

## نقش رایانه در ایجاد استانداردهای زبانی

مصطفی عاصی

پژوهشگاه علوم انسانی و مطالعات فرهنگی

استانداردسازی برای ایجاد هماهنگی در کارها، صرفه‌جویی در وقت و هزینه و انرژی، دستیابی به کیفیت بهتر و بهره‌وری بالاتر در همه‌ی زمینه‌های زندگی، چه به طور رسمی و منظم و چه به طور غیر رسمی و پراکنده، از سوی سازمان‌های خصوصی، ملی و جهانی انجام یافته است. زبان باگوناگونی‌ها و کاربردهای بسیار متنوع خود یکی از عرصه‌های عمده‌ی استانداردسازی در همه‌ی سطوح آن بوده است.

از آنجا که جنبه‌های گوناگون زبان، از واژگان گرفته تا دستور و از گفتار گرفته تا نوشتار، به گونه‌ی گسترده دستخوش استانداردسازی‌های رسمی و غیررسمی می‌گردند، این گفتار تنها به استانداردهای رسمی‌بی اختصاص دارد که در حوزه‌ی زبان فارسی رسمی و تحت تأثیر رایانه تدوین گردیده، به تصویب رسیده و به اجرا درآمده است: از جمله استاندارد ملی کد تبادل اطلاعات هشت بیتی فارسی، استاندارد ملی صفحه کلید فارسی در کامپیوتر، استاندارد جهانی نوسه گردانی خط فارسی، و استانداردهای در دست بررسی یونی‌کد فارسی، آوانویسی فارسی و نشانه‌گذاری متن‌های فارسی.

پیش‌گفتار

برای انجام دادن دقیق هرکاری معمولاً به اندازه‌گیری روی می‌آوریم. و برای انجام دادن

فرهنگ، ۳۸ و ۳۷، بهار و تابستان ۸۰، صص ۹۳-۱۰۵

مکرر همان کار یا کارهای مشابه می‌کشیم از ابزارهای اندازه‌گیری یکسانی سود بریم. هنگامی که قرار باشد نتایج کار در گستره‌ی بزرگ و توسط افراد مختلف در موقعیت‌های گوناگون و بارها و بارها به کار گرفته شود، آنگاه لزوم بهره‌گیری از اندازه‌ها، معیارها و مقیاس‌های یکسان و ثابت آشکار می‌شود. چنین شرایطی معمولاً در حوزه‌ی علم و صنعت پدید می‌آید و می‌توان همه‌ی بررسی‌ها، هم‌سنجی‌ها، معیارگذاری‌ها و توافق‌هایی را که منجر به تعیین و پیشنهاد رویه، الگو، چارچوب، ویژگی یا تعریفی یکسان و پایدار می‌گردد زیر عنوان کلی فرایند استانداردسازی<sup>۱</sup> قرار داد.

ممکن است شیوه، ابتکار یا معیار یک شخص یا یک سازمان به اندازه‌ی مطلوب و مورد پذیرش دیگران باشد که خود به خود به صورت استاندارد درآید و یا اشخاص یا سازمان‌هایی بر سر استفاده از یک الگوی واحد به توافق برسند؛ اما امروزه به منظور ایجاد هماهنگی در میان تولیدکنندگان، پرهیز از تداخل و دوباره کاری، بالا بردن کیفیت و بهره‌وری، ایجاد ایمنی و حفظ منافع مصرف‌کنندگان، ایجاد سهولت و اطمینان در صادرات و واردات و بسیاری از مزایای دیگر، در کشورها سازمان‌های رسمی برای تدوین استانداردهای ملی به‌وجود آمده‌اند. در سطح جهانی نیز نهادهایی برای هماهنگی فعالیت‌های استاندارد و تدوین استانداردهای بین‌المللی تشکیل شده است. استاندارد می‌تواند جنبه‌ی پیشنهادی، ارشادی یا اجباری داشته باشد و پدیده‌های گوناگون را در حوزه‌های دانش، صنعت، و حتی ادب و هنر دربرگیرد. البته روشن است که دقت ریاضی و یکدستی مکانیکی، که در فراورده‌های صنعتی و مفاهیم علمی عاملی بنیادین است، در مباحث ادبی و هنری ممکن است جایی نداشته باشد. و زبان در این میان، از یک سو کاربرد علمی و از سوی نقشی هنری دارد. آنجا که از آن به عنوان ابزار بیان علمی استفاده می‌شود، یا توصیف علمی خود آن هدف قرار می‌گیرد و یا به هر طریقی با پدیده‌های فنی و صنعتی برخورد می‌کند، رهیافتی منظم، قانونمند، یکدست و دور از تفنن، بی‌قاعدگی و پراکنندگی ضرورت می‌یابد.

معیارگذاری در زمینه‌ها و جنبه‌های گوناگون زبانی چه به طور رسمی و چه غیر رسمی انجام می‌پذیرد؛ از تدوین دستور زبان‌های آموزشی و تجویزی گرفته تا آیین‌نامه‌ها و شیوه‌نامه‌های انتشاراتی، از تعیین زبان معیار برای کاربرد در رسانه‌ها تا واژه‌گزینی‌های فرهنگستانی، یا فردی و سازمانی، اصلاح یا تغییر خط و بسیاری موارد دیگر. در زبان فارسی نیز نمونه‌های بسیاری را از این قبیل می‌توان یافت.

### زبان فارسی و رایانه

اگر به یاد بیاوریم که تنها پیدایش صنعت چاپ و ماشین تحریر چه تأثیر دگرگون‌کننده‌ی بر خط فارسی نهادند و بی‌آنکه بحث و جدل‌هایی - آن‌گونه که معمولاً در مورد اصلاح خط پیش می‌آید - برانگیزند بیش‌ترین تغییرات را در جهت یکسان‌سازی آن ایجاد کردند، می‌توانیم قدرت تعیین‌کننده‌ی فن‌آوری و صنعت را در این زمینه دریابیم. نکته‌ی مهمی که باید بدان توجه داشت این است که اگر بتوانیم پیش از دخالت بی‌مقدمه‌ی فن و صنعت، ویژگی‌های اساسی زبان و خط و خواست‌ها و پیشنهادهای صاحب‌نظران را تنظیم و ارائه کنیم، دگرگونی‌ها همسوبا مصالح فرهنگی ما خواهد بود و در غیراین‌صورت، این فن و صنعت است که اراده‌ی خشک و بی‌ملاحظه‌ی خود را بر ما تحمیل خواهد کرد.

رایانه با گسترش برق‌آسای خود در همه‌ی زمینه‌های زندگی تأثیری به مراتب نیرومندتر و گسترده‌تر از صنعت چاپ دارد و هم‌اکنون نیز می‌توان آثارش را بر خط و زبان فارسی دید. از جمله ورود واژگان و اصطلاحات تخصصی این رشته به زبان فارسی است که بخش بزرگی از آن (خواه وام‌واژه و خواه برابر فارسی) کاربرد عام نیز یافته است. همچنین پیدایش شیوه‌های نوین نگارش متن و قلم‌های (فونت‌های) تازه و ابتکاری در خط فارسی است که برخی از آنها به کلی با قلم‌های آشنا و سنتی تفاوت دارد. آنچه ما را وامی‌دارد این دگرگونی‌ها را بپذیریم مزایا و تسهیلاتی است که رایانه فراهم می‌سازد و اگر بخواهیم از امکانات بیش‌تری نیز بهره ببریم، ناچاریم باز هم شرایط و قواعد بیش‌تری را رعایت کنیم که بهای آنها را می‌توان کنار نهادن تفنن یا

بی‌نظمی دانست. استانداردسازی یعنی نفی بی‌نظمی.

به طور کلی، رایانه را می‌توان ابزاری برای پردازش انواع داده‌ها دانست که یکی از این انواع، داده‌های زبانی است. نخستین برخورد رایانه با زبان فارسی با میانجی خط بوده است که خود جنبه‌های گوناگون دارد. چرا که درونداد، ذخیره‌سازی، پردازش و برونداد داده‌های زبانی بیش‌تر و پیش‌تر به صورت نوشتاری انجام می‌شود. و می‌دانیم که خط فارسی خود دارای مشکلات و ناهماهنگی‌های زیادی است<sup>(۱)</sup> و هم از این‌رو، به کارگیری آن در رایانه ایجاد رویه‌های یکسان و استانداردهای رسمی و غیررسمی را ناگزیر می‌سازد. در این گفتار، تنها به چند استاندارد رسمی ملی و جهانی که تاکنون تدوین شده و نگارنده نیز در آنها شرکت داشته است اشاره می‌کنیم.

### زمینه‌ی مسئله

ممکن است تا به حال برای بسیاری از ما اتفاق افتاده باشد که یک متن فارسی را که در رایانه با کمک یک واژه‌پرداز آماده کرده‌ایم، بخواهیم در رایانه‌ی دیگری با نرم‌افزاری دیگر بخوانیم، ویرایش کنیم یا چاپ نماییم و متوجه شویم که به‌جای متن فارسی، نشانه‌های عجیب و غریب بر صفحه‌ی نمایش پدیدار یا روی کاغذ چاپ می‌شود. دلیل اصلی، ناهماهنگی میان کدهای نویسه‌های فارسی دو نرم‌افزار است. این تنها یکی از مشکلات متعددی است که نبود یک مجموعه کد استاندارد به‌وجود می‌آورد.

به دلیل آنکه ماهیت ساختمان درونی رایانه بر پایه‌ی سلول‌های الکترونیکی دوحالتی (بیت)<sup>۱</sup> بنا نهاده شده است، زبان آن زبان صفر و یک است و هر نوع داده، از جمله نشانه‌های نگاشتاری، در رایانه به طور قراردادی به کدهای صفر و یک (0/1) تبدیل می‌شوند. پس نخست باید مجموعه کدهای مورد نیاز را در جدولی مرتب کرد و به حافظه‌ی رایانه سپرد. اگر بخواهیم در یک یا چند رایانه داده‌های معینی را به کارگیریم ناچاریم یا از جدولی واحد استفاده کنیم و یا در هر مرحله جدول‌های قراردادی مبدأ و

مقصد را بررسی و مقایسه کنیم و برنامه‌ی تبدیلی<sup>۱</sup> برای تبدیل کدهای مبدأ به مقصد بنویسیم و اجرا کنیم. با توجه به این که در همه‌ی مراحل درونداد<sup>۲</sup> (تایپ کردن)، انبارش<sup>۳</sup>، پردازش<sup>۴</sup>، برونداد<sup>۵</sup> (نمایش یا چاپ) و جابه‌جایی داده‌ها<sup>۶</sup> میان چند رایانه یا درون شبکه‌های محلی<sup>۷</sup> و جهانی<sup>۸</sup> با همین کدها سروکار داریم؛ چندگانگی و ناسازگاری آنها مشکلات عظیم و اتلاف وقت، انرژی و پول فراوانی را در پی دارد.

یکی از این مجموعه‌ها با نام EBCDIC<sup>۹</sup> در دهه‌ی شصت میلادی توسط شرکت قدرتمند و انحصارگر IBM ایجاد شد و در اواخر همان دهه نیز جدول دیگری به نام ASCII<sup>۱۰</sup> از سوی سازندگان و شرکت‌های دیگر امریکایی برای مقابله با این انحصارگری ارائه گردید که به صورت استاندارد جهانی درآمد. جدول اولیه که ترکیبی از هفت بیت را<sup>۱۱</sup> به کار می‌برد، می‌توانست روی هم ۱۲۸ نشانه‌ی متشکل از حروف الفبا، ارقام عددنویسی و نشانه‌های ویژه را در خود جای دهد. این جدول برای خط‌های غیرلاتینی امکان بسیار محدودی داشت و در نتیجه، شکل گسترش‌یافته‌ی هشت بیتی جدول ASCII با ۲۵۶ امکان عرضه شد که ۱۲۸ جایگاه بعدی می‌توانست در اختیار خطوط دیگر قرار گیرد. متأسفانه از نخستین سال‌های ورود رایانه به ایران، یعنی اوایل دهه‌ی چهل خورشیدی، برای به‌کارگیری خط فارسی پراکنده کاری آغاز شد و هر سازمان، شرکت و فردی که در این زمینه فعالیت می‌کرد به سلیقه‌ی خود و بدون هماهنگی با دیگران، نشانه‌های فارسی را با هر تعداد و به گونه‌های متفاوت و با انواع ترتیب‌ها در جدول می‌چید و به کار می‌برد.<sup>(۲)</sup> این آشفتگی برای کاربران ناخوشایند و آزاردهنده و برای کارشناسان و برنامه‌سازان مشکل‌آفرین و وقت‌گیر بود.

1. converter

2. input

3. storage

4. processing

5. output

6. data transfer

7. local area network (LAN)

8. wide area network (WAN)

9. Extended Binary Coded Decimal Interchange Code

10. American Standard Code for Information Interchange

11. byte

## ۱. استاندارد ملی کد تبادل اطلاعات هشت بیتی فارسی

نخستین بررسی‌ها برای استانداردسازی خط فارسی در رایانه را کمیته‌ی مطالعاتی استاندارد تبادل اطلاعات در فارسی، در سال‌های ۱۳۵۶-۱۳۵۷، انجام داد. در این کمیته نمایندگان از اغلب سازمان‌های دست‌اندرکار داده‌پردازی، دانشگاه‌ها، مؤسسه‌ی استاندارد و تحقیقات صنعتی ایران، معاونت انفورماتیک سازمان برنامه و بودجه، و فرهنگستان زبان ایران (فرهنگستان دوم) شرکت داشتند. نتیجه‌ی بررسی‌ها با عنوان پیشنهاد مقدماتی برای کد استاندارد و تبادل اطلاعات در فارسی، در خرداد ماه ۱۳۵۷ انتشار یافت. در این گزارش، نشانه‌های فارسی به صورت تک‌نمادی<sup>۱</sup> و برپایه‌ی جدول کد ASCII پیشنهاد شده بود و نشان‌دهنده‌ی دوراندیشی و آینده‌نگری پیشنهاددهندگان بود. اما در طی سال‌های ۱۳۵۸-۱۳۵۹، با توجه به محدودیت‌های نرم‌افزاری و سخت‌افزاری آن زمان، اشکالاتی برای این جدول عنوان گردید و جدولی دگرگون‌شده و دونمادی ارائه گردید (- سازمان برنامه و بودجه، ۱۳۵۸).

با پیدایش رایانه‌های شخصی و افزایش توانمندی‌های آنها از یک‌سو، و افزایش شمار برنامه‌سازان و گونه‌گون‌تر شدن کدهای دلبخواهی از سوی دیگر، بار دیگر نیاز به استاندارد کارآ و روزآیند احساس شد. از این‌رو، مؤسسه‌ی استاندارد و تحقیقات صنعتی ایران در سال ۱۳۶۷ با تشکیل کمیسیونی از نمایندگان سازمان‌های رایانه‌یی، استاندارد با شماره‌ی ۲۹۰۰ و با همان شیوه‌ی دونمادی تدوین می‌کند (- مؤسسه استاندارد، ۱۳۶۸) که به‌زودی عدم‌کارایی آن آشکار می‌شود. در نتیجه، در سال ۱۳۶۹، کمیسیون استاندارد کد تبادل اطلاعات هشت بیتی فارسی متشکل از کارشناسان داده‌پردازی، استاندارد و پژوهشگاه علوم انسانی بررسی‌های خود را آغاز می‌کند. این کمیسیون، پس از دو سال، جدولی را ارائه می‌دهد که در آذرماه ۱۳۷۱ در کمیته‌ی ملی استاندارد کامپیوتر با نام استاندارد کد تبادل اطلاعات هشت بیتی فارسی به تصویب می‌رسد و در خردادماه ۱۳۷۲، با شماره‌ی ۳۳۴۲ استاندارد ایران منتشر می‌گردد (- مؤسسه استاندارد، ۱۳۷۲).

## ویژگی‌های استاندارد ۳۳۴۲

برخی از ویژگی‌ها و سودمندی‌های این استاندارد عبارت‌اند از:

۱. استاندارد برای کدهای هشت بیتی تدوین شده است، زیرا مجموعه کدهای شش یا هفت بیتی با ۶۴ یا ۱۲۸ امکان، دیگر پاسخگوی نیازهای پردازش پیشرفته‌ی متن‌های فارسی نیستند و ۲۵۶ جایگاه مجموعه‌ی هشت بیتی به ما امکان می‌دهد همه‌ی نویسه‌ها و نشانه‌های مهم نگاشتاری فارسی را در دسترس داشته باشیم؛ به‌ویژه آنکه به خلاف رایانه‌های قدیمی، دستگاه‌های داده‌پردازی جدید، اطلاعات خود را در محیط‌های هشت بیتی ذخیره، بازیابی و مبادله می‌کنند.

۲. در این استاندارد، رهیافت تک‌نمادی<sup>(۳)</sup> به کار گرفته شده است که خود دارای برتری‌های بسیاری است. نخست آنکه برای هر نشانه فقط یک کد در نظر گرفته شده است و به هیچ روی، نشانه‌ها به شکل‌های گوناگون نویسه‌ها و مشکلات جانبی خط فارسی وابسته نیستند. به سخن دیگر، این رهیافت برخوردی محتوایی با نشانه‌ها دارد و با اصل ساده‌سازی منطقی استاندارد همسواست. روش‌های گوناگون چندنمادی بیش‌تر به دلیل محدودیت‌های نرم‌افزاری و سخت‌افزاری گذشته پدید آمده و مشکلات بی‌شماری را در مبادله‌ی اطلاعات میان افراد، سازمان‌ها و مراکز رایانه‌یی ایجاد کرده بودند. اکنون همه‌ی دستگاه‌های دروندادی، پروندادی و پردازشی می‌توانند به سادگی شکل‌های متفاوت حروف را، با توجه به جایگاهشان در متن، مشخص نمایند.

۳. در این مجموعه برای همه‌ی نشانه‌های خط فارسی از جمله نشانه‌های زیروزیری، همزه، نشانه‌های ویژه و علائم ریاضی جایگاهی مستقل پیش‌بینی شده است در نتیجه، دیگر نباید نگران از میان رفتن یا گم شدن بخشی از داده‌ها باشیم.

۴. در موارد خاص، امکان ایجاد صورت‌ها و ترکیب‌های نوشتاری دلخواه و بیرون از منطق اتصال حروف فارسی نیز هست؛ یعنی با بهره‌گیری از دو نشانه‌ی فاصله‌ی مجازی<sup>۱</sup> و اتصال مجازی<sup>۲</sup> می‌توان انواع پیوسته‌نویسی یا جدانویسی را رعایت کرد.

۵. این استاندارد بر پایه‌ی جدول کد ASCII و هماهنگ با گسترش کد هشت بیتی

1. Pseudo space

2. Pseudo connection

سازمان جهانی استاندارد<sup>۱</sup> طراحی شده است و توانایی همپایی با استانداردهای دیگر را دارد.

۶. این استاندارد برای تبادل داده‌های فارسی در همه‌ی سیستم‌های رایانه‌یی مانند ریزرایانه‌ها<sup>۲</sup>، رایانه‌های متوسط<sup>۳</sup>، رایانه‌های بزرگ<sup>۴</sup>، شبکه‌های انتقال اطلاعات محلی و سراسری<sup>۵</sup>، پایانه‌های هوشمند<sup>۶</sup>، رسانه‌های ذخیره و نمایش اطلاعات<sup>۷</sup> و مانند اینها کاربرد دارد.

۷. تدوین و تصویب این استاندارد راه را برای ایجاد استانداردهای دیگر که به گونه‌یی وابسته به آن هستند هموار کرد. همچنین طراحان و سازندگان نرم‌افزار و سخت‌افزار نیز از این پس می‌توانند با اطمینان بیش‌تری جنبه‌های مختلف خط و زبان فارسی را در تولیدات خود اعمال نمایند و مهم‌تر اینکه کاربران نیز از سرگردانی و صرف وقت رهایی می‌یابند.

البته تحقق کامل این امر نیاز به زمان و احساس مسئولیت دارد؛ چرا که گذشته از مشکل کنار نهادن عادات پیشین، ایجاد تغییرات لازم زمان‌گیر و هزینه‌بر است. اما تردیدی نیست که پیروی از این استاندارد به سود همه و به صلاح زبان فارسی خواهد بود. (۴)

## ۲. استاندارد ملی صفحه‌کلید فارسی در کامپیوتر

پیش از ورود رایانه به ایران، ابزار مکانیکی نوشتن متن فارسی ماشین تحریر بود که با فشردن کلیدهای آن، حروف گوناگون بر کاغذ نقش می‌بست. از آنجا که ماشین تحریر فارسی پس از عربی ساخته شده بود، برخی از ویژگی‌های آن مناسب خط و زبان عربی بود از جمله شکل ظاهری حروف، طرز پیوند آنها با یکدیگر و تعداد نشانه‌ها و آرایش و ترتیب قرارگرفتن آنها روی صفحه‌کلید. دو مورد نخست بیش‌تر به روش نگارش و خوش‌نویسی مربوط می‌شد، اما موارد دیگر عواملی چندگانه داشت. در این

1. ISO

2. micro computers

3. mini computers

4. main frame

5. LAN and WAN

6. intelligent terminals

7. information storage and display media



ماشین تحریرها، محدودیت تعداد کلیدها باعث کنار نهادن برخی از نشانه‌ها (مانند نشانه‌های زیر و زبری یا نشانه‌های ویژه) گردید و جای حروف گوناگون را در موقعیت‌های نزدیک یا دور از انگشتان فعال و بسامد آنها در متن‌های عربی تعیین می‌کرد. در نتیجه الف و لام (ال)، که پرکاربردترین تکواژ عربی را می‌سازند، در میان صفحه و درست زیر انگشتان اشاره جای گرفتند. در به کارگیری این صفحه کلید برای تایپ فارسی تنها اقدامی که صورت گرفت، افزودن حروف خاص فارسی یعنی پ، چ، ژ، گ بود.

ورود رایانه‌های اولیه با محدودیت‌های پردازشی، نمایشی و چاپی خود وضعیت را دشوارتر ساخت. تحولات و پیشرفت‌های بعدی نیز نه تنها کمکی نکرد، بلکه منجر به پیدایش صفحه کلیدهای متعدد و آشفتگی بیش‌تر گردید تا جایی که هر سازمان یا شرکت رایانه‌یی صفحه کلید را به سلیقه‌ی خود تنظیم می‌کرد و حتی برخی از آنها برای کاربردها یا نرم‌افزارهای متفاوت آرایش کلیدهای جداگانه‌ی تک‌نمادی، دونمادی، و چندنمادی داشتند.<sup>(۵)</sup> نتیجه‌ی این هرج و مرج دشواری درونداد، ناسازگاری داده‌ها، و بی‌اعتنایی سازندگان سخت‌افزار به بازار ایران برای تولید صفحه کلیدهای ویژه‌ی فارسی بود.

تصویب استاندارد کد تبادل اطلاعات هشت بیتی فارسی لزوم ایجاد استاندارد را برای صفحه کلید فارسی رایانه مطرح ساخت. در حقیقت، استاندارد صفحه کلید مکمل استاندارد کد تبادل اطلاعات و پشتیبانی‌کننده‌ی آن است، همان‌گونه که خود نیز نماینده و زیرمجموعه‌ی آن به شمار می‌رود. کمیسیون استاندارد کامپیوتر با بررسی صفحه کلیدهای موجود و استاندارد دونمادی که پیش‌تر ارائه شده بود و با در نظر گرفتن استاندارد ۳۳۴۲ و چندین معیار دیگر، در شهریورماه ۱۳۷۳، صفحه کلیدی را طراحی و پیشنهاد کرد که در اجلاس کمیته‌ی ملی استاندارد کامپیوتر به تصویب رسید و در آذرماه همان سال، با شماره‌ی ۲۹۰۱ به عنوان استاندارد رسمی ایران منتشر شد.

ویژگی‌های استاندارد ۲۹۰۱

مهم‌ترین ویژگی این استاندارد همپایی آن با استاندارد کد تبادل اطلاعات هشت بیتی

۲۴۴۶ است. نتیجه‌ی این همپایی، پیدایش ساده‌ترین و آسان‌ترین صفحه‌کلید، یعنی صفحه‌کلید تک‌نمادی فارسی است که کار درونداد متن را بسیار راحت و سریع می‌سازد. پیامد دیگر آن، کاهش نشانه‌های تکراری و باز شدن جای بیش‌تر برای نشانه‌های ویژه مانند نشانه‌های زیرو زبری، تشدید، تنوین، نشانه‌ی اختصاصی برای ریال و چند نشانه‌ی دیگر است، به گونه‌یی که همه‌ی نمادهای جدول‌های استاندارد ۳۳۴۲ بر روی کلیدها جای گرفته‌اند.

در آرایش نمادها بر روی صفحه‌کلید نیز نکات مثبتی وجود دارد، از جمله اینکه با بررسی همه‌ی صفحه‌کلیدهای پیشین و نیز بررسی آماری و بسامدی حروف فارسی در متن‌های متعارف، تا حد امکان جایگاه‌های آشنا و مشترک حفظ شده‌اند. حرف‌های پربسامد - برای صرفه‌جویی در حرکت انگشتان دست - در جایگاه‌های میانی، حروف مشابه نزدیک به هم و اعداد و نشانه‌های ویژه در ردیف بالا قرار گرفته‌اند. همچنین همگونی میان نشانه‌های ویژه‌ی فارسی و لاتینی تا جایی که ممکن بوده رعایت شده است. برای نمادهای فاصله‌ی مجازی و اتصال مجازی PSP و PCN جایگاه‌های مشخصی پیش‌بینی شده است که در مورد فاصله‌ی مجازی می‌توان از shift+space نیز استفاده کرد.

نیاز شدید به این استاندارد از آنجا نمایان می‌شود که در فاصله‌ی کوتاهی پس از اعلام آن، اغلب نرم‌افزارهای ایرانی و برخی از خارجی‌ها آن را به کار بردند و برچسب‌های فارسی با عنوان JSIRI 2901 در تیراژ بالا تولید شد و جانشین انواع دیگر گردید. اکنون که بیش‌تر ماشین‌های تحریر الکترونیکی و هوشمند هستند، امکان بهره‌گیری از این استاندارد در آنها نیز پدید آمده است.

### ۳. استاندارد جهانی نویسه‌گردانی خط فارسی

همواره نمایش شکل نوشتاری و تلفظی نام‌های خاص جغرافیایی و اشخاص و اصطلاحات فرهنگی فارسی در متن‌های غیرفارسی و به‌ویژه اروپایی مشکل و نابه‌سامان بوده است. تنها نگاهی به چند دایرة‌المعارف مانند ایرانیکا، اسلام، بریتانیکا و

یا برخی از کتاب‌های تاریخ، تاریخ ادبیات و مانند اینها نشان می‌دهد که چه شیوه‌های متفاوتی برای نمایش وجود دارد. اغلب آنها آمیزه‌ی نامشخصی از نویسه‌گردانی و آوانویسی و نماینده‌ی مخلوطی از گونه‌های کهن و امروزی فارسی و حتی عربی هستند. در زمینه‌ی کتابداری و اطلاع‌رسانی که نام‌ها، عنوان‌ها و کلید واژه‌ها در مرتب‌سازی، جست‌وجو و بازیابی اطلاعات نقش اساسی دارند، مسئله جدی‌تر می‌شود و هنگامی که صحبت از رایانه و خودکارسازی این فعالیت‌ها به میان آید، این ناهماهنگی‌ها کار را ناممکن می‌سازد.

سازمان جهانی استاندارد (۱۹۹۷) برای حل این مشکل، تدوین یک رشته استانداردهایی را برای نویسه‌گردانی دستگاه‌های نوشتاری آغاز کرده است. هدف از این کار فراهم ساختن وسیله‌ی برای تبادل بین‌المللی پیام‌های نوشتاری به شکلی است که انتقال و بازسازی خودکار آنها را توسط انسان و ماشین ممکن سازد. در این صورت، سیستم تبدیل باید یکدست و کاملاً برگشت پذیر باشد. از آنجا که نخست برای زبان و خط عربی استاندارد نویسه‌گردانی تدوین شده بود و بسیاری از کارشناسان کشورهای مختلف متوجه شدند که خط فارسی نیاز به استانداردی جداگانه دارد، مؤسسه‌ی استاندارد ایران مسئول تهیه‌ی آن گردید. کمیته‌ی فنی ایرانی با شرکت کارشناسانی از مؤسسه‌ی استاندارد، کتابخانه‌ی ملی، فرهنگستان زبان و ادب فارسی و پژوهشگاه علوم انسانی و برخی سازمان‌های دیگر در سال ۱۳۷۴ تشکیل شد و پس از بررسی‌های بسیار، پیشنهاد خود را برای بررسی کشورهای دیگر عضو این کمیته فرستاد. نظریات هیأت‌های هفده کشور، پس از دریافت، بررسی گردید و به جز نظرهای کاملاً موافق، دیگر دیدگاه‌ها یا اعمال گردید و یا به طور مستدل پاسخ داده شد. پس از یک دوره رأی‌گیری، پیشنهاد نهایی در اردیبهشت ماه ۱۳۷۶، در اجلاس سالانه‌ی سازمان جهانی استاندارد به تصویب رسید.

### ویژگی‌های استاندارد جهانی 233-3

در این استاندارد برای هر نماد نوشتاری فارسی یک معادل لاتینی تعیین شده است و نشانه‌های زیروبری و تئوین و تشدید و برخی از نشانه‌های ویژه نیز در نظر گرفته

شده‌اند. هرچند کوشش شده تا جایی که ممکن است مشابهت و تداعی آوایی نیز وجود داشته باشد براساس یکی از اصول کلی استانداردهای نویسه‌گردانی، لزومی نیست که نویسه‌ها نشان‌دهنده‌ی تلفظ، معیارهای زیبایی‌شناختی یا عادات‌های ملی باشند، اما این استاندارد هیچ مغایرتی با استانداردهای ملی و روش‌های سنتی ندارد و ویژگی اصلی این استاندارد سادگی و بی‌ابهامی آن است و اینکه به همان سادگی که می‌توان متن فارسی را به‌طور خودکار به نمادهای لاتینی تبدیل کرد، همانطور نیز می‌توان آن را به خط فارسی برگرداند. البته تلفظ درست برای کسانی ممکن است که با زبان فارسی آشنا باشند. این استاندارد سیستم ساده‌شده‌ی نویسه‌گردانی فارسی به لاتینی به شمار می‌رود و به‌ویژه برای پردازش اطلاعات کتابخانه‌یی در نظر گرفته شده است.

### استانداردهای درست بررسی

تدوین استاندارد فرایندی دایمی است و هرگونه تغییری که در پدیده‌ها (چه علمی، چه صنعتی و چه فرهنگی) به‌وجود آید، تجدیدنظر در استانداردهای پیشین یا تدوین استاندارد تازه را ایجاب می‌کند. استاندارد کد تبادل اطلاعات هشت بیتی می‌تواند تا سال‌ها در بخش بزرگی از دنیای داده‌پردازی کاربرد داشته باشد، اما از مدتی پیش، کارشناسان و پژوهشگران درصدد ایجاد نظامی نوین در این رشته‌اند تا با توجه به توانایی‌های روزافزون رایانه بتوانند هزاران نشانه‌ی نگاشتاری مربوط به همی زبان‌های طبیعی و ساختگی جهان را در آن جای دهند.

یکی از این فعالیت‌ها در سازمانی با نام یونی‌کد<sup>۱</sup> انجام می‌گیرد که کنسرسیومی متشکل از چندین شرکت بزرگ رایانه‌یی است و مجموعه کد شانزده بیتی را مطرح کرده است. چنین مجموعه‌یی می‌تواند تا ۶۵۵۳۶ نماد را در خود جای دهد. این کنسرسیوم اعلام کرد که آمادگی دارد کدهای پیشنهادی کشورهای مختلف را، با رعایت اصول کلی خود، در این مجموعه بگنجانند. در ایران کمیته‌یی برای بررسی این مسئله تشکیل شد و پس از یک سال، نتایج بررسی‌های خود را ارائه داد. پیشنهاد عمده‌ی این کمیته به

1. UNICODE: Universal Code

یونی کد، در نظر گرفتن مجموعه کدهای مستقل برای فارسی و بر مبنای استاندارد هشت بیتی است.

از استانداردهای دیگر در دست بررسی، استاندارد آوانویسی فارسی است که می تواند مکمل نویسه گردانی باشد. همچنین استاندارد برای نشانه گذاری متن های فارسی مورد نیاز است که خود بحث مفصلی را می طلبد.

### پی نوشت ها

۱. برای جزئیات بیش تر ← عاصی، ۱۳۷۲.
۲. برای نمونه، در بررسی مشکلات استاندارد تبادل اطلاعات (← خبرنامه انفورماتیک، ۱۳۷۲)، ۱۸ جدول مختلف از سوی شرکت های گوناگون رایانه پی ارائه شده است.
۳. برخی از مزایای به کارگیری شیوهی تک نمادی: عبارت اند از: کاهش تعداد نشانه ها، آسانی و سرعت درون داد متن، سازگاری سیستم های نمایش و چاپ، عدم وابستگی به نرم افزار و سخت افزار خاص، پرهیز از شیوه های متعدد و یا چند نمادی، و مانند اینها.
۴. تاکنون بسیاری از شرکت های نرم افزاری ایرانی و خارجی به کارگیری این استاندارد را آغاز کرده اند.
۵. برای ملاحظه ی چندین نمونه از صفحه کلیدها (← خبرنامه انفورماتیک، ۱۳۷۲).

### کتابنامه

- خبرنامه انفورماتیک. ۱۳۷۲ (اردیبهشت و تیر). ش ۵۳. صص ۱۸۳-۱۹۴.
- سازمان برنامه و بودجه. ۱۳۵۸ (شهریور). پیشنهاد نهایی کد تبادل اطلاعات در فارسی.
- عاصی، مصطفی. ۱۳۷۲. «بار دیگر خط فارسی؟»، آدینه، ش ۸۸. صص ۵۸ و ۵۹.
- مؤسسه استاندارد و تحقیقات صنعتی ایران. ۱۳۶۸ (مرداد). کد تبادل اطلاعات به زبان فارسی، ش ۲۹۰۰.
- \_\_\_\_\_ ۱۳۷۲ (خرداد). استاندارد کد تبادل اطلاعات هشت بیتی فارسی، ش ۳۳۴۲.
- \_\_\_\_\_ ۱۳۷۳ (آذر). طرز قرار گرفتن حروف و علامت زبان فارسی بر روی صفحه کلید کامپیوتر، ش ۲۹۰۱.



شروېشگاه علوم انسانی و مطالعات فرهنگی  
پرتال جامع علوم انسانی