

زبان فارسی در تبادلِ اطلاعاتِ کتاب‌شناختی

ماندانا صدیق‌بهبزادی

هر کشوری، با هر نظام اقتصادی و در هر مرتبه‌ای از پیش‌رفت علمی و تکنولوژیکی، به ایجاد نظام‌های اطلاعاتی قابل اعتماد و سازمان یافته نیاز دارد تا بتوان آخرین آثار مدون و مشاهدات عینی و اندیشه‌های بشری را، با بهره‌گیری از آنها، گردآورد و در اختیار پژوهش‌گران قرار داد. برای برآوردن نیازهای اطلاعاتی پژوهش‌گران، آگاهی از اطلاعات کتاب‌شناختی مواد و دسترسی به آنها ضرورت دارد و، در این امر، کاربردهای زبان مادری مسئله‌ای مهم و در خور توجه است.

انقلاب تکنولوژیکی، بیش از هر زمان، ملت‌ها را به هم نزدیک و امکان دست‌یابی به اطلاعات را در کوتاه‌ترین مدت و به رغم موانع مرزی فراهم ساخته است. گسترش اطلاعاتی ناشی از استفاده از تکنولوژی جدید در وهله اول با زبان ربط پیدا می‌کند. زبانی که برای دست‌یابی به اطلاعات قابل استفاده نباشد به تدریج توانایی خود را برای ایجاد ارتباط مناسب و استفاده از امکانات فنی از دست می‌دهد. با توجه به این امر، بیشتر کشورها، برای حفظ و تقویت زبان خود، به سرمایه‌گذاری در تکنولوژی زبان و به ایجاد زیرساخت مناسب برای آن اقدام کرده‌اند.

برای آگاه ساختن علاقه‌مندان از مواد منتشر شده و از راه دست یافتن به آنها، فدراسیون بین‌المللی انجمن‌ها و مؤسسات کتاب‌داری (ایفلا)^۱، در سال ۱۹۷۴، «کنترل

1) International Federation of Library Associations and Institutions (IFLA)

کتاب‌شناختی جهانی^۲ را طرح‌ریزی کرد که مقصود از آن ضبط اطلاعات کتاب‌شناختی درباره مواد منتشر شده در جهان بود. برای دسترسی به خود مواد، طرح «دست‌رسی جهانی به انتشارات»^۳، اول بار در نشست ۱۹۷۳ ایفلا، به صورت برنامه‌ای اعلان شد که زیر پوشش آن هر کس - در هر نقطه‌ای از جهان چه کشورهای در حال رشد، چه کشورهای پیش‌رفته، چه در شهرهای بزرگ و چه در شهرهای کوچک - بتواند به مواد منتشر شده دسترسی پیدا کند.

فکر تدوین کتاب‌شناختی در سطح جهانی ابتدا در کنفرانس ۱۹۶۱ ایفلا در پاریس و سپس در کنفرانس ۱۹۶۹ در کپنهاگ مطرح شد. مقصود بالا بردن اطلاعات کتاب‌شناختی به سطح معیار بود. در کنفرانس ۱۹۶۹، خانم سوزان اونوره^۴ طرح ایجاد نظام مبادله اطلاعات کتاب‌شناختی در سطح جهانی را عنوان کرد و متذکر شد که هر کشور موظف است اطلاعات کتاب‌شناختی توصیفی انتشارات خود را، مطابق با معیارهای جهانی، با توجه به ویژگی‌های ملی و فرهنگی، به صورت ماشین‌خوان تهیه کند و در اختیار همگان قرار دهد.^(۱)

طرح «کنترل کتاب‌شناختی جهانی»، چه از نظر ملی و چه از نظر بین‌المللی، نقش حیاتی پیدا کرده است. از مهم‌ترین خصوصیات آن دارا بودن بُعد جهانی است. نیل به مقاصد آن مستلزم بهره‌مندی از ابزارهای گردآوری و پردازش و انتقال الکترونیکی اطلاعات (از طریق نظام‌های رایانه‌ای، نظام‌های پیوسته^۵ و شبکه‌های رایانه‌ای) در محیط‌های چندزبانه است.

در سال ۱۹۸۶، در کنفرانس ایفلا در توکیو، مسائل مربوط به رایانه‌ای کردن زبان‌ها و خط‌های گوناگون (در محیط‌های چندزبانه) مطرح شد. نتایج حاصل از این کنفرانس به شرح زیر بوده است:

۱. تدوین معیارهایی برای آوانویسی زبان‌های گوناگون؛
۲. هم‌کاری کتاب‌خانه‌ها و مراکز اطلاع‌رسانی کشورهای که از زبان و خط غیر لاتینی در توسعه موازین معیاری استفاده می‌کنند؛
۳. توسعه سری نویسه‌های معیار؛

2) Universal Bibliographic Control (UBC)

3) Universal Availability of Publication (UAP)

4) Suzanne Honoré 5) Online

۴. تدوین موازین معیاری و راه‌حل‌های مناسب برای الفبایی کردن و بازیابی همه زبان‌ها و خط‌ها؛

۵. سازگار سازی «معیار بین‌المللی کتاب‌نامه نویسی»^۶ و «یونی مارک»^۷ با زبان‌های غیر لاتینی؛

۶. تشویق تولیدکنندگان نرم افزارها و سخت افزارهای محیط‌های چند زبانه.

معیاری کردن اطلاعات کتاب‌شناختی در سه سطح مطرح شد:

۱. سطح اطلاعات کتاب‌شناختی؛

۲. سطح اطلاعات فهرست‌نویسی؛

۳. معیاری کردن تکنولوژیکی که نقش مهمی در مبادله اطلاعات دارد.

از اقدامات مؤثر در معیاری کردن اطلاعات فهرست‌نویسی می‌توان به موارد زیر اشاره کرد:

۱. تدوین «طرح مارک»^۸ که در سال ۱۹۶۶ به همت کتاب‌خانه کنگره به منظور ایجاد نظام تبادل اطلاعات کتاب‌شناختی ماشین‌خوان شکل گرفت و جهان‌گستر شد.

۲. تدوین معیار بین‌المللی کتاب‌نامه‌نویسی با گرایش به هماهنگ‌سازی تبادل اطلاعات کتاب‌شناختی در نظام‌های رایانه‌ای؛

۳. تدوین ویرایش دوم قواعد فهرست‌نویسی انگلو-امریکن^۹ (۱۹۷۸، ۱۹۸۸) با توجه به مسائل رایانه‌ای اطلاعات کتاب‌شناختی.

هم‌اکنون معیار بین‌المللی کتاب‌نامه‌نویسی و ویرایش دوم قواعد فهرست‌نویسی انگلو-امریکن، با تغییرات و تعدیل‌های جواب‌گویی مسائل خاص زبان و فرهنگ هر کشور، در بیشتر کتاب‌خانه‌های جهان از جمله ایران به کار می‌رود. از مسائل مطرح شده در معیار اطلاعات کتاب‌شناختی و فهرست‌نویسی، طراحی جدول‌هایی است مربوط به

6) International Standard Bibliographical Description (ISBD)

تهیه و تدوین معیار فهرست‌نویسی توصیفی کتاب و سایر مواد کتاب‌خانه‌ای

7) UNIMARC

8) MARC (Machine Readable Catalog Project)

نظامی که نخستین بار کتاب‌خانه کنگره آمریکا به وجود آورد تا سازمان‌دهی و اشاعه اطلاعات کتاب‌شناختی ماشینی برای ثبت در کتاب‌شناسی‌های ملی یا محلی را میسر سازد.

9) Anglo-American Cataloging Rules

آوانویسی یا حرف‌نویسی (نویسه گردانی) به عنوان ابزار کنترل کتاب‌شناختی و کنترل اطلاعات مربوط به مستندات. ضبط صحیح اعلام و اصطلاح‌های بیگانه یا برگردان و بازنویسی آنها به الفبای زبان مادر در تدوین هر نوع ابزار کتاب‌شناختی، از قبیل دایرة‌المعارف‌ها، و در علم کتاب‌داری و اطلاع‌رسانی، مانند تدوین سرعنوان‌های موضوعی و اصطلاح‌نامه‌ها و کتاب‌شناسی‌ها، حایز اهمیت خاصی است. از آنجا که همه آثار هر نویسنده یا درباره هر نویسنده باید، در ابزارهای کنترل کتاب‌شناختی، یک جا و در کنار هم قرار گیرد و بازیابی شود، لذا لازم است شکل ثابت واحدی برای هر نام انتخاب گردد.

در کنترل اطلاعات مربوط به مستندات دو شیوه دگرنویسی مطرح می‌شود: واج‌نگاری^{۱۰} یا آوانگاری^{۱۱} و نویسه گردانی (یا حرف‌نگاری یا حرف‌نویسی)^{۱۲}.

هرگاه در برگرداندن الفبایی به الفبای زبان دیگر تنها به تلفظ کلمات و جملات توجه داشته باشیم و همان تلفظ را با حروف الفبای زبان دیگر ثبت کنیم، در حقیقت، واج‌نگاری یا آوانگاری کرده‌ایم. در این حالت، بسا پیش می‌آید که، هرچند تلفظ درست نقل شده، نتوان از روی صورت «برگشته» صورت نوشتاری اصل را باز یافت و، به اصطلاح، «برگشت پذیری» وجود ندارد. مثلاً «خواهر»، که در آوانویسی به صورت xāhar در می‌آید، در «برگشت» به صورت «خاهر» ظاهر می‌گردد. «سازمان بین‌المللی استاندارد»^{۱۳} اصطلاح transcription را چنین تعریف می‌کند: «فرایند ضبط عناصر واجی یا تک‌واژی یک زبان با استفاده از دستگاه الفبایی خاص».^{۱۴} در دایرة‌المعارف زبان و زبان‌شناسی، این اصطلاح چنین تعریف شده است: «فرایند ضبط عناصر آوایی یا واجی یک زبان با استفاده از دستگاه الفبایی خاص».^{۱۵}

اما نویسه گردانی (حرف‌نگاری یا حرف‌نویسی) شیوه‌ای است که در آن یک حرف از یک نظام نوشتاری را با حرفی از نظام نوشتاری دیگر نشان می‌دهیم و آن، در واقع، نمایش یا نوشتن لغات و حروف یک زبان با حروف یا الفبای زبان دیگر است. در نویسه گردانی، صورت نوشتاری کلمات و جملات مورد نظر است به طوری که در صورت لزوم به آسانی بتوان از روی نویسه گردانی شده صورت نوشتاری اصلی را بازسازی کرد

10) Phonemic transcription

11) Phonetic transcription

12) Transliteration

13) International Organization for Standardization (ISO)

و بنابراین، ارزش آن در همان «برگشت پذیر» بودن آن است. اما دشواری کار در آن است که گاه تلفظ کلمه، پس از برگردان، سخت دگرگون می‌شود، مثلاً eight در نویسه‌گردانی به الفبای فارسی می‌شود: ایگهت. سازمان بین‌المللی استاندارد transliteration را چنین تعریف می‌کند: «فرایند ضبط نمادهای نوشتاری یک نظام الفبایی با استفاده از نمادهای نوشتاری متناظر در دستگاه الفبای زبان دیگر». (۴) در دایرة المعارف زبان و زبان‌شناسی، ذیل transliteration (نویسه‌گردانی) چنین آمده است: «فرایند ضبط نشانه‌های موجود در یک نظام نوشتاری بر حسب نشانه‌های موجود و متناظر آن در نظام نوشتاری دیگر». (۵) کتاب‌داران دورش را در هم می‌آمیزند، به گونه‌ای که خاصیت آوانویسی و حرف‌نویسی، هر دو، حفظ گردد.

موارد به کار بردن دگرنویسی و ضرورت آن

بیشتر کتاب‌خانه‌ها و مراکز اطلاع‌رسانی، برای همه کتاب‌های خود، به زبان‌های گوناگون، برگه‌دان دستی یا ماشینی دارند. برای آن‌که بتوان کلیه برگه‌ها را با الفبای واحد برگه‌آرایی کرد، مدخل اصلی یا شناسه را دگرنویسی می‌کنند. این دگرنویسی معمولاً از زبان‌های بیگانه به زبان بومی انجام می‌گیرد. در مورد کتاب‌هایی که از زبان‌های دیگر به زبان بومی برگردانده می‌شود، لازم است که نام نویسندگان خارجی و سازمان‌ها و مراکز، مطابق با اصول و قواعدی مدون و دقیق، دگرنویسی شود تا در ضبط اسامی خارجی ناهماهنگی و چندگونگی پیش نیاید. راه حل برای برگه‌آرایی و بازیابی در تمام زبان‌ها معیار کردن طرح‌های دگرنویسی است.

در امر دگرنویسی زبان فارسی با دو مسئله روبرو هستیم. یکی مسئله مربوط به کتاب‌خانه‌های خارج از ایران، که در آنها همه خطوط را به الفبای لاتینی (مانند کتاب‌خانه کنگره) یا به الفبای سیریلیک (روسیه) برمی‌گردانند؛ دیگر، مسئله کتاب‌خانه‌های ایرانی که در آنها نام نویسندگان و واژه‌های بیگانه به خط فارسی درمی‌آید.

بررسی دگرنویسی از فارسی به زبان‌های خارجی

ضبط تلفظ اعلام و واژه‌های فارسی در زبان‌های خارجی با ناهماهنگی روبروست. در دگرنویسی اسامی و واژه‌های فارسی به لاتینی قاعده خاصی وجود ندارد. سازمان‌ها و

افراد، بنا بر ذوق و علم خود، هریک به گونه‌ای، صورت نوشتاری کلمات فارسی را به زبان‌های خارجی برگردانده‌اند. مثلاً خراسان به صورت‌های Xorāsān و Khorāsān؛ جلال‌الدین به صورت‌های Jalāledin، Djalāledin، Djālalodin ضبط شده است. در دگرنویسی، برای ج از J و dj، برای ش از sh، s، š، برای خ از x، kh، برای ق از q، gh، k، و برای غ از gh، ġ، q استفاده شده است.

تاکنون، طرح‌هایی که برای دگرنویسی زبان فارسی به کار رفته بیشتر از فارسی به انگلیسی بوده است و، از آنجا که «دگرنویسان» به زبان فارسی و تلفظ صحیح آن آشنا نبوده‌اند، از قواعد دگرنویسی ترکی یا عربی به انگلیسی پی‌روی کرده‌اند. در این دگرنویسی‌ها، و (ضمّه) با u و - (کسره) با i نشان داده شده است. در کتاب‌خانه کنگره، از جدول نویسه‌گردانی فارسی و عربی / انگلیسی، تدوین کتاب‌خانه کنگره و انجمن کتاب‌داران آمریکا، استفاده می‌شود.

سازمان بین‌المللی استاندارد، برای تدوین جداول معیار نویسه‌گردانی زبان‌های بیگانه، از جمله عربی و فارسی، به انگلیسی کوشش بسیار کرده است. کمیته Iso / Te 46 / sc2 جداول مربوط به نویسه‌گردانی خط‌های سیریلیک، عربی، عبری و جز آن را به مرحله نهایی رسانده است.

از جمله طرح‌هایی که به همت ایرانیان در این زمینه تهیه شده باید از رساله دکتری ناصر شریفی،^(۴) در دگرنویسی حروف و کلمات فارسی به انگلیسی نام برد که در تنظیم قواعد دگرنویسی کتاب‌خانه کنگره تأثیر به سزا داشته است. دیگر روش تبدیل اسامی جغرافیایی به لاتین^(۵) است که اداره جغرافیایی ارتش منتشر ساخته و کوششی است برای یک‌دست ساختن نام‌های جغرافیایی مناطق ایران.

مؤسسه استاندارد و تحقیقات صنعتی ایران، در مقام یکی از اعضای سازمان بین‌المللی استاندارد، با توجه به قوانین و قواعد همان سازمان، جدولی برای نویسه‌گردانی حروف فارسی به لاتینی (CD 233-3)، تهیه کرده است. در این جدول، برای هر حرف فارسی یک حرف لاتینی پیشنهاد شده است، مثل š برای ش. در تدوین شیوه پیش‌نهاد شده برای دگرنویسی حروف و کلمات فارسی به لاتینی، به نظام‌های رایج‌ای حرف به حرف نویسی توجه شده و اصل برگشت پذیری ملاک قرار گرفته است. مشکلی که برای زبان فارسی و عربی وجود دارد نویسه‌گردانی زیر و زبر و ضوابط است. تلاش سازمان

بین‌المللی استاندارد بر آن است که استفاده از علائم را محدود سازد و آن را به حدی که ضروری باشد کاهش دهد. این سازمان، در شیوه‌نویسه‌گردانی، بیشتر به برگشت‌پذیری به شکل اولیه توجه دارد تا به تلفظ اصوات.

دگرنویسی زبان‌های بیگانه به فارسی

به دلیل تفاوت فاحش نشانه‌های الفبای فارسی با نشانه‌های الفبای زبان‌های خارجی، برای ضبط اعلام و وام‌واژه‌ها در فارسی مشکلات جدی وجود دارد. مصوت‌های /a/، /o/، /e/ در خط فارسی با زنمود نوشتاری ندارند. اشتباه‌های تلفظی که از این بابت در خواندن پیش می‌آید بسیار است. این شکل به‌ویژه در ضبط اعلام و واژه‌های بیگانه نمودار می‌گردد. ضبط اعلام و واژه‌های بیگانه در فارسی دچار تشتت و چندگونگی است. البته، برحسب میزان نزدیکی دستگاه‌های آوایی (صامت و مصوت) زبان‌های وام‌دهنده با همتای آنها در فارسی، میزان مشکلات ضبط واژه‌های قرصی فرق می‌کند. در مورد دگرنویسی اسامی و واژه‌های خارجی به خط فارسی قاعده خاصی وجود ندارد و هر کس، از مترجم و نویسنده و محقق تا روزنامه‌نویس، بنا بر سلیقه و ذوق یا میزان آشنایی و علم خود، به گونه‌ای دیگر صورت نوشتاری کلمات را از زبان‌های خارجی به فارسی برگردانده است.

برای پرهیز از تشتت و چندگونگی، دگرنویسی اسامی و واژه‌های بیگانه باید بر پایه اصول صحیح و مدون انجام گیرد. در تدوین قواعد باید از دستاوردهای نظری و عملی زبان‌شناسی بهره گرفت. برای رفع تشتت و اجتناب از آن راهی نیست جز وضع و اتخاذ قراردادهای ثابت.

دایرة‌المعارف فارسی،^(۸) به مقتضای نیاز به برگرداندن اسامی خارجی به فارسی، نخستین گام را در این راه برداشت؛ اما در آن، جدول معادل‌های فارسی حروف انگلیسی و ضوابطی که برای دگرنویسی اختیار شده نیامده است. پرناز عظیمی،^(۹) در بخش دوم پایان نامه دوره کارشناسی ارشد خود، به ضوابطی درباره دگرنویسی کلمات از انگلیسی به فارسی اشاره دارد. ادیب سلطانی در راهنمای آماده ساختن کتاب^(۱۰) قواعد و ضوابطی برای دگرنویسی نام‌ها و واژه‌های بیگانه در زبان فارسی پیش‌نهاد کرده است.

کتاب‌خانه ملی ایران درباره ضبط نام‌های بیگانه به خط فارسی از ضوابط و قواعدی

پی‌روی می‌کند و اساس را دور نبودن از صورت مشهور و معمول و نزدیک بودن به تلفظ اصلی (بر اساس ملیت صاحب اثر) قرار می‌دهد و از میان انواع صور نوشتاری که مترجمان برای نام اشخاص اختیار کرده‌اند آن را انتخاب می‌کند که به تلفظ اصلی نزدیک‌تر باشد.

هیچ‌یک از این کوشش‌ها بر اساس بررسی و تجزیه و تحلیل وجوه اشتراک و افتراق واج‌های دو زبان و در چارچوب کار عملی زبان‌شناسی انجام نگرفته است. اما در تدوین شیوه‌نامه ضبط اعلام انگلیسی در فارسی^(۱۱)، که طرح آن با پشتیبانی مالی و فنی مرکز نشر دانشگاهی و کتاب‌خانه ملی جمهوری اسلامی زیر نظر هیئت مشاوران (آقایان دکتر محمدرضا باطنی، دکتر علی محمد حق‌شناس، احمد سمیعی/گیلانی، و کامران فانی)، تهیه شده، به وجوه اشتراک و افتراق واج‌های زبان انگلیسی و فارسی توجه شده است. در تعیین قواعد ضبط اعلام، اصول کلی زیر مورد نظر است:

۱. نام افراد با توجه به ملیت آنان تلفظ می‌شود؛
 ۲. هر آوایی به آوای نظیر یا به نزدیک‌ترین آوا در فارسی برگردانده می‌شود. برگردان آوا با حرفی که در الفبای فارسی برای آن وجود دارد نوشته می‌شود.
 ۳. در دگرنویسی اعلام از قواعد نظام آوایی زبان فارسی پی‌روی می‌شود (هم از نظر واجگان، هم از نظر واج آرایی و هم از نظر عناصر زیر زنجیری)؛
 ۴. صورت نوشتاری اعلامی که پذیرفته شده و جا افتاده باشد حفظ می‌شود؛
- پیش‌نهاد می‌شود که، با استفاده از دستاوردهای نظری و عملی زبان‌شناسی، در مورد همه اعلام زبان‌های بیگانه، قواعد دگرنویسی به فارسی تنظیم و تدوین گردد.

معیارسازی رایانه‌ای

ضوابط معیارسازی رایانه‌ای باید بر نظام‌های مورد استفاده در سطح بین‌المللی مطابقت پذیر باشد. سازمان بین‌المللی استاندارد از آغاز به این نکته توجه نمود که راه‌حل‌های رایانه‌ای برای معیارهای مربوط به خط لاتینی الزاماً در مورد خطوط غیر لاتینی اجراشدنی نیست. در کنترل کتاب‌شناختی رایانه‌ای، خصوصیات زبان‌های غیر لاتینی متضمن مشکلات خاص خود است که با راه‌حل‌های زبان‌های لاتینی چاره نمی‌شود. مسئله سری نویسه‌های معیار در محیط‌های چندزبانه و گداهای نویسه معیار

از مسائل مهمی بود که در «ایفلا ۱۹۹۱» مطرح شد. از جانب مؤسسات بین‌المللی، از جمله سازمان بین‌المللی استاندارد، تلاش‌هایی صورت گرفت تا معیارهایی برای محیط‌های چندزبانه تدوین گردد.

در سازمان بین‌المللی استاندارد، دو گروه به کتاب‌داری اختصاص داده شد: یکی کمیته فنی 46 Tc، که مسئول توسعه معیارهای کتاب‌خانه‌ها و خدمات اطلاع‌رسانی است؛ دیگر کمیته مشترک فنی (JIC) ISO/IEC که مسئول معیارسازی وسایل رایانه‌ای است. در سال ۱۹۸۰، برای برقراری سری نویسه‌های بین‌المللی، از جانب دو گروه کوشش‌هایی شد. کمیته مشترک فنی سری نویسه‌هایی به نام Unicode و کمیته فنی استاندارد ISO/IEC 10646 را هم‌زمان طراحی کرد. ISO/IEC 10646 از این جهت برای کتاب‌خانه‌ها با ارزش بود که بسیاری از نیازهای کتاب‌شناختی را که در Unicode نادیده گرفته شده بود در برمی‌گرفت. هر دو گروه در صددند که، با استفاده از سری کدهای ۱۶ بیتی، معیارهایی برای محیط‌های چندزبانه اختیار کنند.

بدین ترتیب، معیار بین‌المللی ISO/IEC 10646، معروف به سری نویسه‌های جهانی (UCS)، در سال ۱۹۹۲ پذیرفته و در سال ۱۹۹۳ منتشر شد. قسمت اول ISO/IEC 10646/1 و ویرایش Unicode 101، در واقع، یکی هستند. از ویژگی‌های UCS تلاش برای ایجاد سری کدهای ۱۶ بیتی برای همه زبان‌هاست.

ویژگی‌های زبان فارسی در نظام‌های رایانه‌ای

وقتی صحبت از تکنولوژی اطلاعاتی و ارتباطی است، مسئله ربط پیدا می‌کند با اطلاعات راجع به خط و زبان که مواد اطلاعات را در بردارند - با زبان پدیدآورنده اسلوب‌های تبدیل متن به گفتار، ابزارهای جستجوی متن و فهم و تجزیه و تحلیل متون برای طراحی نظام‌های اطلاعاتی مدیریت، ترجمه رایانه‌ای و ابزارهای مناسب برای دسترسی به بانک‌های اطلاعاتی و جز آن.

وقتی از خط فارسی سخن به میان می‌آید، مسائلی مطرح می‌گردد از قبیل راست به چپ‌نویسی، متن‌های دوزبانه با جهات دوگانه، نشانه‌های نوشتاری فارسی و حروف منفصل و متصل، چندگونگی هر یک از حروف - بزرگ، کوچک، آغازی، میانی، پایانی، تنها. برخوردی که در ایران و کشورهای عربی با این مسائل شده دو نوع است: روش

ذخیره (۱) تک‌نمادی (۲) چندنمادی اطلاعات. در روش تک‌نمادی، متن فارسی، صرف‌نظر از شکل نمایش حروف آن، به یکی از اشکال (حروف بزرگ یا کوچک) ذخیره می‌شود؛ مثلاً «کتاب‌داری» به صورت «ک ت اب داری» در پرونده قرار می‌گیرد. برنامه نمایش دهنده اطلاعات وظیفه انتخاب شکل نمایش حروف را، با توجه به محل آنها در کلمه، به عهده دارد. در روش چندنمادی، اطلاعات به خط فارسی به نمادهای نمایشی خود ذخیره می‌شوند. اطلاعات «کتاب‌داری» در پرونده دقیقاً به همین نحو ذخیره می‌شود. در همه مراحل، برای تبدیل داده‌ها، به برنامه‌های رابطی نیاز است تا عمل تبدیل اطلاعات را، برای دستگاهی دیگر، که احتمالاً از اسلوب ذخیره‌سازی دیگر استفاده می‌کند، انجام دهد. بنابراین، هرچه سریع‌تر باید در فکر معیارسازی‌های فارسی بود. کدهای معیار فارسی باید ۸ بایتی و حداقل دو نمادی باشند و بتوانند از ASCII پشتیبانی کنند.^(۱۲)

مسائل دیگری همچون شیوه نوشتاری، شیوه جستجو و مرتب‌سازی نیز مطرح است. برای مرتب‌سازی رکوردهای یک پایگاه اطلاعاتی به خط فارسی و برای جستجوی عبارت‌های فارسی از شیوه برابرگذرانی حرف به حرف استفاده می‌شود. با استفاده از این شیوه، عمل جستجو، مستقل از شیوه نوشتاری متن (کتابشناسی و کتاب‌شناسی)، میسر خواهد بود. حروف تعریف، به طور کلی، در مرتب‌سازی منظور نمی‌شوند. انواع همزه در فارسی، برحسب بافت، در سنجش داده‌ها دخیل می‌شوند. جای تأسف است که متخصصان برنامه‌ریزی رایانه‌ای به جای تولید نرم افزارهای پایه و مطمئن برای خط فارسی، از نرم افزارهای فرنگی برای ذخیره و پردازش اطلاعات به خط فارسی استفاده کرده‌اند.^(۱۳)

در مرتب‌سازی رکوردهای یک پایگاه اطلاعاتی به خط فارسی (برگه آرای الفبایی)، از جهت شیوه املا فارسی، مشکلات دیگری نیز وجود دارد که در رایانه حل نشده است. الف مقصوره، که در خط فارسی عموماً به صورت «ا» یا «ی» نوشته می‌شود، در رایانه، برحسب شکل نوشتاری، مرتب می‌گردد و در دوره الفبایی قرار می‌گیرد؛ مانند: اسحاق / اسحق، اسماعیل / اسمعیل، کبرا / کبری، مرتضا / مرتضی.

به این مقوله کلماتی مانند مشکوة را باید افزود که به صورت مشکات نیز نوشته می‌شود و، در بازیابی رایانه‌ای، برحسب صورت نوشتاری، در دوره الفبایی قرار می‌گیرد.

این مسائل در کنترل کتاب‌شناختی مشکل‌آفرین است. البته امکان معادل‌سازی وجود دارد، یعنی می‌توان مثلاً کبری را معادل کبرا تعریف کرد؛ ولی، در این حالت، فقط یک صورت نوشتاری بیشتر بازبایی نمی‌شود و، در موقع وارد کردن اطلاعات به رایانه، باید هر بار معادل مرتب‌سازی کلمات را هم وارد ساخت که خود بسیار وقت‌گیر است.

از طرفی، در مورد زبان فارسی، نیاز به الگویی واحد برای ذخیره اطلاعات کتاب‌شناختی در رایانه در سطح ملی احساس می‌شود تا بتواند در مراکز اطلاعاتی نقش مؤثری برای هماهنگ‌سازی شیوه‌های ذخیره و بازبایی اطلاعات ایفا کند.

با استفاده از الگوی واحد ذخیره اطلاعات کتاب‌شناختی و تکنولوژی ارتباطات، مراکز اطلاعاتی و تحقیقاتی و کتاب‌خانه‌ها می‌توانند شبکه اطلاعاتی هماهنگی ایجاد کنند. در حال حاضر، در بیشتر کشورهای پیش‌رفته و بعضی از کشورهای در حال رشد، در زمینه اطلاعات کتاب‌شناختی، از فرمت مارک (فهرست‌نویسی ماشین‌خوان) استفاده می‌کنند. این الگوی رایانه‌ای، برای انواع مواد کتاب‌خانه‌ای، یک رشته معیارها برای ذخیره اطلاعات کتاب‌شناختی در بردارد. فرمت مارک، هرچند از روش‌های معیار پی‌روی می‌کند، کشورهای گوناگون، با استفاده از آن، متناسب با نیاز و خط و زبان خود، می‌توانند الگوهای خاص خود را پدید آورند.

در ایران نیز، مانند سایر کشورها، باید الگوی فهرست‌نویسی ملی مناسبی طراحی شود. در حال حاضر، طراحی چنین الگویی در کتاب‌خانه ملی در دست مطالعه است و مرکز اطلاعات و مدارک علمی ایران نیز الگوی مارک ملی خاصی به نام «ایران مارک» را پیشنهاد کرده است.

سری نویسه‌های معیار فارسی

در حال حاضر، نرم‌افزارهای متفاوت، با فرمت‌ها و کدهای فارسی گوناگون، طراحی شده است که بیشتر تنها در همان محیط سازمانی تعریف شده کاربرد دارند. برخی از فارسی‌سازهای موجود معیار مؤسسه استاندارد و تحقیقات صنعتی را پشتیبانی می‌کنند ولی عده آنها زیاد نیست.

با گسترش نرم‌افزارهای فارسی و استفاده آنها از سری نویسه‌های گوناگون، کم‌کم معلوم می‌شود که ذخیره اطلاعاتی آنها در میان برخی برنامه‌های کاربردی مبادله‌شدنی

نیست. از طرفی، با پیش‌رفت‌های اخیر در پردازش، با ذخیره و بازیابی اطلاعات، با نیاز به بهره‌گیری از شبکه‌های رایانه‌ای (راه دور و محلی) و نیاز به اتصال نظام‌های رایانه‌ای گوناگون، ناسازگاری اطلاعات فارسی به طور روزافزونی نمودار می‌گردد. مادام که شبکه محدود به یک سازمان زیرنظر یک مدیریت باشد، مسئله‌ای از نظر انتقال اطلاعات به خط فارسی پیش نخواهد آمد. اما نیاز به نظام‌های روزآمد، که در خور استفاده مشترک از اطلاعات باشد، روز به روز محسوس‌تر می‌شود و فوریت ایجاد یک نظام هماهنگ انتقال اطلاعات به خط فارسی آشکار می‌گردد.

ارجاعات

1. Dorothy ANDERSON, "IFLA's Programme for UBC: the Background and Basis," *IFLA Journal*, (Vol. 1, No. 1, 1975) p.5

2. International Organization for Standardization. *General Principles for the Conversion of One Written Language into Another*. Revised Version. Introductory Note on the General Principles of Transliteration, first published in Iso/Rq (1954).

3. *The Encyclopedia of Language and Linguistics*, Chief editor, ASHER (1994), S. V. "Phonetic Transcription: History".

4. International Organization for Standardization *General Principles for the Conversion...*

5. *The Encyclopedia of Language and Linguistics*, Chief editor, ASHER (1994). S. V. "Phonetic Transcription: History".

6. Nasser SHARIFY. *Cataloging of Persian Works: Including Rules for Transliteration, Entry and Description*. Chicago, ALA, 1959.

۷. روش تبدیل اسامی جغرافیایی به لاتین، تهران، اداره جغرافیایی ارتش، ۱۳۴۵.

۸. دایرةالمعارف فارسی، به سرپرستی غلامحسین مصاحب، تهران، فرانکلین، ۱۳۴۵، مقدمه.

۹. پرناز عظیمی، دگرنویسی کلمات و اسامی از فارسی به انگلیسی و از انگلیسی به فارسی، پایان‌نامه کارشناسی ارشد، دانشگاه تهران، دانشکده علوم تربیتی، ۱۳۵۲.

۱۰. میر شمس‌الدین ادیب سلطانی، راهنمای آماده ساختن کتاب، تهران، سازمان انتشارات و آموزش انقلاب اسلامی، ۱۳۶۵، ص ۱۹۷-۲۹۵.

۱۱. ماندانا صدیق‌بهبودی، شیوه‌نامه ضبط اعلام انگلیسی در فارسی، تهران، مرکز نشر دانشگاهی، کتاب‌خانه ملی جمهوری اسلامی ایران، ۱۳۷۵.

۱۲. بهزاد ترابی، «کدهای انتقال اطلاعات فارسی: راهی که در پیش است»، گزارش کامپیوتر، آذر و دی ۱۳۷۲، ص ۸۴-۸۷.

۱۳. شرکت نرم‌افزار و سخت‌افزار ایران. درد دلی چند، از گروه تحقیق و توسعه نرم‌افزار مدیریت پایگاه‌های اطلاعاتی (نوسا)، (چاپ نشده).

منابع

1. *Automated Systems for Access to Multilingual and Multiscript Library Materials: Problems and Solutions*. Edited by Christine **BOLMEGER**. München: Saur, 1987.
2. *Automated Systems for Access to Multilingual and Multiscript Library Materials*. München: Saur, 1994.
3. **Aliprand**, Joan M. "Nonroman Scripts in the Bibliographic Environment", in *Information Technology and Libraries*, Vol. II, No. 2 (June 1992) pp. 105-119.
4. **AMAN**, Mohammed. "Use of Arabic in Computerized Information Interchange", *Journal of the American Society for Information Science*, 33 (4) July 1984.
5. **Large**, J. A. "The Foreign-Language Barrier and Electronic Information", *Online*, Vol. 14, No. 4, 1990.



پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی