصوره» که گاه به صورت «ی» نوشته می‌شود و گاه به صورت «آف». مثل «تقوا» و «تقوی».^(۷)

- کلماتی که به دو شکل نوشته می‌شوند مانند «تهران» و «طهران».

یکی از شیوه‌های هماهنگ سازی می‌تواند استفاده از سیاهه‌های از قبل تعبیه شده باشد. در این شیوه تمهیداتی اندیشیده می‌شود که کامپیوتر بتواند رسم الخط‌های گوناگون را به صورت واحد بخواند. این سیاهه باید از قبل برای کلمات تهیه شود و ارجاعات در آن رعایت گردد.

نرم افزار می‌تواند از نظام معادل سازی استفاده کند و این کار شبیه تهیه سیاهه از قبل تعبیه شده است و یا می‌تواند در نرم افزار تمهیداتی به کار برد تا هر بار که فهرست نویس، منبعی را فهرست می‌کند، در مقابل هر کلمه صورت دیگر کلمات را هم نوشت که این کار نیز بسیار مشکل می‌باشد.

بهترین راه حل، سیاست گذاری واحد برای یکسان سازی رسم الخط فارسی است. کار هماهنگ کردن رسم الخط فارسی باید توسط مرجعی با اختیارات لازم انجام شود و رسم الخط واحدی را تصویب کند که اجرای آن الزام آور باشد. در این هماهنگ سازی، سرهم و یا جدانویسی کلمات تعریف شود، جمع‌های فارسی یکدست گردد و اختلاف املاءها از میان برود.

اولین قواعد برگه آرای کامپیوتری در فهرست‌های رایانه‌ای در سال ۱۹۶۶ منتشر گردید. در سال ۱۹۸۰، قواعد برگه آرای ALA و قوانین برگه آرای کتابخانه کنگره آمریکا که کاملتر از قواعد برگه آرای ALA بود، تدوین گردید.^(۱) سپس قواعد برگه آرای BLAS^(۲) نشر یافت. قواعد برگه آرای BLAS که براساس اصول سازمان جهانی استاندارد (ISO) تهیه گردیده بود، در سه بخش به شرح ذیل ارائه شد^(۳):

- ۱- نظم حروف
- ۲- اصطلاحات استثنایی و حروف مرکب
- ۳- برگه آرای کردن و برگه آرای نکردن عناصر در شناسه‌ها

در نظام سنتی برگه آرای، فهرست‌ها براساس اصل آنچه که باید باشد^(۴) انجام می‌گیرد. کتابدار می‌تواند بعضی از شناسه‌ها را در محلی از فهرست که به نظرش مناسب‌تر می‌آید، قرار دهد. بر این اساس، اعداد را می‌توان به شکل حرفی در نظر بگیرد و در جای حرفی مربوطه برگه آرای نماید. اما در رایانه، مرتب کردن شناسه‌ها از اصل آنچه که هست^(۵) پیروی می‌کند. هر شناسه همان گونه که نوشته می‌شود، مرتب می‌گردد و نمی‌توان آن را طوری دیگر تلقی کرد. مثلاً عدد «۳» به صورت «۳» و «سه» به صورت «سه» مرتب می‌شود.^(۶) در رایانه هر کاراکتر (حرف، عدد، علائم نقطه گذاری و غیره) خوانده می‌شود و کامپیوتر زمانی می‌تواند «الف» و «ب» را مقایسه کند، که تشخیص دهد «الف» قبل از «ب» می‌آید. با توجه به اهمیت برگه آرای و آرایش داده‌ها در بازیابی اطلاعات در نرم افزارهای کتابخانه‌ای باید برنامه نویسی خاصی انجام گیرد تا بتوان قواعد برگه آرای را در فهرست‌های رایانه‌ای پیاده نمود.^(۶)

برگه آرای و آرایش داده‌ها در فهرست‌های رایانه‌ای فارسی

روش برگه آرای و آرایش داده‌ها در فهرست‌های فارسی، روش حرف به حرف می‌باشد که در این روش، حدی برای کلمات در نظر گرفته نمی‌شود و از ابتدای هر مدخل، حروف با یکدیگر مقایسه می‌شوند، نه کلمه.

نرم افزارهای کتابخانه‌ای و نکات قابل توجه در داده آرای

در این قسمت به نکاتی اشاره می‌شود که برنامه نویسان نرم افزارهای کتابخانه‌ای باید به آن توجه داشته باشند، چرا که این عوامل نقش بسیار مهمی در بازیابی منابع به عهده دارند.

حرف تعریف: حرف تعریف «آل» در اول کلمات عربی خوانده نمی‌شود، بنابراین اگر کتابی و یا نام نویسنده‌ای با حرف تعریف «آل» شروع شود، «آل» در داده آرای نباید در نظر گرفته شود. نرم افزارهای کتابخانه‌ای باید قابلیت داده آرای صحیح این گونه کلمات را داشته باشند و باید این امکان در اختیار فهرست نویسی قرار گیرد تا بتواند مشخص کند که حرف «آل» در اول کلمه در هنگام داده آرای خوانده شود یا نشود. مانند:

«الکتاب» که «آل» آن خوانده نمی‌شود.

«الکل» که «آل» آن خوانده می‌شود.

- ممکن است در عنوان برخی از کتاب‌ها،

عملگرهای محاسباتی به کار رفته باشد که این

عملگرها با توجه به قواعد برگه آرای فارسی باید به

شکل صحیح املاء آنها برگه آرای شوند. مانند «+»

که باید به صورت «به علاوه» در نظر گرفته شود؛

یعنی این که هر کجا که این علامت در رکورد به

کار رود، باید به صورت «به علاوه» خوانده شود، ولی

شکل آن به صورت «+» حفظ گردد. این نکته حتماً

باید در فهرست‌های رایانه‌ای توجه شود و

برنامه‌نویسان به آن توجه نمایند.

- در برگه آرای نام خانوادگی مرکب، سیستم

رایانه‌ای عنصر یا عناصر اضافه شده را باید در

ترتیب الفبایی در نظر بگیرد و حد نام خانوادگی را

«ویرگول» تعریف کند، چرا که علامت «ویرگول»

در نمایه پدیدآور برخلاف نمایه عنوان، معنی‌دار

می‌باشد؛ مانند:

رضا، احمد

رضا، حمید

رضاپور، احمد

رضازاده، علی

رضازاده مشفق، احمد

همان طور که در بالا اشاره گردید، چون سیستم

داده آرای شناسه‌ها در زبان فارسی حرف به حرف

می‌باشد، اگر برنامه نویسان حد نام خانوادگی را

«ویرگول» ندانند، بازیابی نام پدیدآور با مشکلات

خاصی همراه خواهد بود. در برخی از نرم افزارهای

کتابخانه‌ای که در ایران طراحی گردیده‌اند، این

مشکل وجود دارد.

- اعداد: شروع برخی از عناوین ممکن است با

عدد باشد. مانند «۱۰۰ داستان کوتاه» در قواعد فهرست

نویسی کتابخانه ملی جمهوری اسلامی ایران، این عنوان را به شکل زیر می‌نویسند:
۱۰۰ [صد] داستان کوتاه

که در داده آرایی نرم‌افزار کتابخانه‌ای، باید قابلیت این امر را داشته باشد تا بتواند عدد را در حرف برگه آرایی نموده و شکل صحیح آن را حفظ کند. این امر باید با برنامه نویسی خاصی در هنگام طراحی سیستم انجام گیرد.

آرایش شناسه‌ها در سرعنوان‌های موضوعی الکترونیکی

یکی از ابزارهای فهرست نویسی تحلیلی، سرعنوان موضوعی می‌باشد. سرعنوان‌های موضوعی کنگره (Subject Headings of Library of Congress) و سرعنوان‌های موضوعی فارسی، از جمله سرعنوان‌هایی هستند که در ایران کاربرد زیادی دارند. برخی از نرم افزارهای کتابخانه‌ای، پایگاه‌جانبی تحت عنوان سرعنوان موضوعی تهیه کرده‌اند که کمک کار کتابدار در هنگام فهرست نویسی می‌شود.

در ایران نیز برخی نرم افزارهای کتابخانه‌ای به تهیه سرعنوان‌های موضوعی الکترونیکی دست زده‌اند که در داده آرایی دارای مشکلات فراوانی هستند.

باید توجه داشت که در سرعنوان‌های موضوعی فارسی، دو نوع برگه آرایی مطرح می‌شود: یکی چگونگی ترتیب قرار گرفتن شناسه‌های مختلف است که ممکن است به صورت فرم‌های تک واژه‌ای، عبارتی، اسم و صفت مقلوب، اضافه مقلوب، با توضیحگر و... باشد و دیگر این که رسم الخط فارسی چه مشکلاتی را در برگه آرایی الفبایی موجب می‌شود. یعنی این که در داده آرایی علاوه بر این که باید به سیستم الفبایی توجه داشت، باید به نوع شناسه نیز توجه کرد که ممکن است در ترتیب قرار گرفتن برعکس الفبایی عمل کند.

برای مثال اگر انواع شناسه‌ها را در یک کلمه واحد داشته باشیم، آنها را به ترتیب زیر الفبایی می‌کنیم: ابتدا موضوع تک واژه‌ای، سپس تقسیمات تاریخی، و سپس تقسیمات فرعی و پس از آن به ترتیب شناسه مقلوب، شناسه با توضیحگر، شناسه با نقطه و شناسه عبارتی می‌آید. این امر نشان می‌دهد که نرم افزارهایی که به طراحی پایگاه اطلاعاتی می‌پردازند، باید به همه علامت‌ها و قواعد داده آرایی شناسه‌های موضوعی توجه داشته باشند و آنها را در برنامه نویسی به کار برند. صرف الفبایی کردن شناسه‌ها نه تنها کمکی به فهرست نویسی نمی‌کند، بلکه او را دچار سرگردانی می‌نماید. سرعنوان‌های موضوعی الکترونیکی که در ایران

طراحی شده‌اند، صرفاً به الفبایی کردن شناسه‌ها پرداخته‌اند که این امر مشکلات زیادی را برای استفاده از آن به وجود آورده است. باید توجه داشت که سرعنوان‌های موضوعی، دستنامه هستند و باید به راحتی از آنها استفاده کرد. برای مثال نمونه‌ای از شناسه‌های سرعنوان موضوعی چاپی و الکترونیکی در ذیل می‌آید:

نمونه‌ای از یکی از صفحات سرعنوان‌های موضوعی فارسی

آب

- آلودگی

- افزایش منابع

- باکتری شناسی

- تجزیه و آزمایش

- تصفیه

آب بخشی

آب بها

آبادان

در نرم افزارهای کتابخانه‌ای اگر به علائم به کار رفته در سرعنوان‌های موضوعی توجه نداشته باشند، تصویر بالا به شکل زیر نشان داده می‌شود:

در نمونه بالا سرعنوان "آب" و در ذیل آن تقسیمات فرعی آن آورده شده است که نشان می‌دهد "آب" شناسه اصلی است و شناسه‌هایی که با "آب" مشخص شده‌اند، تقسیم فرعی می‌باشند. یک فهرست نویس تقسیمات فرعی یک مدخل اصلی را باید در زیر آن ببیند تا به ارتباط آنها با همدیگر پی ببرد.

نمونه‌ای از آرایش داده‌ها در سرعنوان‌های موضوعی الکترونیکی:

آب

آب - آلودگی

آبادان

آب - افزایش منابع

آب - باکتری شناسی

آب بخشی

آب بها

آب - تجزیه و آزمایش

در شکل الکترونیکی توجه می‌شود تا مدخل اصلی و مدخل‌های فرعی بدون توجه به فرعی و اصلی بودن آنها مرتب می‌شود و هیچ فهرست نویسی نمی‌تواند سرعنوان‌های فرعی یک موضوع اصلی را پیدا کند که همین امر باعث شود که این پایگاه قابلیت استفاده نداشته باشد.

نحوه مرتب کردن تاریخ یک کشور نیز نمونه‌ای خوب جهت تفاوت دو نوع سرعنوان چاپی و الکترونیکی است. چرا که در سرعنوان چاپی شناسه‌های مربوط به تاریخ یک کشور به ترتیب

سال حکومت می‌باشد و بدین شیوه به راحتی می‌توان فهمید که چه حکومت‌هایی به ترتیب بر یک کشور حکومت کردند. اما در شکل الکترونیکی این نحوه ترتیب به صورت الفبایی می‌باشد و به راحتی تشخیص این که کدام حکومت جلوتر از حکومت دیگر بر آن کشور تسلط داشته‌اند ممکن نیست.

داده آرایی رده‌ها در نرم افزار:

در داده آرایی رده‌های کنگره و دیویی، نرم افزارها باید نکات خاصی را توجه نمایند. در رده‌بندی کنگره ابتدا باید حروف الفباء و سپس ترتیب اعداد در نظر گرفته شود. آنچه باید نرم افزارها در برگه آرایی رده کنگره رعایت نمایند این است که بعد از شماره اصلی مشابه، اعشار حرفی را باید قبل از اعشار عددی مرتب نمایند و این نیازمند به برنامه نویسی ویژه‌ای است. چراکه در منطق کامپیوتر، اعشار عددی قبل از اعشار حرفی قرار می‌گیرد و این برخلاف قواعد داده آرایی رده‌ها می‌باشد. مانند: «AG ۱۲۱ / ۵» باید بعد از «AG ۱۲۱ / ۳» قرار گیرد.

در سیستم کامپیوتری می‌توان این مشکل را با نوشتن اعداد اعشاری عددی هم سطح شماره اصلی رفع نمود.

در رده‌بندی دیویی علاوه بر برگه آرایی به ترتیب شماره، آنچه نرم افزارها باید به آن توجه داشته باشند این است که در برگه آرایی، رده ۰ فا ۴ (رده زبان‌های ایرانی) باید قبل از ۴۲۰ و رده ۰ فا ۸ قبل از ۸۱۰ قرار گیرد. بنابراین، برنامه نویسان نرم افزارهای کتابخانه‌ای باید به این امر توجه داشته باشند.

1. ALA filing rules\Filing comitte. Resource and Technical services, provision, American Library Association.- Chicago: ALA, 1980.
2. BLASE Filing Rules\British Library Filing Rules Committee.- London: British Library. 1980.
3. File as if.
4. File is as.
5. Eric Hunter. Cataloging.- 3rd.- London: Library Assosiation Publication, 1991, pp.112-113.

۶. رحمت الله فتاحی. فهرست کامپیوتری و تفاوت‌های آن با برگه‌دان. فصلنامه کتاب، دوره پنجم، شماره ۱۰۲ (بهار و زمستان ۱۳۷۳): ص ۲۶-۲۷.

۷. در مورد قواعد برگه‌آرایی و مشکلات رسم الخط فارسی به منابع زیر رجوع شود:

- فرهاد وزیری، دستورنامه برگه‌آرایی، ویرایش ۲ (تهران: کتابخانه ملی، ۱۳۷۶).
- سعید اکبرنژاد، دستورنامه برگه‌آرایی، فصلنامه کتاب، دوره هشتم، شماره ۳ (پاییز ۱۳۷۶)، ص ۱۴۸-۱۴۴