آشفتگی نگارش نام پدیدآورندگان ایرانی در پایگاه اطلاعاتی آی اس آی

مریم خسروی
عضو هیئت علمی، پژوهشگاه علوم و فناوری اطلاعات ایران

چکیده: ضبط اسامی ایرانی در منابع خارجی دچار نشست و چندگانگی است و از قاعده خاصی پروری نمی‌گردد. پژوهش‌های تخصصی - بی‌پایان است، به‌منظور بررسی آشفتگی نام و نگارش اسامی مؤلفان ایرانی، کمک‌کشی در پایگاه اطلاعاتی آی‌اسآی، طی سال‌های 2007 و 2008 داده‌است. در این راستا، فهرست کامل از آثار ایرانی در پایگاه اطلاعاتی Web of Science برای سال‌های 2007 و 2008 استخراج شد. از این‌ها، با نگارش متن‌ها، جامعه آماری مقالات چاپ شده، ایرانیان و درصد اسامی مؤلفان ایرانی، نگارش آشفتگی، نگارش اسامی در پایگاه اطلاعاتی آی‌اس آی، بود. نتایج نشان داد که نگارش اسامی مردان بیش از زنان دارای نوبت نگارش بوده است. همچنین، بیشترین میزان نوبت نگارش (در یک سوم آنها) مربوط به کوشش‌های نام بوده است.

کلیدواژه‌ها: نام مؤلفان ایرانی، آشفتگی، نگارش، پدیدآورندگان، پایگاه اطلاعاتی آی‌اس آی
1. مقدمه

نام مؤلف در تمامی پایگاه‌های اطلاعات کتابشناسی، یکی از مهم‌ترین گرینه‌های جستجو است و همیشه برای هر یک از چنین منابع و پایگاه‌های وابسته مدل‌سازی می‌شود. هر پایگاه و اجزای یک از منابع "استفسار" از نظر روش و اجزای مختلف فناوری جستجو سانه‌ها است. البته عنوان که کاربران با کافی‌گیری اطلاعات در پایگاه‌ها، همیشه باعث هم‌آفرینی اطلاعات، نام مؤلف برخی‌ها را را در آن جدوله موضوعی به شناسند و سپس به دنبال تایید آن مؤلف می‌گردد. با حوزه کاری خود را که توسط وی پیشرفتی تحریک در آن را است بیابد. جستجوی نام مؤلفان فقط برای پایگاه منابع علمی به فضای مطالعه و پژوهش کاربرد ندارد، بلکه کاربردهای دیگری نیز دارد، ازجمله در بررسی‌های کمی که در علم‌های به آزمایی یک نویسنده برای پاسخ‌های موضوعی، با میزان استناد به آن‌ها، محاسبه شاخص‌های مثل شاخص هرز، شاخص ج، انجام می‌شود، از نام مؤلف استفاده فراوان می‌شود.

نام مؤلف در ترتیب نیز، هرگونه تنش و تنوع در نگارش اسامی مؤلفان می‌تواند مشکل‌آفرین باشد. این نوع نیز در مورد اسامی نویسندگان ایرانی در پایگاه‌های اطلاعاتی آپاس آن وجود دارد و مشکلاتی را هم ایجاد نموده است. مواجهه به نامی که امکان‌ها تأسیس آن به چند صورت امکان‌دار است، با برازیان آن مدرک را دچار اختلال می‌کند و از طرف دیگر، نمی‌توان مدت‌یافته که برای استفاده به فلان مدرک کاربر تمامی شکل‌های احتمالی نگارش نام نویسنده آن را در ذهن داشته باشند و در نتیجه بخاطر به‌عده‌های بی‌ضبط گوناگون رجوع، کننده صرف‌نظر از اینکه نیازی به تایید آنها نیست ناشی نمی‌شود.

نام باعث مثال، نام یکی از ارسال‌های دانشمندان ایرانی، با 14 املای مختلف در پایگاه اسکوپوس و نیز با املاه‌های گوناگون در پایگاه‌های آپاس آن تابش شده است (مهمان‌زاده 1388). بنابراین، جستجو برای یکی از این نام‌ها، کافی‌گیری اطلاعات این نام مؤلف هدایت نمی‌کند. حاصل این نوع در نگارش و نیز امکان‌های تایید مشکلات عدیده دیگری را ایجاد می‌کند، از جمله اشتباه در محاسبه شاخص‌های علم‌سنجی نظیر شاخص هرز و نمای صنعتی کاربرد مناسب‌تر از طرف دیگر که به دانشمندان مربوط می‌شود (نظریه استادانه، آثار، و غیره).

از طرف دیگر از لحاظ قانون‌ها، تاریخ‌هایی ضبط اساسی از زبان فارسی، در متون و مدخل‌ها یکی از ارزش‌های کتابشناسی باعث سردرگمی و آشکار شدن نگارش می‌شود. (تصوق بهزادی 1381) (بعاونان مثل، برای نوشتن نام مهدی به زبان لاتین یک نفر آن را با خصوصیت

1. Stars Search
2. Hirsch Index
3. ISI
4. scopos
آنلاین نگارش نام بیدرودی‌های ایرانی در... | خسرو مهدی و دیگر

مشکلاتی است که بر سر این پژوهشگران قرار دارد. مواجهه با نامی که قبلاً جستار کردند، نجات به عدم توقف در پژوهشکده می‌شود.

مالک لیست اینک به توصیه نگارش فقط مختلف نام مؤلفان و برگردانند آنها به زبان انگلیسی نیست. این امر درباره ضبط دیگر عراص و برگردانند آنها به زبان فارسی نیز مطرح است. آنها که دیگر متخصصان عناوین کردند مسأله چشمتگر دیگری، ضبط نادیده اعلام شده است. بیشتر متخصصان ایرانی در حین برگردان اسامی شبه قاره از حروف لاتین به حروف فارسی دچار خطای شده‌اند. این خطای ناشی از اختلافهای در ضوابط آوانگاری شده قاره و ایران است. نیود آگاهی کافی در متخصصان ایرانی از صنایع مخصوص الفبا اردو/ هندی و تطبیق دادن نظام آوانگاری ایران با نظام شبه قاره موجب خطا می‌شود. بنابراین به

یادداشتی از نظام آوانگاری شبه قاره به‌سرعت زیر است:\n
الف) برای تصفیه الگوی نظام آوانگاری شبه قاره به‌بان ترتیب است: به = په = th = hp = thh = که = gh = k = v یا th = شده = ۴ یا th = باید ثبت هر gh با یک گاه مبنای است، متخصصان ایرانی را دچار خلط و خطا می‌کند.

ب) برای فتحه "أ" و الگوی محدود و الگو "ه" به‌طور معمول از یک آوا بینی "ا" استفاده کرده که در برگردان فارسی عدم تمیز آنها از یکدیگری موجب اشتباه شده است. گاهی فتحه "ب" و "پ" نیز در استفاده نشده است. مثال: "مدراس" (مدرس. الیه آگاهی در نظام دیق آوانگاری شبه قاره برای تمیز

الف محدود فتحه از آوانگاری "ا" استفاده شود "م")

ج) برای "کش" از "ک" استفاده می‌شود که در برگردان فارسی گاهی به "ک" تبدیل

شده است. مثال: "پیشاور"<PRESHAWAR>

در "ب" رای ضمه "ب" از "ب" استفاده می‌شود که متخصصان ایرانی آن را "و" خلط

ب) برای قرآن "آ" استفاده می‌شود که متخصصان ایرانی "آ" را "و" خلط

کردند. مثال: "قلهان"<MULTAN>

پ) برای قرآن "آ" استفاده می‌شود که متخصصان ایرانی "آ" را "و" خلط

با توجه به این مقدار و مشکل موجود در مورد نگارش لاتین اسامی نویسنده‌های ایرانی،

این پژوهش سعی دارد تا آشفتگی نگارشی اسامی پیداکردن ایرانی را در پایگاه آی‌اس آی مورد مطالعه قرار دهد و میزان و چگونگی این آشفتگی را مشخص نماید.

در همین راستا، این پژوهش سعی دارد به سوالات زیر پاسخ دهد:
1. تعداد مقالات چاپ شده ایرانیان در این دوره 2007 و 2008 چقدر بوده است؟
2. تعداد مؤلفان ایرانی در دو مقاله در این پایگاه اطلاعات در سال‌های 2007 و 2008 چگونه است؟
3. پراکنده‌گی جنسیتی مؤلفان مقالات ایرانی در این دوره بوده است؟
4. چه میزان از اسامی مؤلفان ایرانی در دو پایگاه اطلاعات آی‌اس آی هستند؟
5. چه میزان از اسامی (که دارای آشکاری نگارش هستند) مناسب به اسامی زنان و چه میزان آن مربوط به اسامی مردان نویسندگه است؟
6. نوع آشکاری نگارش اسامی چگونه است؟

2. پیشنهاد پژوهش

تا جهت که نگارندگی امضاء در این مورد پژوهشی، چچ در زبان فارسی و چه در زبان این از طرف زبان‌شناسان و کتابداران) انجام گرفته و با حضور این پژوهش منابع و بعید است. دلیل آن تصور شد که در زبان‌های دیگر، به‌ویژه زبان‌هایی که اکثریت یا همه زبان‌های فارسی در این باره مشکلی نیست که نیاز به پژوهش داشته باشند. مهم‌ترین امکانات دارد پژوهشی به‌زبان ملی و نه انگلیسی، برای بررسی این اشکال مسائلی می‌باشد. بررسی نگارش اسامی مؤلفان به زبان انگلیسی به‌طور منطقی فقط در زبان‌هایی رخ می‌دهد که ناگیرا از آوانویس اسامی خود به ادامه زبان انگلیسی هستند. در این میان، زبان‌های فارسی و عربی با اشتباه به‌دلیل استفاده خاص خود از حروف صامت و سفارش و حروف صدا و عدم وجود استاندارد در مورد آوانویسی شرایط خاصی دارند و وجود صدای و حروف که که در زبان‌های فارسی شده و که نیاز به مسیر می‌شود، یکی از دلایل بررسی این مشکل است که این امر باعث شود یکی یا دویکی به‌دلیل سلیقه خود دست به برگرداندن نام خود به زبان انگلیسی بزند.

در جستجوی م有信心ی نیز پژوهش‌هایی درباره چگونگی ضبط زبان ایفا، صدای بی‌هست (۱۳۸۱) کتاب می‌بیند که هم‌اکنون سمینار داده شود و نیز امکان تغییری جغرافیایی در منابع فارسی نشان داده شود و نیز امکان تغییری جغرافیایی در منابع فارسی نشان داده شود و نیز امکان تغییری جغرافیایی در منابع فارسی نشان داده شود و نیز امکان تغییری جغرافیایی در منابع فارسی نشان داده شود و نیز امکان تغییری جغرافیایی در منابع فارسی

1. Names disambiguation  
2. Ambiguities names
که تا حد کمی به این موضوع مرتبط است، انجام شده است که به‌طور مختصر، به برخی از آنها اشاره می‌شود:

اسامی ابهام‌دار، مسأله چالش برانگیز است که دارای سابقه تاریخی طولانی است. در زمان مذکور اطلاعات ذیل واقع، با مسأله به‌نام ابهام‌داری اسامی مواجه هستیم. ابهام اسامی ممکن است به دو دلیل رخ دهد: شخصی چند نام داشته باشد و یا چند نفر یک نام داشته باشد.


در حیطه کتابخانه‌های دیجیتال، بی‌کردن اسامی ابهام‌دار در استاده‌های کتابشناسی، امری عادی است. ابهام ممکن است به‌علت وجود چندین مؤلف با اسامی مشابه و یا اسامی متفاوت برای یک مؤلف پیش‌پایید. به‌طور اخیر، در مسأله ابهام‌دار مکرر و ابسانده‌های شناختی ممکن است در این رابطه صورت پیگیرد. استاده‌های مکرر به این صورت است که چند مؤلف، یک نام و یا نام‌های یک نام‌دهنده، در هنگام تلفظ مشابه به‌همنام باشند، درمانند. برای مثال، دی پریرا، دی پریرا (Masada and Adachi 2007) از نظر هنگام تلفظ جزئی نزدیکی ممکن است به‌صورت متفاوت بوده و ممکن است به‌صورت متفاوت "دنیلسون پری" و "دی پریرا" می‌باشد. (Pereira et al. 2009)

 belongings ممکن است بر کیفیت گرد آوری اطلاعات علمی تأثیر بگذارد و با می‌تواند باعث کاهش عملکرد بازیابی اطلاعات و جستجو در وب باشد؛ حتی ممکن است باعث شناسایی و استاده‌ای اختصار نقش دارد تا این مؤلفان شود. این مسئله در مسئله درست به‌عنوان شاخصی از یک مؤلف گوناگونی اسامی، اسامی همانند، تلفظ غلط اسامی، و یا نام‌های مساعار. دو نوع ابهام اسامی را در مقالات پژوهشی و کتابشناسی های (استاده‌ها) می‌توان مشاهده نمود. نوع اول اینکه یک نام ممکن است به‌عنوان "دنیلسون پری" و "دی پریرا" می‌باشد. (Han, Zha, and Giles 2005) از داشته‌گاه رایس ارجاع داده شد و با دی و پریرا، دی پریرا، دی پریرا از آزمایشگاه تحقیقات AT & T و یا "دی پریرا، دی پریرا" از داشته‌گاه پویا باشد.


روش پژوهش. جامعه آماری، و روش گردآوری داده‌ها

پژوهش حاصل یک مطالعه توصیفی - پیش‌پایی است. جامعه پژوهش شامل همه

3. روشن نشاند (2009)
نویسندگان و مؤلفان ایرانی است که مقاله‌ای با مقاله‌های خود را در سال‌های 2007 و 2008 در پایگاه اطلاعات آگ‌اس آی به‌جای رسانده‌اند. تعداد کل آن‌ها پس از اعمال محذوفیت مورد نظر ( فقط مقاله و افراد دارای نام کوچک)، طبق رکورد‌ها به‌تستی در این پایگاه اطلاعات و در این پروژه زمان‌شماره 2999492 رکورد پیدا کرد. در این پژوهش، اکثر جامعه آماری مورد مطالعه قرار گرفت. لازم به اشاره است که افزایش هم‌کننده، همکاری و صدور کننده، نکته مختصری برای شده بود.

ولی امکان جستجو و تعیین جنسیت آن‌ها از طریق اینترنت ممکن بود، در این بار قرار گرفتند.

برای گردآوری اطلاعات، ابتدا فهرست تمام اشارات ایرانیان در پایگاه اطلاعاتی Web of Science سال‌های 2007 و 2008 به صورت tagged format (marked list) در قالب دسته‌های پاینتایی استخراج شد. سپس، این رکوردها در همکرد شد و با استفاده از نرم‌افزار Access 2007، مولفان و اناضیفگی آن‌ها تاکنون (برای افراد دارای چندین مقاله و یا دایر کننده تایبی) آشنایی (تاوان)، مشاهده می‌شود، چرا که در هر حرف n جمله نام آنها درج شد. در مرحله بعد، برای هر یک از آن‌ها تعیین جنسیت (به‌صورت نام کوچک) انجام شد.

لازم به ذکر است که در این مرحله، گاهی به‌عنوان تاریخ نویسندگی از اسهام‌های محقق (مانند حسنی، شاهین و یا نام و نامی) که صدور و اچ‌نگاری مشابه دارند، از جستجو در اینترنت استفاده شد. در مرحله آخر، برای هریک از آن‌ها کد مشخصی گرفت. در این مرحله، در صورتی که نام فرد مورد بررسی (دارای چندین مقاله) دارای آشنایی نگارش بود، کد تعیینی برای نام دارد توانایی نگیری تعیین می‌شود.

برای بررسی آشنایی نوع نگارش، 10 از اسهام‌های گوناگونی نگارش به عنوان نمونه، مورد مشاهده قرار گرفتند. تا توجه آشنایی نگارش اساسی آن‌ها مشخص شود. در پایان، اطلاعات جمع‌آوری شده با استفاده از روش‌های آماری توصیفی، چون توزیع فراوانی و درصد، مورد تجزیه و تحلیل قرار گرفتند.

4. نتایج

1- بررسی وضعیت مقالات جامعه آماری پژوهش در دستورالعمل 1 مشاهده می‌شود:

1. Loet Leysdorff
جدول ۱. جامعه آماری بررسی جنسیت

<table>
<thead>
<tr>
<th>جنسیت</th>
<th>مطالعات چاپ شده</th>
<th>تعداد</th>
</tr>
</thead>
<tbody>
<tr>
<td>زن</td>
<td>۲۱/۶۷٪ ۸۷۴۲</td>
<td></td>
</tr>
<tr>
<td>مرد</td>
<td>۷۸/۳۳٪ ۳۱۴۵۶</td>
<td></td>
</tr>
<tr>
<td>جمع</td>
<td></td>
<td>۱۱۹۸۸</td>
</tr>
</tbody>
</table>

آنچنان که در جدول ۱ و نمودار ۱ نشان داده می‌شود، از تعداد ۱۹۹۸۱ مقاله چاپ شده در این پایگاه اطلاعاتی، مقالاتی که توسط مردان در این پایگاه اطلاعاتی به چاپ رسیده است بیش از ۷۵ برابر زنان است. (البته لازم به اشاره است که این نتایج با توجه به محدودیت‌های پژوهش، به‌لحاظ اینکه برای برخی از اسامی مؤلفان، شناسایی جنسیت آنها مقدور نبود، احتمال‌های دارد).

جدول ۲. انتشار مقاله توسط ایرانیان

<table>
<thead>
<tr>
<th>سال انتشار</th>
<th>درصد</th>
<th>جمع</th>
</tr>
</thead>
<tbody>
<tr>
<td>٢٠۰۷</td>
<td>٤۴/۳۳٪</td>
<td>١٧۳۰</td>
</tr>
<tr>
<td>٢٠۰۸</td>
<td>٥٥/۶۷٪</td>
<td>٢٢۴۶۸</td>
</tr>
<tr>
<td>جمع</td>
<td></td>
<td>٣٩۹۹۸</td>
</tr>
</tbody>
</table>

نمودار ۱. مقایسه سال‌های مورد مطالعه و چاپ مقاله توسط ایرانیان

۱-۴۲ بروز جنسیت مؤلفان و سال انتشار
جدول ۳. تعداد مقالات چاپ شده توسط ایرانیان در این دوره سال به‌تفاوت جنسیت آنها

نشان می‌دهد:

۵۱
جدول ۳ تعداد مقالات چاب‌شده

<table>
<thead>
<tr>
<th>جنسیت</th>
<th>سال چاب</th>
<th>۲۰۰۸</th>
<th>۲۰۰۷</th>
</tr>
</thead>
<tbody>
<tr>
<td>زن</td>
<td>۳۷۷۳</td>
<td>۴۹۶۹</td>
<td>۲۱/۲۸</td>
</tr>
<tr>
<td>مرد</td>
<td>۵۹۶۶</td>
<td>۱۷۲۹۹</td>
<td>۷۸/۷۲</td>
</tr>
<tr>
<td>جمع</td>
<td>۹۷۳۹</td>
<td>۲۲۲۹۸</td>
<td>۱۰۰</td>
</tr>
</tbody>
</table>

براساس جدول ۳ و نمودار ۲ در سال ۲۰۰۷ تعداد ۱۷۳۳ مقاله توسط ایرانیان در این پایگاه اطلاعاتی به چاب رسیده است، در حالی که این امر در سال ۲۰۰۸ افزایش داشته است (۲۴۲۴ مقاله).

از تعداد ۱۷۳۳ مقاله چاب‌شده در سال ۲۰۰۷، تعداد ۳۷۷۳ مقاله توسط زنان نگارش شده و سهم مردان در سال ۲۰۰۷ تا حد ۲/۵ برابر زنان بوده است. در سال ۲۰۰۸ نیز ۲۴۲۴ مقاله به چاب رسیده که ۴۹۶۹ مقاله توسط زنان و ۱۷۲۹۹ مقاله دیگر مربوط به مؤلفان مرد بوده است. در کل، از تعداد ۳۹۹۹۸ مقاله به چاب رسیده در این دو سال، ۸۷۶۲ مقاله آن توسط زنان و ۳۱۲۵۶ مقاله توسط مردان (۲/۳) نگارش شده است.

نمودار ۲ مقایسه سال چاب مقاله و جنسیت
4-3. بورسی آنفلتکی: تکارش اساسی

جدول ۴ تعداد مؤلفان ایرانی دارای مقاله در این دو سال و تعداد اساسی دارای آشفتگی نگارش را نشان می‌دهد (برخی از مؤلفان دارای چند مقاله بوهداند):

<table>
<thead>
<tr>
<th>تکارش اساسی</th>
<th>جنسیت</th>
</tr>
</thead>
<tbody>
<tr>
<td>آشفتگی نگارش</td>
<td>چمن</td>
</tr>
<tr>
<td>تعداد</td>
<td>درصد</td>
</tr>
<tr>
<td>۱۹۸۱۴</td>
<td>۱۹۲۸</td>
</tr>
<tr>
<td>عدم آشفتگی نگارش</td>
<td>چمن</td>
</tr>
<tr>
<td>تعداد</td>
<td>درصد</td>
</tr>
<tr>
<td>۱۸۹۶</td>
<td>۱۸۸۸</td>
</tr>
</tbody>
</table>

همانطور که مشاهده می‌شود در کل ۱۷۲۳۳ مؤلف در این دو سال در این پایگاه اطلاعاتی مقالاتی را به‌پایه رسانده‌اند که از این تعداد، ۱۲۸۱ (۷۴٪) مؤلف در نگارش نام و یا نام خانوادگی خود دارای آشفتگی نگارش بوهدان و ۱۴۹۵۲ (۸۵٪) آنها همیشه به یک صورت عمل نموداند.

در مقایسه آشفتگی نگارش، اساسی زنان (۱۲/۱۵٪) به‌پایه کمتر از مردان (۸۸/۸۸٪) بوهد است.

4-4. بورسی نوع آنفلتکی: تکارش اساسی

برای برسی نوع تنوع نگارش اساسی، تعداد ۲۲۰ مورد از نگارش اساسی دارای تنوع، بررسی شد. گوناگونی مشاهده شده در جدول ۵ قابل رزیت است (الزم به اشاره است که یکی از اساسی دارای چندین نوع نگارش نام بوهداند):

<table>
<thead>
<tr>
<th>شناسه نوع</th>
<th>تعداد</th>
<th>کوئونویسی</th>
</tr>
</thead>
<tbody>
<tr>
<td>۱۹۸۱۴</td>
<td>۹۸</td>
<td>غلظ نویسی و حروف صامت</td>
</tr>
<tr>
<td>۲۸/۲۳</td>
<td>۹۶</td>
<td>دوقسمتی نام نام کوچک</td>
</tr>
<tr>
<td>۱۸/۱۷</td>
<td>۵۷</td>
<td>عدم رعایت فصله</td>
</tr>
<tr>
<td>۱۰۰۴</td>
<td>۳۴</td>
<td>حروف صدادار</td>
</tr>
<tr>
<td>۱/۸۰۰</td>
<td>۳</td>
<td>دوقسمتی نام نام خانوادگی</td>
</tr>
<tr>
<td>جمع</td>
<td>۱۰۰</td>
<td>۲۳۲</td>
</tr>
</tbody>
</table>

53
همانطور که ملاحظه می‌شود، 52/29% از گوناگونی های مشاهده شده مربوط به اختصار و کوتاه‌نویسی است. پس از آن، غلط‌های تایپی به میزان 28/06% قرار دارد. بنابراین در حالی است که دو قسمتی بودن نام کوچک نیز می‌تواند باعث 17/12% از گوناگونی ایجاد شده‌باشد. عدم رعایت فاصله و یا گذاشتن خط فاصله بین دو قسمت نام کوچک و یا همان خوانش‌گی نیز باعث 75/11% از نوع نگارش‌ها بوده است. حروف صدادار در انتهای نام نگارش و در انتهای دو قسمتی بودن نام خوانش‌گی باعث 58/0% از نوع نگارش‌ها شده‌اند.

4- انواع آشفتگی‌های نگارش مشاهده شده

در ادامه، جند صورت از آشفتگی‌های نگارش اساسی مشاهده شده، آورده می‌شود:

کوتاه‌نویسی شامل نوشتن فقط حروف اول نام، به عنوان مثال:

Abadpour, A
Abadpour, Arash
Aalami, M
Aalami, Mehran
Bazgirb, Ayoob
Bazgir, Ayoob
Bazrafshan, Afireza

غلط تایپی و حروف صامت شامل:

دو قسمتی بودن نام مانند:

Abassi, Mohammad Reza
Abassi, Mohammad R.

دو قسمتی بودن نام خوانش‌گی همچنین:

Bozorgmehry, Bozarjomehry
Bozorgmehry, B.
Forouzandeh-Moghadam, Mehdi
Forouzandeh, M.

عدم رعایت فاصله و گذاشتن خط فاصله بین دو قسمت نام، شامل:

Barqi, Mohamadali
Abassi, Mohamad- Hassan

حروف صدادار مانند:

Abassi, Massoumeh
Abassi, Massoumeh
Abassi, Shahriar
Abassi, Shahryar

لازم به اشاره است که اینچنین نوع، درباره نام خوانش‌گی نیز صادق بود.

Ebrahimnezhad, Yahya
Ebrahim-Nezhad, Yahya
Ebrahimnezhad, Y

و یا?

Bezaatpour, Abolfazl
Bezaatpour, Abolfazl
بحث و نتیجه‌گیری

نام‌ها کلماتی هستند که از هزاران سال پیش فقط برای شخصی و تمیز افراد از یکدیگر وضع شده‌اند. چگونگی ضبط نام افراد در نوشته‌ها، به‌خصوص در فرهنگ مکانیک و آمده و در کتاب‌های مرجعی نظر دانشجویان با مسائل مسالمه‌کننده به‌طور تجربه و تأثیر مطرح می‌شود. از آنجا که ضبط اسامی ایرانی در منابع انگلیسی دچار تغییر و آشفتگی است و چون در این جوانه‌ای از قواعدی پروری نشده است، هم‌رسی به سیلیقه و دوز خود، ضبط و صورتی‌ها انگلیسی است. انتخاب کرده، در نتیجه یک اسم به دو یا چند شکل نوشته شده است (تصدیق بهره‌داری 1381). این امر باعث شده است که بازیابی تمام مقالاتی که از طرف یک نویسنده به‌جای رسمیه است، قابل بازیابی نباشد و یا به‌خاطر نشان اسامی، جذبین مقاله به یک مولف اصلی می‌شد.

در همین راستا، این پژوهش به بررسی نحوه ثبت اسامی پیدا آورده‌های ایرانی در بانک اطلاعات آی‌اسی در سال‌های 2007 و 2008 پرداخته است تا مشخص کند که چه درصدی از اسامی مؤلفان ایرانی در این پایگاه دارای آشفتگی نگارش هستند.

نتایج نشان داد که از 19998 مقاله جاپان شده ایرانی در این پایگاه اطلاعاتی در این دوره سال، که متعلق به 17233 مؤلف است (البته ممکن است از این تعداد بیشتر باشد، ولی به‌عنوان تشکیل نام، یک نفر درنگ و گرفته شده باشد)، تعداد 14592 نفر از مؤلفان برای نگارش نام خود همیشه به یک صورت عمل نموده و 2817 نام دارای آشفتگی نگارش بوده است؛ همچنین،

منظره‌ای از آشفتگی نگارش نام در این پژوهش، هرگونه اختلاف نوشتاری است، از جمله کوتاهی‌پیش نام و یا نام خانوادگی (برای مثال، برای نوشتن نام اکبر‌نژاد، عطف نام شده است)، گفته‌های خی، خط Ali Reza، سعید، سعید، محمود، محمود، محمود و یا Saeid، Saeed، / Masoud، Masoud، Mahmoud، Mahmoud، Mahmoud سعید... که به صورت های منظره‌ای است. این آشفتگی نگارش در بسیاری از اسامی و نام‌های خانوادگی (به‌عنوان مثال، برای نوشتن نام محمود، محمود، محمود و یا موسی، موسی، موسی، Mousavi، Moosavi، Mousavi، Mousavی) مثال، برای نوشتن نام خانوادگی صفا، موسی، اکبر‌نژاد، برای به‌چند صورت زیر عمل شده است:

Safai, Safaie, Safaiy, Safae
Akbarnegajad, Akbarnezhad
Mosavi, Moosavi, Mousavi, Mousavی)

در کد نگارش و همچنین نظم الفبای آنها، مشکلاتی عدیدی را ایجاد می‌کرد که باعث شد نتایج بار کد نگارش آنها تعوض شد.

55
در بررسی‌ای که به منظور نوع آشفتگی نگارش اسمی بعمل آمده، تعداد ۲۰۰ مورد از نام‌هایی که به عنوان دارای آشفتگی نگارش نشانه‌شده بودند، مورد بازیابی قرار گرفتند و شش نوع آشفتگی نگارش را به‌زحمتی در همانهایی بافته شد. اما بیشترین نوع این مربوط به اختصار و کوتاه‌نویسی (۲۱/۵۹٪) بود.

۶. به‌شناسه‌ها

با توجه به ناهماهنگی ضبط اطلاعات در منابع و اثر‌هایی کتاب‌خانه‌ای:
- نیاز به تدوین شیوه‌هایی واحد برای برطرف کردن تشخیص موجود از سوی وزارت‌خانه

surnames, initial name, middle name, family)

- متغییر که در آن مفاهیم ظرفیت و قابلیت ارتباط

به‌طور دقیق تعیین شده است و از یکدیگر افتراق می‌یابند، در اثر‌هایی کتاب‌خانه‌ای و

می‌توان نتیجه‌گیری احتمال می‌شود.

- فهرست مستند اسمی برای یکداستن نام و نام خانوادگی پیدا آورندگان، به منظور

حصول اطلاعاتی از یکدیگر کاربرد آنها نیست.

- همچنین بهتر است شیوه‌های ملی برای آوانویس اسمی ایرانی به انگلیسی تهیه شود

نا در موارد مختلف برای رعایت یکداستی از آن استفاده شود. به عنوان مثال، برگرداندن

نام افراد به انگلیسی برای درج در پاهورت و نیز درج اسمی مؤلفان ایرانی به انگلیسی

در آثار آنان.

۷. منابع

صدیقی‌نژادی، آنالیز. ۱۳۸۱. شناسه‌های ضبط اطلاعات ناهنجاری. نشر آثار.

مهم‌نژاد، علی. ۱۳۸۵. تئوری و عمل اطلاعات. نشر انتشارات نوش‌نامه‌های عارف. نگارش‌های ایرانیاسلامی. جلد: میزان آن‌های ناهنجاری. (دسترسی در http://www.hawzah.net/Hawzah/Articles/Articles.aspx).


Masada, T., and J. Adachi. 2007. Citation data clustering for author name disambiguation.


Using web information for author name disambiguation.

Study of Spelling Negligence of Iranian Authors' Names in "ISI" Database

Maryam Khosravi*
Faculty Member in IRANDOC

Abstract: The recording of Persian names suffers from negligence and no rule is being followed in this regard. This paper was a descriptive-quantitative study, it also examined negligence recording of names of Iranian authors in the ISI database. In line with the study goals, the complete list of papers by Iranians during 2007 and 2008 in the Web of Science Database were extracted. Of other purpose of this research was to determine their gender and their way of spelling (consistently or negligently) were surveyed and each of the authors was dedicated a record. The results showed men's names were more negligently recorded compared with that of women.

Keywords: Names of Iranian authors, spelling negligently, spelling of authors' names, ISI database

*Corresponding author: khosravi@irandoc.ac.ir