**نام مقاله: استفاده از ضريب وزني در سرعنوان هاي موضوعي در راستاي بهبود بازيابي اطلاعات**

**نام نشريه: فصلنامه كتابداري و اطلاع رساني (اين نشريه در www.isc.gov.ir نمايه مي شود)**

**شماره نشريه: 41 \_ شماره اول، جلد 11**

**پديدآور: دكتر مرتضي كوكبي**

**مترجم:**

**چكيده**

**در فهرستنويسي موضوعي، فهرستنويس براي بيان محتواي موضوعي كتاب در دست فهرستنويسي، اغلب از سرعنوان(هاي) موضوعي استفاده مي‌كند. نحوة قرار گرفتن يك سرعنوان يكسان، در فهرستبرگه‌هاي متفاوت، ممكن است متفاوت باشد. يك سرعنوان ممكن است تنها سرعنوان تخصيص يافته به يك فهرستبرگه، يا با يك يا چند سرعنوان ديگر همراه باشد. در حالت دوم، بسته به رتبه سرعنوان در ميان سرعنوانهاي ديگر، آن سرعنوان به طور نظري، داراي وزن متفاوتي خواهد بود. امّا در تمامي اين موارد، اين سرعنوانها به گونه­اي مساوي در نظر گرفته مي‌شوند؛ نكته­اي كه مي­تواند بر بازيابي اطّلاعات تأثيري منفي بگذارد. نوشتة حاضر تلاش مي­كند براي حلّ اين مشكل در فهرستنويسي سنّتي، نرم افزارهاي كتابخانه­اي، و مارك ايران، راه حلهايي پيشنهاد نمايد.**

**كليدواژه­ها: فهرستنويسي موضوعي، ضريب وزني، نرم افزارهاي كتابخانه­اي، مارك ايران.**

**مقدّمه**

**«وزن‌دهي»[2] كه با عبارتهاي ديگري مانند «ارزشگذاري وزني اصطلاح» نيز در متون مربوط به بازيابي اطّلاعات و از جمله متون كتابداري ديده مي­شود، عبارت است از «ارزشگذاري اصطلاحات به كار رفته در عبارت جستجو، كه طي آن بالاترين ارزش و وزن به اصطلاحاتي داده مي­شود كه مربوط‌ترين يا مفيدترين اصطلاح براي درخواست جستجو باشند» (كينن[3]، 1378، 84). «پوليت»[4] نيز در كتاب «نظامهاي ذخيره و بازيابي اطلاعات: خاستگاه، توسعه و كاربردها» در مبحث «وزن­دهي اصطلاح»، چنين مي­نويسد: «اهميت مفروض يك اصطلاح را مي­توان با فراواني رويداد آن در پايگاه اطلاعاتي به هم پيوند زد: هر چه فراواني اصطلاح بيشتر باشد، بين مدارك مربوط و ساير مدارك كمتر تفاوت مي­گذارد». (پوليت، 1380، 111). در اين كتاب، حدود چهار صفحه به مفهوم وزن­دهي اصطلاح اختصاص داده شده است. «هارتلي»[5] و ديگران نيز در كتاب «اصول و روشهاي جستجوي پيوسته»، در هنگام بحث پيرامون «بازيابي رتبه­اي» به وزن دهي واژه­ها اشاره مي­كنند و هدف اين وزنها را «پديد آوردن امكان تطبيق خودكار پرسشها با ركوردهاي مدارك از طريق محاسبة ارزش تطبيق رياضي مي­دانند» (هارتلي و ديگران، 1380، 286). «حسن‌زاده» در مقاله­اي با عنوان «تأثير مدلهاي بازيابي اطلاعات بر ميزان ربط»، دربار? ضريب وزني نيز مطالبي را نوشته است (حسن زاده، 1383). «بيزا ـ ييتس و ريبرو ـ نتو»[6] در صفحات 184 تا 192 و 215 تا 217 از كتاب خود با عنوان «قلمروهاي نو در بازيابي اطلاعات»، در مورد وزن دهي و وزن دهي احتمالي سخن مي‌­گويند (بيزا ـ ييتس و ريبرو ـ نتو، 1384). همچنين، «گراسمن و فريدر»[7] در فصل دوم كتاب خود با عنوان «بازيابي اطلاعات: الگوريتمها و روشهاي اكتشافي» كه به راهبردهاي بازيابي اختصاص دارد، در مبحث «مدل فضاي برداري»، محاسبه ضريب وزني يك اصطلاح در يك مدرك را «تركيبي از فراواني اصطلاح (df) و فراواني مدرك معكوس (idf)» تعريف مي­كنند (گراسمن و فريدر، 1384، 21).**

**امّا نكته­اي كه بايد مورد اشاره و تأكيد قرار گيرد، اينكه وزن­دهي در نظامهاي بازيابي كه از روش نمايه‌سازي پس‌همارا[8] استفاده مي­كنند، به كار گرفته مي‌شود. نظامهاي بازيابي مبتني بر نمايه سازي پيش ـ همارا[9] نمي­توانند از وزن­دهي استفاده كنند. تا آن جا كه نگارنده مي­داند، تاكنون دربار? استفاده از ضريب وزني يا وزن دهي در نظامهاي بازيابي مبتني بر نمايه‌سازي پيش­همارا، مانند برگه دان سنّتي كتابخانه و نظامهاي بازيابي كه بر ساختار برگه‌دان متّكي هستند، حدّاقل در زبان فارسي مطلبي منتشر نشده است. آنچه در ادامه مي‌­آيد، استفاده از ضريت وزني را در اين گونه نظامهاي بازيابي اطّلاعات مورد بررسي قرار مي­دهد.**

**استفاده از ضريب وزني در نظامهاي بازيابي مبتني بر نمايه‌سازي پيش‌ـ‌همارا**

**در فهرستنويسي موضوعي، فهرستنويس براي بيان محتواي موضوعي كتاب در دست فهرستنويسي، اغلب از سرعنوان(هاي) موضوعي استفاده مي‌كند. اين سرعنوان (ها) در مراحل بعدي، در بالاي نسخه­اي از همان فهرستبرگه تايپ مي‌شوند و برگة مادر را به برگة موضوع تبديل مي كنند. در برابر هر سرعنوان موضوعي كه به كتاب داده مي‌شود يك برگة موضوع توليد مي شود و در برگه دان موضوع در رديف الفبايي خود قرار مي‌گيرد. در تمامي مراحل فهرستنويسي موضوعي، فهرستنويس تلاش مي‌كند اصول متداول فهرستنويسي موضوعي، و از جمله «ويژگي» را رعايت كند. «ويژگي، يعني به جاي موضوع اعم و كلي تر بايد سعي شود سرعنوان موضوعي اخص و مستقيم‌تر به كار برده شود؛ يعني، موضوع مورد نظر در قالب واژه يا اصطلاح خاص مورد استفاده كاربران ارائه شود. مفهوم مستقيم يعني سرعنوانها به طور مستقيم و بدون واسطة واژه‌هاي هم‌خانواده در فهرست بيايند تا كاربر سريع تر و آسان تر به موضوع مورد نظر خود دست يابد». (فتّاحي و طاهري، 1384، 64).**

**نحو? قرار گرفتن يك سرعنوان يكسان در فهرستبرگه‌هاي متفاوت، ممكن است متفاوت باشد. براي نمونه، سرعنوان «آمار» ممكن است در فهرستبرگه­اي به منزلة تنها موضوع، به كتابي داده شود كه دربار? آمار به طور كلّي است. همين سرعنوان ممكن است در فهرستبرگه­اي ديگر، در كنار يك يا چند سرعنوان ديگر و در جايگاهي متفاوت قرار گيرد، مثلاً در كنار سرعنواني ديگر مانند «احتمالات». در اين حالت دوّم، بسته به تعداد سرعنوانهاي داده شده به كتاب و نيز، جايگاه سرعنوان «آمار» در ميان اين سرعنوانها، تقسيمهاي متفاوتي پيش مي آيد. معمولاً فرض بر اين است كه اگر «آمار» سرعنوان اوّل است موضوع اصلي كتاب در دست فهرستنويسي، «آمار» است يا حدّاقل بخش عمد? آن به «آمار» اختصاص دارد. امّا وزن موضوعي اين سرعنوان با وزن سرعنوان حالت اوّل متفاوت است زيرا اگر فرض را بر اين بگذاريم كه وزن سرعنوان «آمار» در حالت اول، عدد يك است در حالت دوم، بسته به تعداد سرعنوانهاي داده شده، اگر سرعنوان «آمار»، سرعنوان اول باشد، وزن آن معادل كسري از يك، و اگر سرعنوان دوم يا سوم باشد، وزني كمتر خواهد داشت. امّا در تمامي اين حالتها، سرعنوان «آمار»، در بالاي برگة موضوعي خاص خود قرار مي‌گيرد و هنگامي كه فهرستبرگه هاي مربوط به حالتهاي بالا در برگه‌دان موضوعي برگه‌آرايي مي‌شوند، در عمل داراي وزنهايي يكسان به شمار مي‌آيند؛ در حالي كه عملاً اين گونه نيست زيرا در حالت اوّل، يعني زماني كه تنها يك سرعنوان «آمار» به كتاب داده شده است، كلّ كتاب به آمار اختصاص دارد؛ در حالي كه در حالتهاي دوم، تنها بخشي از كتاب به آمار مربوط است. در اين جا، شايد اين نكته مطرح شود كه كاربر مي تواند با بررسي ساير سرعنوانهاي موضوعي داده شده در بخش فهرست تحليلي هر فهرستبرگه، رتبه هر سرعنوان و تعداد سرعنوانهاي ديگر در آن فهرستبرگه را تشخيص دهد؛ امّا در پاسخ مي توان گفت كه حتّي اگر فهرستبرگه هاي معمولاً در هم فشرد? موجود در برگه‌دانهاي كتابخانه‌ها چنين فرصتي را براي كاربر ايجاد كند، بعيد است او چنين بررسي‌اي را انجام دهد، ضمن آنكه در كنار هم قرار گرفتن سرعنوانهاي موضوعي و ساير شناسه افزوده‌ها، تشخيص تعداد سرعنوانها و رتبه هر سرعنوان را براي كاربر معمولاً ناآشنا به رموز فهرستنويسي، دشوار مي سازد.**

**حالت سومي نيز معمولاً رخ مي دهد و آن اين است كه سرعنوان «آمار»، به تنهايي موضوع قرار نمي گيرد و تقسيم فرعي بر سرعنواني اصلي است؛ مانند «آموزش و پرورش ـ آمار». در اين حالت نيز، اگر تنها يك سرعنوان به كتاب داده شود، تقسيم فرعي «آمار» وزني كمتر از سرعنوان اصلي يعني «آموزش و پرورش» دارد، امّا اگر سرعنواني مانند «آموزش و پرورش ـ ايران ـ آمار» به كتابي داده شود، وزن تقسيم فرعي «آمار» در اين حالت، احتمالاً از حالت اوّل نيز كمتر خواهد بود. اين حالت در بازيابي خودكار اطّلاعات در نرم افزارهاي كتابخانه­اي، و در حالت جستجوي «موضوع كليدواژه­اي» پيش مي آيد كه در آن، رايانه در هنگام جستجوي مدارك مربوط به «آمار»، ممكن است حتي تقسيمهاي فرعي «آمار» را نيز بازيابي كند. در چنين حالتي نيز، مدارك مربوط به «آمار» با وزنهايي متفاوت ـ و نه لزوماً به ترتيب موضوعي وزني ـ در كنار هم بازيابي خواهند شد.**

**شايان ذكر است، در تمامي اين حالتها، سرعنوانهايي كه با رعايت اصل «ويژگي» ارائه شده اند در عمل، بازيابي را به گونه­اي «ويژه» انجام نمي‌دهند؛ زيرا مثلاً سرعنوان «آمار»ي كه سومين سرعنوان از سه سرعنوان يك كتاب است، نمي‌تواند هموزن سرعنوان «آمار»ي باشد كه تنها سرعنوان يك كتاب است. امّا در عمل، به شرحي كه در بالا آمد چنين پنداشته مي شود و سرعنوان «آمار»ي كه سومين سرعنوان از سه سرعنوان يك كتاب است، به همان انداز? تك سرعنوان «آمار»، «ويژه» در نظر گرفته مي‌شود.**

**آنچه در اين نوشته خواهد آمد، تلاش براي ارائه راهكارهايي براي حلّ اين مشكل و بهبود بازيابي موضوعي اطّلاعات در برگه‌دانهاي كتابخانه، نرم افزارهاي كتابخانه­اي و مارك ايران است. براي ارائه راه حل براي سه حالت برگه‌دانهاي كتابخانه، نرم افزارهاي كتابخانه­اي و مارك ايران، بايد به اين نكته توجّه داشت كه وزن هر سرعنوان، تابعي از سه متغير زير است:**

**1.تعداد سرعنوانهاي داده شده به يك كتاب**

**2.مرتبه آن سرعنوان در بين سرعنوانهاي داده شده**

**3.حجم يا تعداد صفحات كتاب.**

**متغير حجم يا تعداد صفحات كتاب، از اين نظر مي تواند مهم باشد كه وزن يك سرعنوان «آمار»، در كتابي با تعداد صفحاتي با حدود 300 صفحه، كه يكي از سرعنوانهاي آن «آمار» است و به طور نظري، حدود 100 صفحه از آن به آمار اختصاص مي‌يابد، بيشتر از وزن سرعنوان «آمار» در كتابي مثلاً 80 صفحه­اي است كه تنها سرعنوان موضوعي آن «آمار» است. البتّه، از آنجا كه معمولاً حجم كتابها در حدود معيني قرار دارد، متغير حجم يا تعداد صفحات كتاب، چندان نمي تواند در اين مورد تأثير گذار باشد.**

**نكتة ديگري كه بايد مورد توجّه قرار گيرد، اين است كه در تمامي موارد مورد بحث، فرض بر اين است كه فهرستنويس در مواردي كه چند سرعنوان به كتابي مي‌دهد، سرعنوان اوّل را به منزلة مرتبط‌ترين يا سرعنوان عمده، و سرعنوان (هاي) بعدي را به منزلة سرعنوانهايي كه اهميت آنها از سرعنوان اوّل به سرعنوانهاي بعدي كاهش مي‌يابد، در نظر مي‌گيرد. در اين مورد بايد اذعان داشت، نمايه ساز انساني همواره در معرض اشتباهات انساني شامل پيشينة ذهني و يا ساير موارد قرار دارد، امّا در اين جا مي توان استدلال كرد چه از ضريب وزني استفاده بشود چه نشود، اين اشتباهات در فهرستنويسي موضوعي اجتناب‌ناپذير است.**

**راه‌حلّ پيشنهادي براي فهرستبرگه‌ها**

**راه حلّ پيشنهادي براي فهرستبرگه‌ها عبارت است از افزودن دو عدد در پايان هر سرعنوان موضوعي، به گونه­اي كه عدد اوّل كه در كمانك گذاشته مي‌شود. رتبه سرعنوان و عدد دوم كه بدون كمانك است، تعداد سرعنوانهاي داده شده به كتاب را نشان دهد. مي‌توان به طور قراردادي، براي تك سرعنوان از هيچ عددي استفاده نكرد. بدين ترتيب:**

**«آموزش و پرورش»، يك سرعنوان است.**

**«آموزش و پرورش (1)2»، اوّلين سرعنوان از دو سرعنوان است.**

**«آموزش و پرورش (1)3»، اولين سرعنوان از سه سرعنوان است.**

**«آموزش و پرورش (2)3»، دومين سرعنوان از سه سرعنوان است.**

**«آموزش و پروش (2)5»، دومين سرعنوان از پنج سرعنوان است.**

**اگر اين اعداد در هنگام برگه آرايي در نظر گرفته شود، كاربر در هنگام جستجوي مطالبي دربار? آموزش و پرورش، نخست سرعنوانهايي را مي‌يابد كه تكي و در نتيجه، داراي بيشترين وزن موضوعي هستند. در مراحل بعدي، سرعنوانهايي را پيدا مي كند كه اولين هستند، اما تنها سرعنوان نيستند. اينها اگر چه از نظر موضوعي وزن سرعنوانهاي گروه اوّل را ندارند، امّا نسبت به سرعنواني مانند «آموزش و پرورش(2)5» كه دومين سرعنوان از پنج سرعنوان است، وزن بيشتري دارند. با انجام چنين كاري، كاربر مطالب مربوط را از مرتبط‌ترين تا كم ارتباط‌ترين پيدا مي كند. در ادامه، چند نمونه از يك نسخه از كتابشناسي ملّي ايران (كتابشناسي ملي ايران، 1375، 1، 297) ارائه خواهد شد. اگر چه همة نمونه ها از فهرستبرگه هاي كتابشناسي ملّي ايران گرفته شده، امّا به دليل صرفه‌جويي در فضاي نوشته، تنها عناوين و سرعنوانهاي موضوعي هر فهرستبرگه ارائه خواهد شد:**

**آمار و كاربرد آن در مديريت**

**1. آمار 2. مديريت ـ روشهاي آماري.**

**آمار كاربردي در اقتصاد و بازرگاني**

**1. آمار 2. آمار بازرگاني 3. اقتصاد ـ روشهاي آماري.**

**آمار و احتمال مقدماتي**

**1. احتمالات 2. آمار.**

**آمار كاربردي**

**1. علوم اجتماعي ـ روشهاي آماري 2. مديريت ـ روشهاي آماري 3. اقتصاد ـ روشهاي آماري 4. آمار.**

**با بررسي چهار سرعنوان «آمار» در بالا، مشخص مي‌شود وزنهاي موضوعي يكساني ندارند، زيرا رتبه هر سرعنوان و تعداد سرعنوانهاي ديگر موجود در هر فهرستبرگه، متفاوت است. امّا در فهرستنويسي سنّتي، تمام اين سرعنوانها، هنگامي كه در بالاترين سطر فهرستبرگه تايپ و فهرستبرگه‌هاي مربوط در جاي الفبايي خود برگه‌آرايي مي‌شوند، عملاً وزني يكسان خواهند داشت. در حالي كه اگر به روش بالا براي هر يك ضريب وزني درخور تعيين شود و اين ضريبها در برگه‌آرايي نيز منظور شوند، بازيابي از مرتبط‌ترين تا كم ارتباط‌ترين انجام و بدين ترتيب اصل ويژگي نيز رعايت مي شود. بدين ترتيب، سرعنوانهاي «آمار» به ترتيب زير ضريب دار مي شوند:**

**كتاب اول: آمار(1)2**

**كتاب دوم: آمار(1)3**

**كتاب سوم: آمار(2)2**

**كتاب چهارم: آمار(4)4**

**ترتيب برگه‌آرايي نيز طبيعتاً همان ترتيب بالاست؛ يعني براي سرعنوانهاي همانند، نخست شماره هاي مساوي درون كمانك در كنار هم و به ترتيب شماره هاي بعدي از كم به زياد مرتّب مي شوند و آن گاه، شماره هاي درون كمانك كه مقدار بيشتري دارند، مي‌آيند. بديهي است، در نمونة بالا، تنها سرعنوانهاي «آمار» با ضريب وزني نشان داده شده‌اند و سرعنوانهاي ديگر نيز بايد براي بازيابي‌هاي مربوط، ضريبهاي وزني مناسب داشته باشند.**

**راه‌حلّ پيشنهادي براي نرم افزارهاي كتابخانه­اي**

**در نرم افزارهاي كتابخانه­اي معمولاً دو نوع جستجوي موضوعي امكان پذير است: «موضوع سرعنواني» و «موضوع كليدواژه­اي». اين عبارتها در نرم‌افزارهاي متفاوت ممكن است به شكلهاي متفاوتي به كار بروند. در جستجو از طريق موضوع سرعنواني، فرض بر اين است كه كاربر عين سرعنوان موضوعي را در محلّ جستجو و گزينة «موضوع سرعنواني» را در «فيلد جستجو» وارد مي‌كند. در اين حالت، شماره گذاري سرعنوانها به روش بالا مي‌تواند راه حلّ مناسبي براي تعيين وزن سرعنوانها باشد. در اين حالت، تنها وجود الگوريتمي ساده، رايانه را قادر مي‌سازد تا سرعنوانهاي سنگين وزن تر را در آغاز و سبك وزن تر را در مراحل بعدي بازيابي كند. بديهي است، به دليل ناآشنايي اغلب كاربران با ساختار سرعنوانهاي موضوعي، آنان ترجيح مي دهند بيشتر از جستجوي «موضوع كليدواژه­اي» استفاده كنند.**

**در فرايند جستجوي «موضوع كليدواژه­اي»، فرض بر اين است كه رايانه، كليدواژه‌هاي مندرج در سرعنوانهاي موضوعي را بازيابي مي كند. در اين حالت، اگر كاربر كليدواژ? «آمار» را در محل جستجو، و گزينة «موضوع كليدواژه­اي» را در «فيلد جستجو» وارد نمايد، رايانه بايد كليد واژ? «آمار» را، افزون بر سرعنوانهايي كه داراي واژ? «آمار» به منزلة سرعنوان اصلي هستند، در سرعنواني مانند «آموزش و پرورش ـ ايران ـ آمار» نيز پيدا كند. در چنين حالتي، اين كه واژ? «آمار» در كجاي سرعنوان قرار دارد نيز اهميت پيدا مي‌كند. نكته قابل بررسي اين است كه وقتي از مفهوم اهميت موضوعي در بخشهاي مختلف يك سرعنوان سخن به ميان مي‌آيد، نمي‌توان با قاطعيت اظهار نظر كرد كه در سرعنواني مانند «آموزش و پرورش ـ ايران ـ آمار» كدام يك از دو مفهوم ايران و آمار اهميت بيشتري دارند، زيرا چنين تركيبي به نظر مي‌رسد بيشتر تابع دستورات كتاب سرعنوانهاي موضوعي باشد تا منطقي ديگر. در نتيجه، مهم اين است كه آيا واژگان «ايران» و «آمار» به منزلة بخش اوّل يا مهم سرعنوان آمده اند يا نه. بدين ترتيب، آنچه در اين حالت مهم است، اتّخاذ تمهيدي است تا به وسيلة آن رايانه بتواند بخش اوّل را از بخشهاي بعدي در يك سرعنوان تميز دهد. بديهي است، اين نكته نيز همچنان باقي است كه اين بخش اوّل در سرعنواني با چه رتبه­اي و در ميان چند سرعنوان قرار دارد.**

**نكتة ديگري كه در جستجوي «موضوع كليدواژه­اي» بايد مورد بررسي قرار گيرد، دو حالتي است كه جستجو در آنها انجام مي شود: جستجو بدون استفاده از عملگرهاي بولي و جستجو با استفاده از اين عملگرها. اين دو حالت در ادامه بررسي مي‌شوند.**

**در حالت اول، هنگامي كه كاربر جستجوي «موضوع كليدواژه­اي» را با يك تك واژه يا عبارت انجام مي دهد، اگر تك واژه در سرعنواني با يك يا چند تقسيم فرعي وجود داشته باشد، رايانه تنها كافي است پس از شناسايي اينكه سرعنواني كه كليد واژ? مورد جستجو در آن قرار دارد، داراي چه رتبه­اي و در ميان چند سرعنوان ديگر است تشخيص دهد كه آيا كليد واژ? مورد نظر كاربر در بخش اوّل آمده است يا نه. مثال زير اين نكته را روشن‌تر مي‌كند. فرض كنيم كاربري در هنگام جستجوي «موضوع كليدواژه­اي»، واژ? «آمار» را در محل جستجو وارد كرده است. آن چه روشن است اينكه كاربر، قصد دارد كتابي دربار? آمار بيابد امّا اين احتمال وجود دارد كه رايانه، سرعنواني مانند «آموزش و پرورش ـ ايران ـ آمار» را، به دليل وجود واژ? «آمار» در آن بازيابي كند كه منظور نظر كاربر نيست. امّا اگر سرعنوان بالا به روش زير شماره‌گذاري شود: «آموزش و پرورش (1) ـ ايران ـ آمار» به گونه‌اي كه شمار? يك نشان دهد بخش اوّل، «آموزش و پرورش» است و الگوريتم بازيابي به گونه­اي نوشته شده باشد كه در اين حالت، تنها كليدواژه هايي را پيدا كند كه با شمار? يك مشخّص شده‌اند، اين بازيابي نادرست يا ريزش كاذب رخ نخواهد داد.**

**در حالت دوم، الگوريتم بازيابي بايد به گونه­اي نوشته شود كه رايانه بتواند سه كليد واژ? مورد نظر كاربر را در سه حالت مختلف بررسي كند. دليل اين كه از سه كليدواژه نام برده شد اين است كه به نظر مي رسد نرم افزارهاي كتابخانه­اي حداكثر تا سه محل براي سه كليدواژه فراهم مي‌كنند. اكنون فرض كنيم كاربري كتابي در مورد «آمار آموزش و پرورش ايران» مي خواهد. او احتمالاً بنا به عادت زبان طبيعي خود، در محل اول، كليدواژ? «آمار»، در محل دوم، كليدواژ? «آموزش و پرورش»، و در محل سوم، كليدواژ? «ايران» را تايپ خواهد كرد. در صورتي كه سرعنوان «آموزش و پرورش(1) ـ ايران ـ آمار» در پيشينه‌هاي مربوط وجود داشته باشد، رايانه هر سه كليدواژه را از اين نظر كه كدام يك در بخش اوّل سرعنوانهاي با تقسيمات فرعي قرار مي گيرد آزمايش خواهد كرد و پس از گزينش سرعنوانهايي كه بخش اوّل آنها «آموزش و پرورش» است، در تقسيمات فرعي اين سرعنوانها به جستجوي دو كليدواژ? ديگر خواهد رفت. بدين ترتيب، اگر در هر سرعنوان، افزون بر مشخّص شدن بخش اوّل، معين شده باشد كه سرعنوان، داراي رتبه چندم و در ميان چند سرعنوان است، بازيابي مدارك مرتبط به صورتي بهينه انجام خواهد گرفت. شايان ذكر است، در اين جا منظور از بخش اوّل هر سرعنوان، تنها بخش اوّل در سرعنوان با تقسيمهاي فرعي نيست بلكه در سرعنوانهاي با نقطه نيز مي باشد.**

**نمونه‌هاي زير از يك نسخه كتابشناسي ملّي ايران (كتابشناسي ملي ايران، 1375، 1، 456ـ457) احتمالاً به درك بهتر موضوع كمك مي‌كند:**

**رويش و پرواز: چهار پاره**

**1. شعر فارسي ـ قرن 14 ـ مجموعه‌ها.**

**... كشتي طوفان كربلا ...**

**1. شعر مذهبي ـ قرن 14 ـ مجموعه‌ها. 2. شعر فارسي ـ قرن 14 ـ مجموعه‌ها.**

**مدايح و مراثي آل محمد (ص): مجموعه زيبا و جانسوز شعراي بنام و معاصر ايران ... به ضميمه­اي از سروده‌هاي اسلامي**

**1. چهارده معصوم ـ شعر. 2. شعر مذهبي ـ قرن 14 ـ مجموعه‌ها. 3. شعر فارسي ـ قرن 14 ـ مجموعه‌ها.**

**عاشقانه‌ها: گزينه سرودهاي شاعران امروز ايران**

**1. شعر عاشقانه فارسي ـ قرن 14. 2. شعر آزاد ـ مجموعه‌ها. 3. شاعران ايراني ـ قرن 14. 4. شعر فارسي ـ قرن 14 ـ مجموعه‌ها.**

**در چهار كتاب بالا كه تنها عناوين و سرعنوانهاي موضوعي فهرستبرگه‌هاي آنها در اين جا ارائه شده است، سرعنوان «شعر فارسي ـ قرن 14 ـ مجموعه‌ها». در چهار رتبه متفاوت در چهار فهرستبرگه ديده مي شود. بديهي است، چنانچه اين سرعنوانها به روش بالا داراي ضريب وزني شوند، در بازيابي، سرعنوان كتاب چهارم با ضريب وزني (4)4 در آخر و سرعنوان كتاب اوّل بدون ضريب وزني (به معناي يك از يك) در آغاز قرار مي‌گيرد. امّا در اين جا لازم مي‌آيد اين سرعنوانها ضريب ديگري نيز داشته باشند، زيرا قرار است مورد جستجوي «موضوع كليدواژه­اي» نيز قرار گيرند. بدين ترتيب و با توجّه به استدلال بالا كه تنها بخش اوّل هر سرعنوان در اين مرحله اهميت دارد، بخشهاي اوّل هر سرعنوان (در اين جا، «شعر فارسي») بايد ضريبي داشته باشند تا بر مبناي آن به منزلة كليدواژ? اصلي استخراج گردند. از آنجا كه اين ضريب تنها در سرعنوانهاي با تقسيم فرعي كارآيي دارد، افزودن شمار? يك در كمانك به سرعنوان اصلي، احتمالاً براي اين منظور كفايت مي‌كند. بدين ترتيب، سرعنوانهاي موضوعي چهار كتاب بالا به صورت زير در مي‌آيد:**

**1. شعر فارسي(1) ـ قرن 14 ـ مجموعه‌ها.**

**1. شعر مذهبي(1) ـ قرن 14 ـ مجموعه‌ها(1)2. 2. شعر فارسي(1) ـ قرن 14 ـ مجموعه‌ها(2)2.**

**1. چهارده معصوم(1) ـ شعر(1)3. 2. شعر مذهبي(1) ـ قرن 14ـ مجموعه‌ها(2)3. 3. شعر فارسي(1) ـ قرن 14 ـ مجموعه‌ها(3)3.**

**1. شعر عاشقانه فارسي(1) ـ قرن 14(1)4. 2. شعر آزاد(1) ـ مجموعه‌ها(2)4. 3. شاعران ايراني(1) ـ قرن 14(3)4. 4. شعر فارسي(1) ـ قرن 14 ـ مجموعه‌ها(4)4.**

**در اين شرايط، رايانه در پاسخ به كاربري كه خواهان مجموعه‌هاي شعر فارسي در قرن 14 است مدارك داراي سرعنوانهاي موضوعي زير را به ترتيب زير بازيابي مي‌كند:**

**1. شعر فارسي(1) ـ قرن 14 ـ مجموعه‌ها.**

**2. شعر فارسي(1) ـ قرن 14 ـ مجموعه‌ها(2)2.**

**3. شعر فارسي(1) ـ قرن 14 ـ مجموعه‌ها(3)3.**

**4. شعر فارسي(1) ـ قرن 14 ـ مجموعه‌ها(4)4.**

**در اين حالت، مدارك مربوط با توجّه به كليد واژ? اصلي كه با شمار? (1) مشخّص شده و بر حسب رتبه سرعنوانها بازيابي مي شوند.**

**راه حلّ پيشنهادي براي مارك ايران**

**مارك ايران در بازيابي اطّلاعات همانند ساير نرم افزارهاي كتابخانه­اي عمل مي‌كند؛ يعني هم مي تواند فهرستبرگه توليد كند و هم به منزلة ابزاري براي بازيابي اطّلاعات مورد استفاده قرار گيرد، با اين تفاوت كه بر يك استاندارد جهاني مبتني است.**

**در مارك ايران در بلوك ـ ـ 6 كه بلوك تجزيه و تحليل موضوعي نام دارد، هشت فيلد به شرح زير براي سرعنوانهاي موضوعي وجود دارد (مارك ايران، 1381، 205):**

**600 نام شخص به منزله موضوع**

**601 نام تنالگان به منزله موضوع**

**602 نام خاندان به منزله موضوع**

**604 پديدآور و عنوان به منزلة موضوع**

**605 عنوان به منزله موضوع**

**606 موضوع (اسم عام يا عبارت اسمي عام)**

**607 نام جغرافيايي به منزله موضوع**

**610 اصطلاحهاي موضوعي كنترل نشده**

**همگي اين فيلدها تكرار پذيرند و درون آنها نخستين فيلد فرعي يعني $a به عنصر شناسه­اي اختصاص دارد كه مثلاً در فيلد 606 «حاوي اصطلاح، به همان شكلي است كه سرعنوان موضوعي مورد استفاده تعيين كرده است». (مارك ايران، 1381، 230). فيلدهاي فرعي $j (تقسيم فرعي شكلي)، $x (تقسيم فرعي موضوعي)، $y (تقسيم فرعي جغرافيايي)، و $z (تقسيم فرعي دوره­اي) افزون بر ساير فيلدهاي فرعي، كمابيش در تمامي اين فيلدها وجود دارند.**

**اكنون براي نشان دادن راه حلّ پيشنهادي در مارك ايران، نمونه هاي بالا را در فيلدهاي مارك ايران مي گذاريم:**

**1. آمار. 2. مديريت ـ روشهاي آماري.**

**6061#$aآمار$2nli@**

**6062#$aمديريت$xروشهاي آماري@**

**1. آمار. 2. آمار بازرگاني. 3. اقتصاد ـ روشهاي آماري.**

**6061#$aآمار$2nli@**

**6062#$aآمار بازرگاني$2nli@**

**6062#$aاقتصاد$xروشهاي آماري$2nli@**

**1. احتمالات. 2. آمار.**

**6061#$aاحتمالات$2nli@**

**6062#$aآمار$2nli@**

**1. علوم اجتماعي ـ روشهاي آماري. 2. مديريت ـ روشهاي آماري. 3. اقتصاد ـ روشهاي آماري. 4. آمار.**

**6061#$aعلوم اجتماعي$xروشهاي آماري$2nli@**

**6062#$aمديريت$xروشهاي آماري$2nli@**

**6062#$aاقتصاد$xروشهاي آماري$2nli@**

**6062#$aآمار$2nli@**

**چنانكه مشاهده مي‌شود، در تمام موارد بالا فيلد 606 بر حسب لزوم تكرار شده است و در همة فيلدها، فيلد فرعي $a به منزلة «عنصر شناسه­اي» وجود دارد. اگر چه در صفحة 21 از دستنامة مارك ايران در مورد «ترتيب فيلدها» نوشته شده كه «نظم خاصي براي ضبط فيلدها در يك ركورد پيش بيني نشده است. نظم حاكم بر فهرست راهنماي فيلدها ممكن است با ترتيب ضبط فيزيكي فيلدها يكي نباشد». امّا در مارك ايران، و مثالهاي ارائه شده در انتهاي «دستنامة مارك ايران» نيز نشان مي‌دهد مي‌توان نظم فيلدهاي موضوعي را بر حسب نظم ارائه آنها در يك فهرستبرگه انتخاب نمود. اين نكته در مارك ايران در مورد نظم فيلدهاي فرعي نيز صدق مي‌كند؛ بدين معنا كه جز در مواردي خاص كه فيلدهاي فرعي از نظمي خاص پيروي مي‌كنند، ترتيب خاصّي براي فيلدهاي فرعي در نظر گرفته نشده است. در مثالهاي ارائه شده در انتهاي كتاب «دستنامة مارك ايران» نيز نظم فيلدهاي فرعي موضوعي، همان ترتيب سرعنوان اصلي و تقسيمهاي فرعي در سرعنوانهاي موضوعي است. بدين ترتيب، هم در پيشينة مارك ايران رتبه هر سرعنوان مشخّص مي‌شود و هم مي توان با افزودن يك الگوريتم ساده، اين امكان را فراهم ساخت كه رايانه هنگام توليد فهرستبرگه، با احتساب فيلدهاي موضوعي و رتبه هر سرعنوان، ضرايب وزني لازم را در انتهاي هر سرعنوان موضوعي اضافه كند. لازم به توضيح است، اگر چه در تمام نمونه‌هاي بالا از فيلد606 استفاده شده است، امّا استفاده از فيلدهاي موضوعي ديگر نيز تفاوتي به وجود نمي آورد و راه حلّ پيشنهادي به قوّت خود باقي است.**

**در نمونه هاي بالا، سرعنوانهاي با تقسيم فرعي نيز وجود دارند و به اعتباري، مسألة تعيين بخش اوّل يا مهم هر سرعنوان با استفاده از فيلد فرعي «عنصر شناسه­اي» تا حدودي روشن شده است. با اين حال، يكي از نمونه‌هاي مربوط به جستجوي «موضوع كليدواژه­اي» براي وضوح بيشتر، با استفاده از نشانه گذاري مارك ايران، در زير ارائه مي‌شود:**

**1. شعر عاشقانه فارسي ـ قرن 14. 2. شعر آزاد ـ مجموعه‌ها. 3. شاعران ايراني ـ قرن 14. 4. شعر فارسي ـ قرن 14 ـ مجموعه‌ها.**

**6061#$aشعر عاشقانه فارسي$zقرن 14$2nli@**

**6062#$aشعر آزاد$jمجموعه‌ها$2nli@**

**6062#$aشاعران ايراني$zقرن 14$2nli@**

**6062#$aشعر فارسي$zقرن 14$jمجموعه‌ها$2nli@**

**در نمونة بالا، باز هم مشاهده مي شود كه ترتيب فيلدهاي موضوعي را مي‌توان همان ترتيب سرعنوانهاي موضوعي در نظر گرفت و فيلد فرعي $a (عنصر شناسه­اي) نيز مشخّص مي كند كه جستجوي موضوع كليدواژه­اي، چگونه در سرعنوان اصلي مي تواند انجام شود.**

**نتيجه‌گيري و پيشنهادها**

**در صورت پذيرش پيشنهادهاي بالا بهتر است:**

**1.اين ضرايب وزني در شناسه هاي كتابشناسي ملّي ايران ـ چه به صورت چاپّي و چه به صورت لوح فشرده ـ اعمال شود تا از اين پس در كتابخانه هاي گوناگوني كه از اين كتابشناسي در راستاي فهرستنويسي بهره مي‌گيرند، به كار رود.**

**2.در نرم افزارهاي كتابخانه­اي نيز اين ضرايب ـ خواه به صورت رويكردي آينده نگر و خواه به صورت رويكردي گذشته‌نگر ـ انجام شود.**

**3.در نرم افزارهاي آتي فرمت مارك ايران، الگوريتمهاي لازم براي توليد اين ضريبها در فهرستبرگه‌هاي توليدي از طريق جستجوي «موضوع كليدواژه­اي» در مارك ايران در نظر گرفته شود.**

**منابع**

**ـ بيزا ـ ييتس، برتيه ريكاردو و نتو ريبرو (1384). قلمروهاي نو در بازيابي اطلاعات، ترجمة علي حسين قاسمي با همكاري سيروس آزادي، علي جوامع، تهران: چاپار.**

**ـ پوليت، ا. استون (1380). نظامهاي ذخيره و بازيابي اطلاعات: خاستگاه، توسعه و كاربردها/ ترجمه محمدحسين دياني، جعفر مهراد. ـ شيراز: كتابخانه منطقه­اي علوم و تكنولوژي شيراز.**

**ـ حسن‌زاده، محمد (1383). تأثير مدلهاي بازيابي اطلاعات بر ميزان ربط. اطلاع‌شناسي، 2 (1)، 64ـ89.**

**ـ فتاحي، رحمت الله و مهدي طاهري (1375). فهرستنويسي: اصول و روشها ـ ويراست4. ـ تهران: كتابدار.**

**ـ كتابشناسي ملي ايران، فروردين 1349ـ . ـ تهران: كتابخانه ملي ايران، 1349 ـ (شماره 83، نيمه اوّل 1375)**

**ـ كينن، استلا (1378). فرهنگ فشرد? كتابداري و اطلاع‌رساني، ترجمه و تدوين فاطمه اسدي كرگاني، مقابله و ويرايش عبدالحسين آذرنگ. ـ تهران: كتابدار.**

**ـ گراسمن، ديويد ا. و فريدر افير (1384). بازيابي اطلاعات : الگوريتمها و روشهاي اكتشافي؛ ترجمه جعفر مهراد و سارا كليني، مشهد : كتابخانه رايانه­اي ؛ شيراز : كتابخانه منطقه­اي علوم و تكنولوژي.**

**ـ مارك ايران/ كميتة ملي مارك ايران. ـ تهران : كتابخانه ملي جمهوري اسلامي ايران، 1381.**

**ـ هارتلي، آر. ج. ... [و ديگران] (1380). اصول و روشهاي جستجوي پيوسته/ ترجمة زاهد بيگدلي. ـ مشهد : كتابخانه رايانه­اي.**

**--------------------------------------------------------------------------------**

**1. دانشيار گروه كتابداري دانشگاه شهيد چمران اهواز: kokabi80@Yahoo.com**

**2. Term weighting.**

**1. Keenan.**

**2. Pollitt.**

**3. Hartley.**

**4. Baeza-Yates and Ribeiro-Neto.**

**5. Grossman and Frieder.**

**1. Post-coordinate indexing.**

**2. Pre-coordinate indexing.**