



وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

گزارش طرح

شناسایی مسائل و ارائه راهکارهای زبانشناختی

سازمان دهی و بازیابی اطلاعات

در سامانه اطلاعات علمی ایران (گنج)

مجری

زهرا محمودزاده

همکار

زهرا دهسرایي

۱۳۹۷

به نام خدا

حقوق: پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی « شناسایی مسائل و ارائه راهکارهای زبان‌شناختی سازمان‌دهی و بازیابی اطلاعات در سامانه اطلاعات علمی ایران (گنج) » پیرو قرارداد شماره ۳۳/۳۵۷۹ مورخ ۹۶/۳/۲۳ میان پژوهشگاه علوم و فناوری اطلاعات ایران (کارفرما) و خانم زهرا محمودزاده (مجری) اجرا شده است.

این گزارش و تمامی حقوق مادی آن براساس «قانون حمایت حقوق مؤلفان و مصنفان و هنرمندان، مصوب سال ۱۳۴۸ و اصلاحیه‌های بعدی آن و همچنین آیین‌نامه‌های اجرایی این قانون متعلق به پژوهشگاه علوم و فناوری اطلاعات ایران است و هر گونه استفاده از تمامی یا پاره‌ای از آن، شامل: نقل قول، تکثیر، انتشار، کاربرد نتایج، تکمیل و مانند آنها به صورت چاپی، الکترونیکی یا وسایل دیگر فقط با اجازه کتبی پژوهشگاه امکان‌پذیر است. نقل قول در حد هزار واژه در انتشارات علمی مانند کتاب و مقاله با درج اطلاعات کامل کتاب‌شناختی، نیازی به مجوز پژوهشگاه ندارد.

صحت مندرجات گزارش برعهده مجری طرح پژوهشی است.

صفحه تأييد پژوهشگاه

**شناسایی مسائل و ارائه راهکارهای زبانشناختی سازمان‌دهی و بازیابی اطلاعات در
سامانه اطلاعات علمی ایران (گنج)**

مجری طرح: زهرا محمودزاده، هیئت علمی، پژوهشگاه علوم و فناوری اطلاعات ایران

نشانی: تهران، پژوهشگاه علوم و فناوری اطلاعات ایران، خیابان انقلاب، نبش فلسطین

تلفن: ۶۶۴۹۴۹۸۰

وبگاه: irandoc.ac.ir

رایانامه: Mahmoodzadeh@irandoc.ac.ir

همکار(ان) طرح: زهرا دهمسرای، مدیر واحد ساماندهی اطلاعات، پژوهشگاه علوم و فناوری اطلاعات ایران

چکیده

در این پژوهش ابتدا با بررسی میدانی پایگاه اطلاعات علمی ایران (گنج)، مسائل زبانشناختی بازیابی و ساماندهی اطلاعات در این پایگاه بدست آمد و سپس راهکارهایی ارائه شد. مسائل نمایه سازی شامل موارد زیر است: ترجمه به فارسی یا انگلیسی اطلاعات کتابشناختی، استفاده از معادل های متعدد برای یک واژه لاتین، یافتن معادل فارسی کلیدواژه های لاتین، تعیین اختصار فارسی برای یک واژه لاتین، نوشتن صورت های مخفف لاتین به اشکال مختلف، انواع خطاهای املائی و نگارشی در هر دو سطح فارسی و لاتین، مسائل مربوط به درج علائم ریاضی و فرمولهای شیمی در چکیده و عنوان، مطالب نامربوط در چکیده. برای حل مسائل زبانشناختی ساماندهی و نمایه سازی به نظر می رسد علاوه بر بهبود قابلیت نرم افزار نمایه سازی، لازم است دانش تخصصی در زمینه علوم مختلف و نیز زبان فارسی و زبانشناسی بکار گرفته شود.

مسائل بازیابی به دو دسته تقسیم می شوند الف) صورت های متفاوت نوشتاری ناشی از کاربرد علامتهای مختلف نوشتاری: تشدید، همزه، تنوین، آوا، نقطه، خط تیره، صورت های متفاوت کسره اضافه و صورت های مختلف جمع و صورت های مختلف اختصارات و ب) صورت های متفاوت ناشی از مسائل معنایی، صرفی و نحوی: کلمات چندمعنا با املائی یکسان (اما صورت آوایی یکسان یا متفاوت)، صورت های متفاوت واژگانی با معنای یکسان یا مرتبط و صورت های کوتاه / کامل اسامی یا سازمان ها. بررسی ادبیات مربوط به بازیابی اطلاعات نشان می دهد که بکار گیری روشها و تکنیک های پردازش زبان طبیعی تأثیر معناداری در بهبود دقت و بازخوانی سیستم های بازیابی اطلاعات دارد.

کلیدواژه ها: سامانه اطلاعات علمی ایران (گنج)، بازیابی متن، ساماندهی متن، پردازش متن، ساختار زبانی متن

فهرست مطالب

عنوان	شماره صفحه
۱. مقدمه	۲
۲. بیان مسئله	۲
۱-۲. پیشینه پژوهش	۳
۲-۲. مراحل اجرایی پژوهش	۴
۳-۲. پرسش‌های پژوهش	۴
۴-۲. هدف‌های پژوهش	۵
۲. مقدمه	۷
۱-۲. تکمیل اطلاعات کتابشناختی و فایل‌های ضمیمه	۷
۱-۱-۲. مسائل مربوط به تکمیل اطلاعات کتابشناختی و فایل‌های ضمیمه	۱۱
۲-۲. نمایه‌سازی اطلاعات کتابشناختی	۱۳
۱-۲-۲. عناوین و تمام متن‌ها	۱۳
۲-۲-۲. پدیدآوران	۱۴
۳-۲-۲. اطلاعات تخصصی پایان‌نامه	۱۵

۱۵.....	اطلاعات عمومی	۴-۲-۲
۱۶.....	چکیده‌ها و مراجع	۵-۲-۲
۱۹.....	کلیدواژه‌ها	۶-۲-۲
۲۰.....	موضوع‌ها	۷-۲-۲
۲۰.....	مسائل زبانشناختی نمایه‌سازی	۳-۲
۲۱.....	کلیدواژه‌ها	۱-۳-۲
۲۳.....	پدیدآوران	۲-۳-۲
۲۴.....	چکیده و عنوان	۳-۳-۲
۳۶.....	نتیجه‌گیری	۴-۲
۳۶.....	مرحله تکمیل اطلاعات کتابشناختی و فایل‌های ضمیمه	۱-۴-۲
۳۷.....	مرحله نمایه‌سازی	۲-۴-۲
۴۱.....	مقدمه	۳
۴۱.....	بررسی لاگ‌های کاربران	۱-۳
۴۳.....	مسائل زبانشناختی بازیابی	۲-۳
۴۳.....	اشتباه و اشکال در صورت زبانی متن	۳-۲-۱
۴۴.....	صورت‌های متفاوت نوشتاری	۲-۲-۳
۴۸.....	مسائل معنایی-صرفی-نحوی	۳-۲-۳
۴۹.....	مسائل فنی	۴-۲-۳

نتیجه گیری.....	۳-۳	۵۳
مقدمه	۴	۵۶
مسائل سازماندهی و نمایه سازی و راهکارهای پیشنهادی	۱-۴	۵۶
دانش تخصصی	۱-۱-۴	۵۷
جستجو، تصمیم گیری و بهبود قابلیت نرم افزار نمایه سازی	۲-۱-۴	۵۸
ویرایش و اصلاح متن	۳-۱-۴	۵۹
خطایابی و رفع اشتباهات تایپی و نگارشی	۴-۱-۴	۶۰
خطایاب وفا	۵-۱-۴	۶۰
نتیجه گیری.....	۲-۴	۶۳
مسائل بازیابی و راهکارهای پیشنهادی.....	۳-۴	۶۳
مدل های زبانشناختی بازیابی اطلاعات	۱-۳-۴	۶۴

فهرست جدول ها، شکل ها و نمودارها

عنوان	شماره صفحه
نمودار ۱-۱: اطلاعات ورودی و خروجی مرکز اطلاعات و مدارک علمی ایران (اقتباس از احترامی ۱۳۹۵).....	۳
جدول ۲-۱: اطلاعات کتابشناختی مدرک (اقتباس از احترامی ۱۳۹۵).....	۷
نمودار ۱-۲: فرایند نمایه سازی مدرک (اقتباس از احترامی ۱۳۹۵).....	۱۱
شکل ۱-۲: مثالی برای اشتباه در ثبت اطلاعات.....	۱۲
شکل ۲-۲: پنجره اطلاعات کتابشناختی در سامانه نمایه سازی.....	۱۳
شکل ۳-۲: پنجره عنوان در سامانه نمایه سازی.....	۱۴
شکل ۴-۲: پنجره پدیدآوران در سامانه نمایه سازی.....	۱۴
شکل ۵-۲: پنجره اطلاعات تخصصی پایان نامه در سامانه نمایه سازی.....	۱۵
شکل ۶-۲: پنجره اطلاعات عمومی در سامانه نمایه سازی.....	۱۶
شکل ۷-۲: پنجره چکیده تمام نما در سامانه نمایه سازی.....	۱۶
شکل ۸-۲: پنجره چکیده راهنما در سامانه نمایه سازی.....	۱۷
شکل ۹-۲: پنجره چکیده لاتین در سامانه نمایه سازی.....	۱۷
شکل ۱۰-۲: پنجره فهرست مندرجات در سامانه نمایه سازی.....	۱۸
شکل ۱۱-۲: پنجره فهرست منابع فارسی در سامانه نمایه سازی.....	۱۸

- شکل ۲-۱۲: پنجره فهرست منابع لاتین در سامانه نمایه سازی ۱۹
- شکل ۲-۱۳: پنجره کلیدواژه ها در سامانه نمایه سازی ۱۹
- شکل ۲-۱۴: پنجره موضوع ها در سامانه نمایه سازی ۲۰
- جدول ۲-۱: ساخت زبانی اطلاعات کتابشناختی ۲۰
- جدول ۲-۲: مثالی برای معادلهای متعدد واژه های لاتین ۲۱
- جدول ۲-۳: صورتهای نوشتاری متعدد برای یک کلیدواژه ۲۲
- جدول ۲-۴: مثالی برای صورت اختصاری فارسی یک کلیدواژه لاتین ۲۲
- شکل ۲-۱۵: مسائل درج فرمولها و علائم ریاضی و شیمی ۲۵
- شکل ۲-۱۶: مسائل درج فرمولها و علائم ریاضی و شیمی ۲۵
- شکل ۲-۱۷: اشتباهات تایپی و نگارشی در چکیده ۲۷
- شکل ۲-۱۸: اشتباهات تایپی و نگارشی در چکیده ۲۸
- شکل ۲-۱۹: اشتباهات تایپی و نگارشی در چکیده ۲۹
- شکل ۲-۲۰: اشتباهات تایپی و نگارشی در چکیده ۳۰
- شکل ۲-۲۱: اشتباهات تایپی و نگارشی و همگراف ۳۰
- شکل ۲-۲۲: اشتباهات تایپی و نگارشی و همگراف ۳۱
- شکل ۲-۲۳: اشتباهات تایپی و نگارشی در چکیده ۳۱
- شکل ۲-۲۴: اشتباهات تایپی و نگارشی و همگراف ۳۲
- شکل ۲-۲۵: اشتباهات تایپی و نگارشی در چکیده ۳۳
- شکل ۲-۲۶: اشتباهات تایپی و نگارشی در چکیده ۳۳
- شکل ۲-۲۷: اشتباهات تایپی و نگارشی در چکیده ۳۴

- شکل ۲-۲۸: اشتباهات تایی و نگارشی در چکیده ۳۵
- شکل ۲-۲۹: عدم انطباق عنوان لاتین با فارسی مدرک ۳۶
- نمودار ۱-۳: ساختار زبانی عبارت جستجو ۴۱
- جدول ۱-۳: صورت های متعدد نوشتاری در بازیابی ۴۳
- شکل ۱-۳: صورتهای متعدد نوشتاری کسره اضافه ۴۵
- شکل ۲-۳: اشتباهات تایی در چکیده ۴۶
- شکل ۳-۳: اشتباهات ناشی از تفاوت در کدهای نوشتاری ۴۹
- شکل ۴-۳: رکوردهای دارای اطلاعات ناقص ۵۰
- شکل ۵-۳: رکورد دارای چکیده نادرست ۵۰
- شکل ۶-۳: عنوان نادرست ۵۱
- شکل ۱-۴: مثالی برای همه ارجاعات ۵۸
- شکل ۲-۴: چکیده دارای خطاهای املائی ۶۰
- شکل ۲-۴: چکیده دارای خطاهای املائی ۶۱

فصل اول

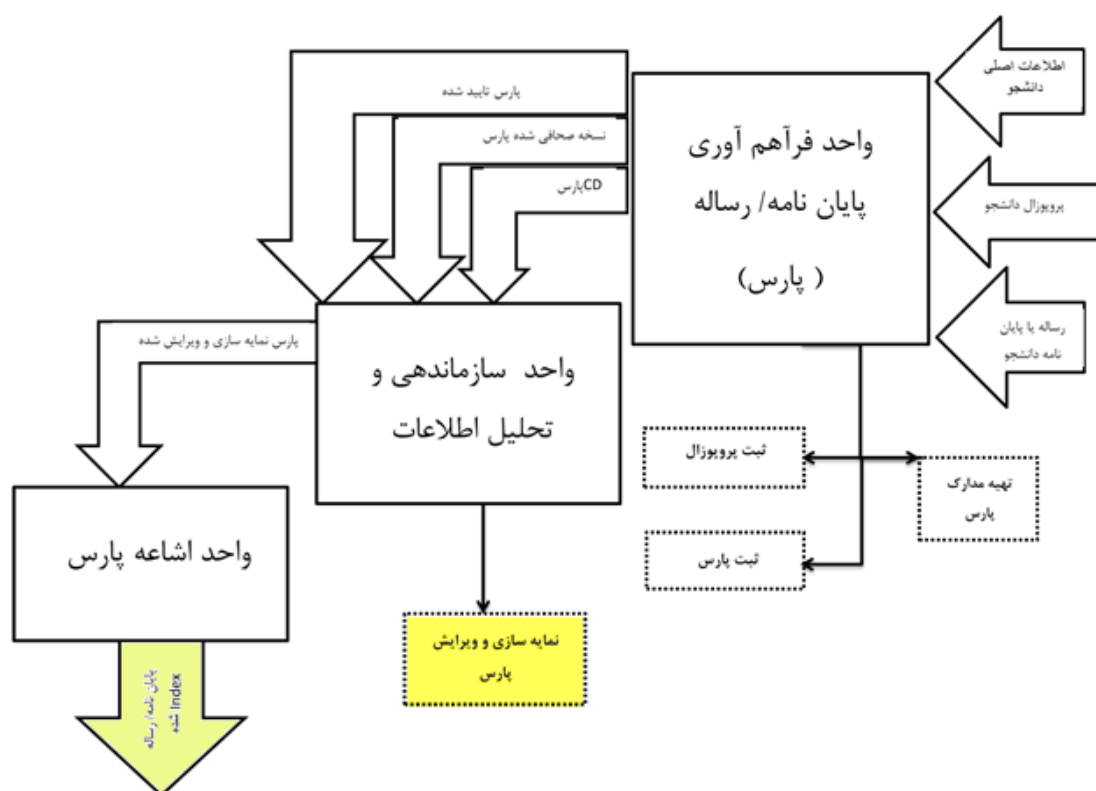
معرفی طرح پژوهشی

۱. مقدمه

پردازش متن یکی از چالش‌هایی است که سیستم‌های بازیابی اطلاعات با آن روبرو هستند. ساختار زبانی متن شامل ویژگی‌های ساختوازی-نحوی-معنایی و لغوی متن تأثیر بسزایی بر دقت بازیابی دارد. در سامانه گنج به عنوان مثال، مسائل خط فارسی از جمله عواملی است که سبب کاهش دقت بازیابی می‌شود. هدف از این طرح شناسایی این عوامل زبانی، بررسی و دسته‌بندی آنها بر اساس ادبیات موجود و در نهایت ارائه راهکارها یا راه‌حل‌هایی برای رفع یا بهبود آنها بر پایه مطالعات و روش‌های بکار گرفته شده در پژوهش‌های موجود در این زمینه است.

۲. بیان مسئله

در این طرح هدف بر آن است تا حد امکان مسائل بازیابی و سامان‌دهی اطلاعات به ویژه مسائلی که ریشه در پردازش متن و ساختار زبانی آن دارد، بر اساس ادبیات موجود شناسایی شود. پس از مطالعه و دسته‌بندی آنها، راه‌کارهای لازم بر اساس ادبیات موجود ارائه شود. نمودار جریان ورودی‌ها و خروجی‌های سه بخش معاونت اطلاعات علم و فناوری ایران (معاونت آفا) در نمودار زیر درج گشته است:



نمودار ۱-۱: اطلاعات ورودی و خروجی مرکز اطلاعات و مدارک علمی ایران (اقتباس از احترامی ۱۳۹۵)

۱-۲. پیشینه پژوهش

در زمینه مسائل و مشکلات پردازش متن فارسی پژوهش‌های متعددی صورت گرفته است که به تعدادی از آنها در بخش «ذکر منابع» اشاره شده است. نتایج گزارش‌های متعدد تدوین شده از سوی مرکز اطلاعات و مدارک علمی ایران به عنوان مثال «گزارش بررسی پایگاه اطلاعات گنج و ارائه راه‌حل‌ها» و نیز گزارش دفتر تحقیق و توسعه (یکپارچه سازی کاراکترها و شناسایی اشکال‌های فراداده مدارک ویرایش شده در نسخه تازه پایگاه اطلاعات علمی ایران گنج، اسفند ۹۵) نشان می‌دهد که سیستم بازیابی اطلاعات گنج در پردازش متن فارسی با مسائلی روبرو است. بررسی منابع زیر نیز نشان می‌دهد که مشکلات بازیابی اصطلاحات علمی شیمی نیز به طور عمده برخاسته از مسائل خط و زبان فارسی است. راه‌حل ارائه شده در این منابع به طور کلی شامل استاندارد کردن، یکسان‌سازی و ارجاع‌دهی کلیدواژه‌های تخصصی است.

- سمائی، سید مهدی. و رجبی، تقی. (۱۳۷۸) مشکلات یکسان‌نویسی اسامی ترکیب‌های شیمیایی آلی در زبان فارسی، نخستین هم‌اندیشی مسائل واژه‌گزینی و اصطلاح‌شناسی، تهران: فرهنگستان زبان و ادب فارسی.

- رجبی، تقی. و همکاران (۱۳۷۹)، بررسی اطلاعات علمی شیمی و مهندسی شیمی ایران در بانک جامع اطلاعات مرکز به منظور بهینه سازی آنها (پایان نامه‌ها)، گزارش طرح پژوهش، تهران: مرکز اطلاعات و مدارک علمی ایران.

- رجبی، تقی. و همکاران (۱۳۸۴). بهینه‌سازی اطلاعات علمی شیمی و مهندسی شیمی (سمینارها) در بانک جامع اطلاعات پژوهشگاه اطلاعات و مدارک علمی ایران و طراحی نرمی‌افزار بانک ترکیبات شیمیایی، گزارش طرح پژوهش، تهران: پژوهشگاه اطلاعات و مدارک علمی ایران.

در گزارش طرح زیر آمده است که هدف از انجام آن تدوین سیستمی است که با جستجو و دسته‌بندی کلیدواژه‌های لاتین، صورت‌های متفاوت کلیدواژه‌های فارسی شامل صورت‌های مترادف، جمع /مفرد عبارت، اشکال مختلف تایپ شده و صورت‌های دارای معنای مرتبط را بازیابی و در اختیار کاربر قرار دهد.

- پارسی اصفهانی، امیر. (۱۳۷۹)، ساماندهی کلیدواژه‌های پایگاه‌های مرکز اطلاعات و مدارک علمی ایران، گزارش طرح پژوهشی، تهران مرکز اطلاعات و مدارک علمی ایران

۲-۲. مراحل اجرایی پژوهش

گزارش‌های پژوهش طی دو گام زیر تهیه و تدوین می‌شود.

الف- شناسایی مسائل پردازش متن در سازمان‌دهی و بازیابی اطلاعات سامانه اطلاعات علمی ایران به روش مطالعه و مصاحبه

ب- ارائه راهکار برای مسائل بالا به روش مطالعه

۳-۲. پرسش‌های پژوهش

الف- مسائل زبانشناختی سازمان‌دهی و بازیابی اطلاعات در سامانه اطلاعات علمی ایران (گنج) کدامند؟

ب- چه راه‌حلهایی برای حل هر یک از مسائل منتج از ساختار زبانشناختی متن در سازمان‌دهی و بازیابی وجود دارد؟

۴-۲. هدف‌های پژوهش

الف- شناسایی مسائل زبانشناختی پردازش متن برای سازمان‌دهی و بازیابی اطلاعات در سامانه اطلاعات علمی ایران (گنج)

ب- بررسی و دسته‌بندی آنها بر اساس ادبیات موجود

ج- ارائه راهکارها یا روش‌هایی که برای بهبود یا رفع آنها در ادبیات مربوط به آن موجود است.

فصل دوم

مسائل زبانشناختی سازمان‌دهی اطلاعات

۲. مقدمه

در این فصل مسائل زبانشناختی که در مرحله سازمان‌دهی اطلاعات مشاهده می‌شود، براساس مطالعه میدانی و کتابخانه‌ای معرفی و بررسی می‌شوند. این مرحله خود شامل دو مرحله تکمیل اطلاعات کتابشناختی و فایل‌های ضمیمه و سپس نمایه‌سازی مدارک علمی است.

۱-۲. تکمیل اطلاعات کتابشناختی و فایل‌های ضمیمه

مدارک علمی شامل پایان‌نامه و رساله ابتدا در سامانه ثبت پژوهشگاه علوم و فناوری اطلاعات ایران ثبت و ذخیره می‌شود. ثبت مدارک علمی مذکور در سامانه ثبت شامل (۱) ذخیره اطلاعات کتابشناختی مدرک و (۲) ذخیره فایل‌های ضمیمه به صورت PDF و WORD توسط کاربر (دانشجو) است. اطلاعات کتابشناختی مدرک شامل موارد زیر است.

جدول ۱-۲: اطلاعات کتابشناختی مدرک (اقتباس از احترامی ۱۳۹۵)

نام فیلد	ویژگی‌های فیلد	الزامی	توضیحات
شماره دانشجویی	فقط عدد	*	کمترین مقدار صفر
مقطع تحصیلی	C	*	
سازمان مادر	C	*	
دانشگاه / پژوهشگاه	C	*	
دانشکده / پژوهشکده	C	*	لیست دانشگاه‌ها بر اساس سازمان مادر، پر

می شود. (در صورتی که در لیست نباشد، امکان اضافه کردن به صورت آزاد وجود دارد).			
حداکثر ۱۰۰۰ کاراکتر	*		عنوان پایان نامه / رساله
	*	C	موضوع اصلی
لیست رشته ها بر اساس موضوع اصلی، پر می شود. در صورتی که رشته در لیست وجود نداشته باشد هم رشته و هم گرایش به صورت آزاد باید وارد شوند.	*	C	رشته
لیست گرایش ها بر اساس رشته، پر می شود. در صورتی که گرایش در لیست وجود نداشته باشد، این فیلد باید به صورت آزاد وارد شود.	*	C	گرایش
امکان انتخاب چندمورد	*	C	زبان پایان نامه
امکان انتخاب چندمورد	*	C	چکیده پایان نامه به زبان
امکان انتخاب چندمورد	*	C	مشخصات ظاهری پایان نامه
فرمت تاریخ شمسی	*		تاریخ دفاع پایان نامه
کمترین مقدار صفر	*	C	تعداد صفحات
حداقل یک یا دو استاد راهنما	*		نام به فارسی
	*		نام خانوادگی به فارسی
توضیحات	الزامی	ویژگی های فیلد	نام فیلد
رایانامه یا کدملی برای جستجو استفاده می شود و ورود یکی از این دو فیلد الزامی است.	*		کد ملی
			رایانامه
			تلفن همراه

	نام به زبان انگلیسی		
	نام خانوادگی به زبان انگلیسی		
استاد مشاور	نام به فارسی	*	همانند استاد راهنما (با این تفاوت که مقدار حداقل تعداد برای استاد مشاور صفر است)
	نام خانوادگی به فارسی	*	
	کد ملی	*	
	رایانامه		
	تلفن همراه		
	نام به زبان انگلیسی		
	نام خانوادگی به انگلیسی		
توضیحات			
چکیده		*	حداکثر ۵۰۰۰ کاراکتر
فهرست مطالب		*	
فهرست جدول‌ها / شکل‌ها / نمودارها			
فهرست منابع فارسی		*	
فهرست منابع انگلیسی			
کلیدواژه		*	حداقل تعداد کلیک واژه‌ها دو مورد است در صورتی که کلید واژه در لیست نباشد امکان ورود به صورت آزاد وجود دارد
عنوان انگلیسی		*	صرفاً حروف و اعداد به زبان انگلیسی حداکثر ۱۰۰۰ کاراکتر
چکیده انگلیسی		*	صرفاً حروف و اعداد به زبان انگلیسی ۵۰۰۰ حداکثر کاراکتر
فایل Word		*	قالب‌های فایل doc, docx مجاز
فایل PDF		*	قالب فایل pdf مجاز

فایل در فرمت های دیگر			
-----------------------	--	--	--

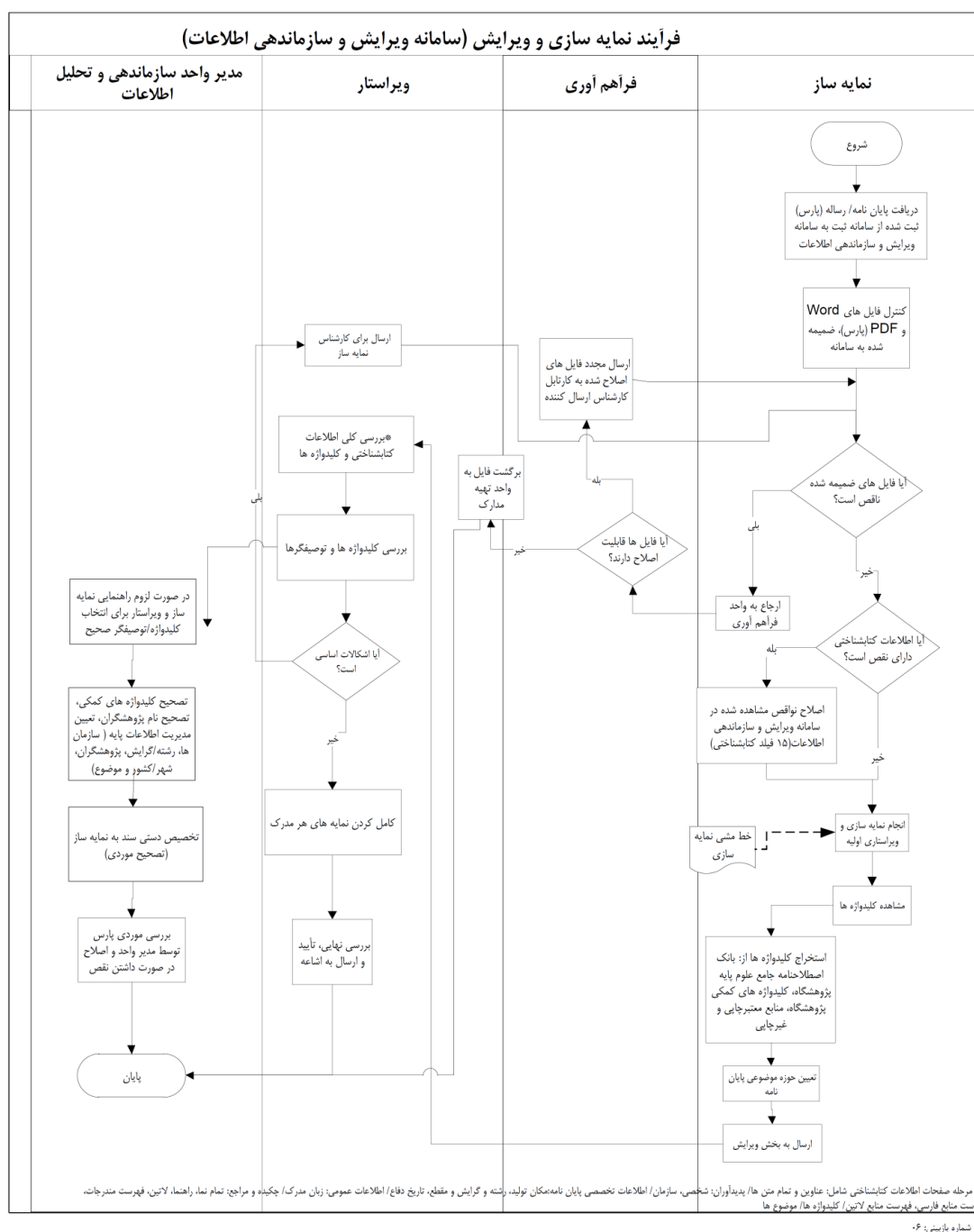
مدرکی که در سامانه ثبت پژوهشگاه گردآوری می‌شوند سپس برای نمایه سازی و ویراستاری وارد سامانه ویرایش و سازماندهی اطلاعات علمی می‌شوند. سازماندهی مدارک به طور کلی در سه سطح صورت می‌گیرد:

۱- نمایه‌سازی: کارشناس نمایه‌ساز

۲- ویرایش: ویراستار

۳- کنترل نهایی: مدیر واحد سازمان‌دهی

این سه سطح در فلوچارت زیر قابل مشاهده است.



نمودار ۱-۲: فرایند نمایه سازی مدرک (اقتباس از احترامی ۱۳۹۵)

۱-۱-۲. مسائل مربوط به تکمیل اطلاعات کتابشناختی و فایل های ضمیمه

اولین مسائل نمایه ساز در رابطه با فایل های ضمیمه و اطلاعات کتابشناختی که ثبت شده اند، به صورت زیر می تواند باشد:

الف- مدرک ثبت‌شده اطلاعات کتاب‌شناختی دارد اما فاقد فایل‌های ضمیمه PDF و WORD است. در این صورت مدرک مذکور به بخش فراهم‌آوری مدارک ارجاع داده می‌شود.

ب- فایل‌های ضمیمه PDF و WORD دارد اما اطلاعات کتاب‌شناختی ندارد یا ناقص است. در این صورت نمایه‌ساز با جستجو و بررسی فایل‌های ضمیمه، اطلاعات کتاب‌شناختی را استخراج می‌کند.

ج- علاوه بر آن که اطلاعات کتاب‌شناختی ندارد یا ناقص است، فایل‌های ضمیمه PDF و WORD نیز ناقص است. در این حالت نمایه‌ساز با جستجو در سایت دانشگاه محل اخذ مدرک سعی در یافتن فایل‌های ضمیمه و تکمیل اطلاعات کتاب‌شناختی مدرک می‌نماید (شکل زیر). مسئله قابل توجه در اینجا مدارکی است که همچنان ناقص باقی می‌مانند یعنی اطلاعات کتاب‌شناختی یا فایل‌های ضمیمه آنها با جستجوی نمایه‌ساز نیز پیدا نمی‌شوند و ناقص باقی می‌مانند. به نظر می‌رسد این مدارک در کارتابل نمایه‌ساز و ویراستار در ذخیره موقت باقی می‌ماند.

دانشگاه آزاد اسلامی

واحد گرمسار

دانشکده مهندسی برق، رایانه و فناوری اطلاعات، گروه کامپیوتر

پایان نامه برای درج کارشناسی ارشد رشته جغرافیا و برنامه ریزی توریسم، دانشگاه آزاد اسلامی واحد گرمسار

عنوان

آسیب شناسی امنیت گردشگری در استانهای مرزی، مطالعه موردی استان سیستان و بلوچستان

استاد راهنما

آقای دکتر سید جمال دریاباری

استاد مشاور

آقای دکتر سعید کریم پور

نگارش

طاهره غلامی

سال تحصیلی ۹۳-۹۲

ب

شکل ۲-۱: مثالی برای اشتباه در ثبت اطلاعات

شکل بالا نمونه‌ای از اشتباه در مدارک پیوست و ثبت اطلاعات از سوی کاربر را نشان می‌دهد: اشتباه در ثبت نام دانشکده (دانشکده مهندسی برق) و درج نام استاد راهنما به صورت ناقص. نمایه‌ساز با

مراجعه با فایل پیوست و یا جستجو در سایت دانشگاه و گردآوری اطلاعات، در نهایت نام دانشکده را به صورت دانشکده علوم انسانی و نام استاد را به سیدجمال‌الدین دریاباری تصحیح نمود.

د- در مواردی که اطلاعات کتاب‌شناختی یا فایل‌های ضمیمه به زبان انگلیسی است، نمایه‌ساز اگر عنوان انگلیسی را به فارسی ترجمه نماید در داخل کروش در بخش عنوان فارسی می‌آورد و اگر تخصصی باشد و یا نتواند در عنوان فارسی خط تیره گذاشته می‌شود. اگر مدرک به زبان خارجی دیگری به جز انگلیسی باشد، یا در مورد درستی ترجمه فارسی مدرک تردید باشد، در عنوان فارسی ترجمه نمی‌آید.

۲-۲. نمایه‌سازی اطلاعات کتابشناختی

اطلاعات کتابشناختی ذخیره شده در مرحله ثبت مدرک به صورت زیر در سامانه ویرایش و نمایه‌سازی قابل مشاهده و در دسترس هر یک از سه سطح اشاره شده در بالا یعنی نمایه‌ساز، ویراستار و مدیر واحد سازماندهی است.

محدود سازی اسناد					
عناوین و تمام متن‌ها	پیداآوران	اطلاعات تخصصی پایان نامه	اطلاعات عمومی	چکیده‌ها و مراجع	کلیدواژه‌ها
موضوع‌ها					

شکل ۲-۲: پنجره اطلاعات کتابشناختی در سامانه نمایه‌سازی

به عبارت دیگر نمایه‌ساز، ویراستار و در نهایت مدیر واحد سازماندهی مدارک علمی را بر اساس هفت موردی که در شکل بالا آمده است، سازماندهی می‌کنند و در سامانه گنج ذخیره می‌نمایند.

۲-۲-۱. عناوین و تمام متن‌ها

در این قسمت عنوان فارسی و لاتین مدرک مورد نظر وارد می‌شود. شکل زیر

گزارشات

عناوین

سنتز و بررسی خواص الکتریکی و اپتیکی نانوپودر ZnO آلاینده با اکسید مس

عنوان اصلی

☐ آیا عنوان دارای فرمول است؟

Synthesis and study of electrical and optical properties of ZnO nanopowders doped with copper oxide

عنوان لاتین

عنوان قراردادی

عنوان فرعی

قابل های تمام متن

دریافت فایل PDF

دریافت فایل Word

ارجاع به سیستم کنترل

شکل ۲-۳: ینجره عنوان در سامانه نمایه سازی

ساختار زبانی عنوان فارسی یا لاتین یک مدرک (پایان نامه، مقاله، طرح پژوهشی) معمولاً یک یا چند عبارت اسمی است. ممکن است گاهی به صورت یک جمله نیز باشد.

۲-۲-۲. پدید آوران

پدیدآوران شامل نام و نام خانوادگی نویسنده یا نویسندگان، استاد راهنما و مشاور در پایان نامه‌هاست.

صفحه: استادان ویرایش و سازماندهی اطلاعات علمی - کاربر: زهرا دانشورایی - نقش: مدیر سیستم

نویسندگان: مدیریت کاربران - مدیریت اطلاعات پایه - ابزار گزارشات

محدود سازی استاد

کاربرانت

اطلاعات شخصی پاپان نامه

اطلاعات عمومی

چکیده ها و مراجع

کلیدواژه ها

موضوع ها

پدیدآور نام ها

پدیدآور شخصی

نام	نام خانوادگی	نقش	ایمیل	کد ملی	تلفن	تلفن همراه
مهرداد	پارسایی برارجانی	پدیدآور				
سعید	فرشته	استاد مشاور				
علیرضا	حسینزاد دویی	استاد راهنما				
علیرضا	حسینزاد دویی	استاد راهنما				

نام: مهرداد

نام خانوادگی: پارسایی برارجانی

نقش: پدیدآور

ایمیل:

کد ملی:

تلفن:

تلفن همراه:

محدود سازی اولیه پدیدآورندگان

محدود سازی اولیه پدیدآور انتخاب شده

ویرایش پدیدآور انتخاب شده

افزودن پدیدآور جدید

حذف پدیدآور

دریافت شده

ویرایش شده

اضافه شده

حذف شده

ویرایش پاپان نامه

ویرایشگر سازمان	دانشگاه	دانشگاه	نقش	آدرس	آدرس و پست	ایمیل	تلفن
وزارت علوم، تحقیقات و فناوری	دانشگاه سمنان	دانشگاه تهران	دانشگاه				

شکل ۲-۴: ینجره یدید آوران در سامانه نمایه سازی

شناسایی مسائل و ارائه راهکارهای زبانشناختی سازمان‌دهی و بازیابی اطلاعات در گنج □ ۱۵

ساختار زبانی نام پدیدآوران واژه و عبارت است. نقص در درج نام پدیدآوران دقت بازیابی را کاهش می‌دهد.

۳-۲-۲. اطلاعات تخصصی پایان‌نامه

اطلاعات تخصصی شامل رشته/گرایش، مقطع تحصیلی و شهر محل اخذ مدرک است.

شکل ۲-۵: پنجره اطلاعات تخصصی پایان‌نامه در سامانه نمایه‌سازی

دامنه و نوع این اطلاعات در سامانه از پیش تعیین شده است. کاربر و یا نمایه‌ساز برای ثبت اطلاعات تنها از بین گزینه‌های موجود انتخاب می‌کند. بنابراین این بخش در حوزه بررسی مسائل زبانشناختی قرار نمی‌گیرد.

۴-۲-۲. اطلاعات عمومی

اطلاعات عمومی نیز فیلدهای مشخصی دارد که توسط کاربر تعیین می‌شوند.

شکل ۲-۸: پنجره چکیده راهنما در سامانه نمایه سازی

چکیده راهنما معمولاً یک یا چند پاراگراف را تشکیل می‌دهد.

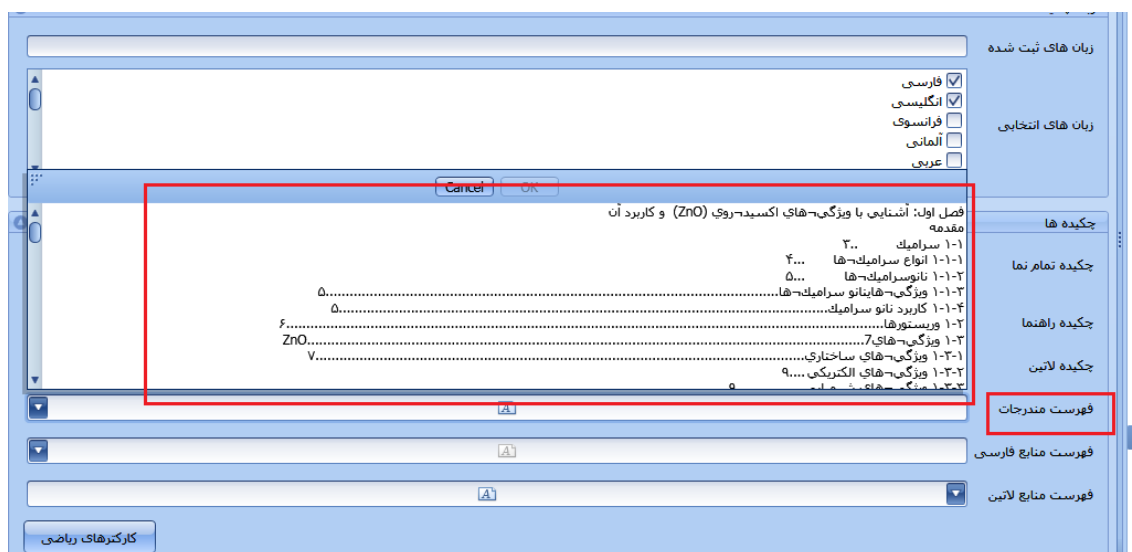
ج- چکیده لاتین

چکیده لاتین معمولاً به زبان انگلیسی است و کاربر در سامانه وارد می‌کند. چکیده لاتین نیز می‌تواند از یک یا چند پاراگراف تشکیل شده باشد. در برخی موارد کاربر چکیده انگلیسی را وارد نمی‌کند. یا چکیده فارسی وارد شده با چکیده درج شده در فایل پایان نامه همخوانی ندارد.

شکل ۲-۹: پنجره چکیده لاتین در سامانه نمایه سازی

د- فهرست مندرجات

فهرست مطالب به همراه شماره صفحه آن در قسمت فهرست مندرجات ثبت می‌شود.

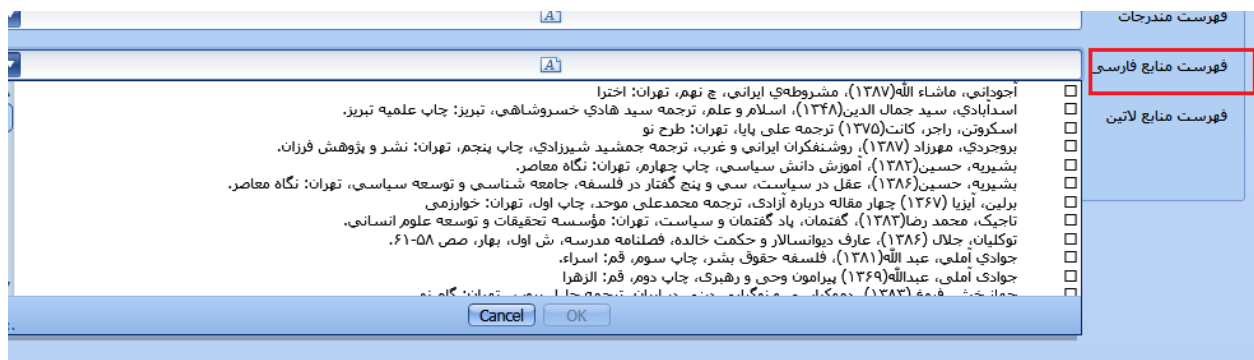


شکل ۲-۱: پنجره فهرست مندرجات در سامانه نمایه سازی

ساختار زبانی فهرست مندرجات عبارت اسمی است.

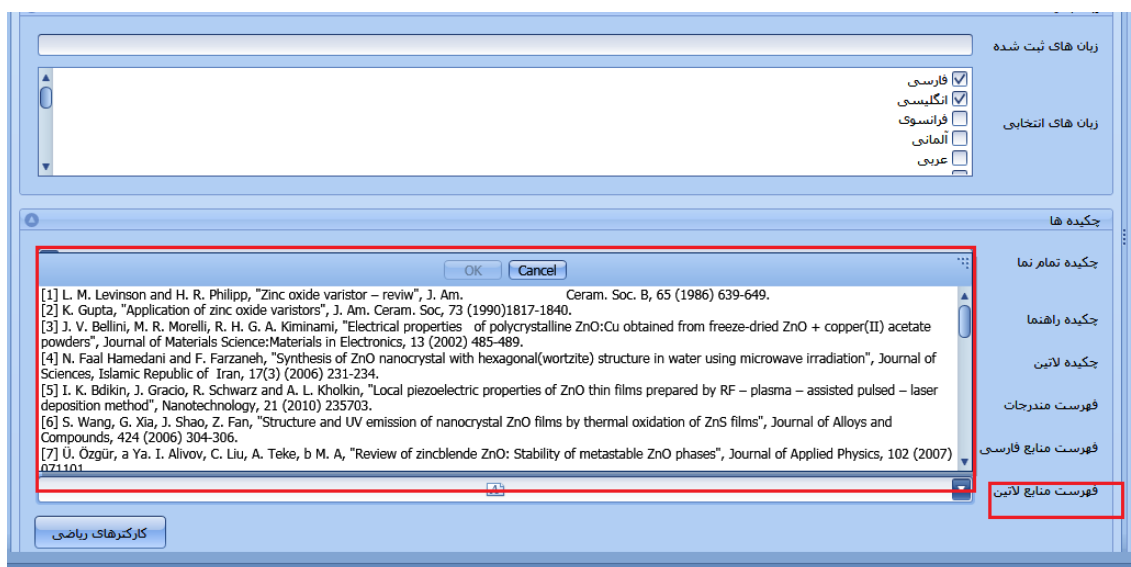
ه - فهرست منابع فارسی

فهرست منابع فارسی یا لاتین را کاربرد می کند و بنابراین خطاهای بسیاری می تواند روی دهد.



شکل ۲-۱۱: پنجره فهرست منابع فارسی در سامانه نمایه سازی

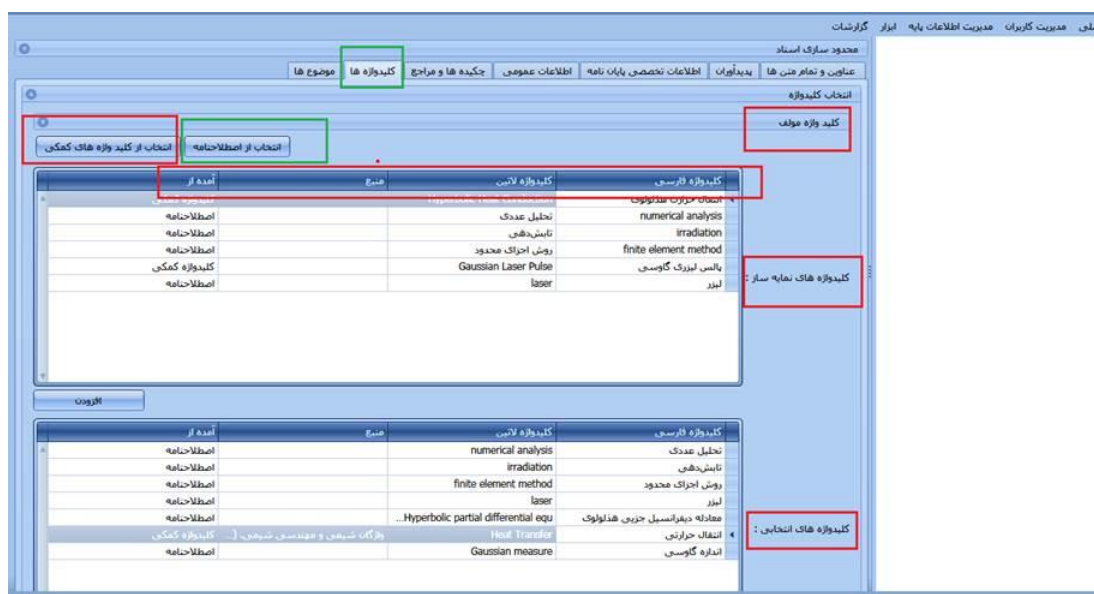
و - فهرست منابع لاتین



شکل ۲-۱۲: پنجره فهرست منابع لاتین در سامانه نمایه سازی

۲-۲-۶. کلیدواژه‌ها

ساختار زبانی کلیدواژه‌ها معمولاً عبارت اسمی یا صفتی است. علاوه بر آن صورت‌های مخفف فارسی یا لاتین نیز می‌تواند به عنوان کلیدواژه بکار رود. کلیدواژه‌ها از دو منبع است: ۱. اصطلاح نامه‌های پژوهشگاه؛ ۲. کلیدواژه‌های کمکی. کلیدواژه‌ایی که در هیچ‌کدام از این دو نباشد، پیشنهاد و بررسی و سپس وارد کلیدواژه‌های کمکی می‌شود.

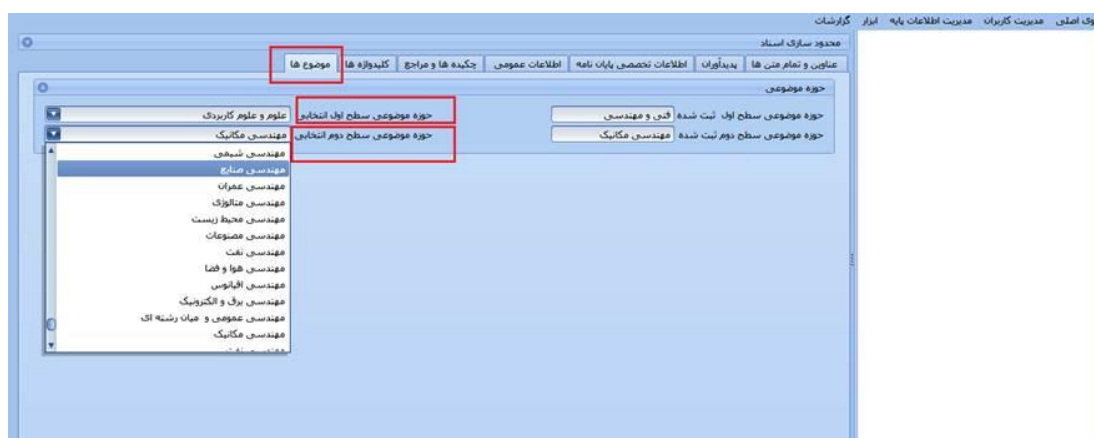


شکل ۲-۱۳: پنجره کلیدواژه‌ها در سامانه نمایه سازی

شناسایی مسائل و ارائه راهکارهای زبانشناختی سازمان‌دهی و بازیابی اطلاعات در گنج □ ۲۰

۲-۲-۷. موضوع‌ها

فیلدهای حوزه موضوعی مدرک نیز از پیش تعیین شده‌اند و فقط کاربر یکی را انتخاب می‌کند. البته مسائلی که ممکن است نمایه ساز با آن روبرو شود، تغییر حوزه‌های موضوعی یا حذف و ایجاد حوزه‌های موضوعی جدید است. بنابراین اگر این تغییرات بر میزان بازیابی تأثیر بگذارد، به عنوان مثال با تغییر یک حوزه موضوعی، مدارک پیشین در اختیار کاربر قرار نگیرد، از میزان دقت بازیابی کاسته می‌شود.



شکل ۲-۱۴: پنجره موضوع‌ها در سامانه نمایه سازی

۲-۳. مسائل زبانشناختی نمایه سازی

بر اساس آنچه گفته شد در حوزه‌های زیر که کاربر اطلاعات کتابشناختی را وارد می‌کند و نمایه ساز به تصحیح یا طبقه‌بندی آنها می‌پردازد، مسائل زبانشناختی ناشی از ساختار زبانی اطلاعات می‌تواند روی دهد. ساختار زبانی این حوزه‌ها در زیر آمده است.

جدول ۲-۱: ساخت زبانی اطلاعات کتابشناختی

انواع اطلاعات کتابشناختی	ساخت زبانی
چکیده (تمام‌نما، راهنما)	پاراگراف: جمله
عنوان پدیدآوران شخص فهرست مندرجات فهرست منابع	عبارت (بیش از یک واژه)
کلیدواژه	واژه یا عبارت

طبیعی است که دامنه مسائل زبانی در سطح پاراگراف و جمله بیشتر از سطح عبارت و واژه است. در سطح پاراگراف مسائلی از قبیل علائم سجاوندی، درج فاصله یا نیم‌فاصله، صورت‌های مختلف نوشتاری و غیره مسائل متنوعی را ایجاد می‌کند.

با بررسی میدانی مدارک موجود در پایگاه گنج مسائل زبانی در سطح هر یک از اطلاعات کتابشناختی به صورت زیر قابل مشاهده است.

۲-۳-۱. کلیدواژه‌ها

الف- معادل‌یابی واژه‌های لاتین

یکی از مسائل ساماندهی، کاربرد معادل‌های متعدد فارسی برای یک واژه لاتین است. هرچند عکس آن نیز امکانپذیر است یعنی وجود صورت‌های متفاوت لاتین برای یک واژه فارسی. مثال زیر نمونه‌ای از آن را نشان می‌دهد.

جدول ۲-۲: مثالی برای معادل‌های متعدد واژه‌های لاتین

معادل انگلیسی	معادل فارسی	منبع/تعریف
Taxonomy	علم رده‌شناسی، طبقه‌شناسی، طبقه‌بندی‌شناسی، رده‌بندی‌شناسی، میانی رده‌بندی، تاکسونومی	قرهنگ جامع کتیداری و اطلاع‌رسانی، مختاری معمار
Taxonomy	طبقه‌شناسی، طبقه‌بندی، رده‌شناسی	واژه نامه کتیداری و دکوماتسیون
Taxonomy	آرایه‌شناسی	قرهنگستان زبان و ادب فارسی
Taxonomy	تاکسونومی، طبقه‌بندی، علم رده‌بندی، تاکسونومی، دانش طبقه‌بندی، رده‌بندی، طبقه‌شناسی	مرکز نشر
Taxonomy	رده‌بندی شناسی	دانشنامه کتیداری و اطلاع‌رسانی، مرحوم یوری سلطانی
Taxonomy	ردگان شناسی	دائرة المعارف جامع دریایی و بندری
Taxonomy	تاکسونومی	میترا صمیمی، مقاله: تاکسونومی و نقش آن در سازماندهی اطلاعات
Taxonomy	آرایه‌شناسی	اصطلاحنامه علوم کشاورزی
Taxonomy	علم رده‌بندی	اصطلاحنامه علوم زمین ^{۱۱}
Taxonomy	آرایه‌شناسی	اصطلاحنامه علوم زیستی رده‌بندی جانوری را به آرایه‌شناسی جانوری، رده‌بندی گیاهان را به آرایه‌شناسی گیاهی ارجاع داده است
Taxonomy	رده‌شناسی، رده‌بندی جانداران	قرهنگ دانش شناسی؛ محمدرضا محمدی قر
Taxonomy	آرایه‌بندی	قرهنگ نوین کشاورزی و منابع طبیعی، جلد چهاردهم شیلات. قرهنگستان علوم جمهوری اسلامی ایران (بررسی اصول کلی رده‌بندی علمی گیاهان و جانوران)
Classification	رده‌بندی	قرهنگستان زبان و ادب فارسی
Classification	رده‌بندی، طبقه‌بندی	واژگان کتیداری و اطلاع‌رسانی
Classification	رده‌بندی	دانشنامه کتیداری و اطلاع‌رسانی، مرحوم یوری سلطانی
Classification	رده‌بندی، طبقه‌بندی	قرهنگ جامع کتیداری و اطلاع‌رسانی، مختاری معمار

مشکل ساماندهی این است که آیا نمایه‌ساز تنها باید یک معادل را در نظر بگیرد یا همه معادل‌های فارسی یک واژه لاتین را. از سویی در موارد بسیاری که معادل‌های فارسی با یکدیگر همپوشانی دارد (مثل taxonomy و classification که هر دو برابرهای فارسی یکسانی دارند) تکلیف چیست؟

به نظر می‌رسد ساماندهی این موارد باید به گونه‌ای صورت گیرد که معادل‌های متعدد در کنار هم آورده شود. در این صورت بازیابی آنها نیز بهتر صورت می‌گیرد.

جدول ۲-۳: صورتهای نوشتاری متعدد برای یک کلیدواژه

اچ ۹. ان ۲. نوع دوم ویروس آنفلوآنزا نوع آ اچ ۹. ان ۲. ویروس آنفلوآنزای نوع آ. اچ ۹. ان ۲. نوع دوم ویروس آنفلوآنزا نوع آ.	H _۱ N _۲ Subtype Influenza A Virus
--	---

در مثال بالا کاربرد علائم نگارشی نظیر نقطه، خط فاصله و رقم سبب شده تا صورت‌های متفاوت نگارشی حاصل شود. این تفاوت‌های نگارشی رکوردهای بازیابی متفاوتی را بدست می‌دهد.

ب- صورت‌های اختصاری

گاهی براساس معادل فارسی یک اصطلاح لاتین، صورت مختصر آن ساخته می‌شود. به عنوان مثال در نمونه‌های زیر صورت‌های مختصر «مجگم و ماد» بر اساس معادل فارسی اصلاحات ساخته شده است.

جدول ۲-۴: مثالی برای صورت اختصاری فارسی یک کلیدواژه لاتین

اصطلاح لاتین	صورت اختصاری لاتین	معادل فارسی	صورت اختصاری فارسی
sex addicts anonymous	(SAA)	معتادان جنسی گمنام	(مُجْگَم) (اعراب گذاری در بازیابی اطلاعات و جستجوی کاربر فارسی زبان انجام نمی‌شود)
voluntary counselling and testing	(VCT)	مشاوره و آزمایش داوطلبانه	(ماد) (کاربر از این اختصار استفاده نمی‌کند اختصار فارسی و انگلیسی همخوانی ندارد)

بهتر است ساماندهی این موارد به نحوی صورت گیرد که همراه با صورت اختصار یافته، معادل عبارت نیز ذخیره شود به گونه‌ای که توسط کاربر در بازیابی قابل مشاهده باشد.

ت-آوانگاری و حرف‌نگاری کلیدواژه‌های لاتین

آوانگاری و یا حرف‌نگاری کلیدواژه‌های لاتین نیز از جمله مسائلی است که نمایه‌ساز با آن روبرو است. در برخی موارد که به ویژه وقتی که برای صورت‌های مخفف لاتین معادلی در فارسی موجود نیست، ممکن است نمایه‌ساز دست به آوانگاری یا حرف‌نگاری صورت لاتین بنماید.

۲-۳-۲. پدیدآوران

• نام‌های خاص

نام خاص افراد به سه صورت پدیدآور، استاد راهنما و استاد مشاور ساماندهی می‌شود. یکی از مواردی که در ساماندهی مدارک مسئله‌ساز است تنوع‌ها و صورت‌های متفاوتی است که برای نام یک فرد در مدارک دیده می‌شود. به عنوان مثال در زیر نام یک فرد در مدارک متفاوت به شکل‌های مختلف دیده شده است: گاهی به صورت کوتاه (فیض فلاح شمس) گاهی کامل (میرفیض شمسی لیالستانی). ساماندهی مواردی از این دست با مشکلات متعددی روبرو است.

الف- اینکه آیا همه این نام‌ها متعلق به یک فرد است؟ پاسخ به این سوال نیازمند بررسی، جستجو و گاهی دشوار است. سؤالی که در اینجا مطرح می‌شود این است که آیا جستجو و یکدست کردن این صورتهای متفاوت در حیطه وظایف یک نمایه‌ساز است؟ چگونه می‌توان صحت نتایج را تأیید نمود؟ بنابراین به نظر می‌رسد بررسی این صورتهای متفاوت و یکدست کردن آنها کاری وقتگیر است که در عین حال ممکن است صورت نهایی بدست آمده نیز درست نباشد.

فلاح شمس لیالستانی

میر فیض فلاح شمس لیالستانی

فیض فلاح شمس

میرفیض شمسی لیالستانی

فیض فلاح لیالستانی

میر فیض فلاح لیالستانی

ب- اینکه این نام‌های متعدد با صورت‌های نوشتاری متعدد نیز نوشته شده‌اند. به عنوان مثال کلمه «الله» در نام زیر به چند صورت نوشته شده است. این موارد به همراه عدم رعایت قواعد نگارشی مثل فاصله یا نیم‌فاصله، مشکلات تاییی علاوه بر مشکل نمایه‌سازی بر دقت بازیابی نیز تأثیر منفی دارد.

یداله مهر علی زاده
یدالله مهر علی زاده
یدا مهر علی زاده
ید اله مهر علی زاده
یداله مهر علی زاد
مهر علی زاده
اله مهر علی زاده

آنچه در مرحله ساماندهی مهم است این است که این صورت‌های متعدد به چه صورت نوشتاری نمایه‌سازی شوند به نحوی که بازیابی همه این موارد متعدد به درستی صورت گیرد و بتواند کاربر را به مدارک درست هدایت کند. مشکل ساماندهی در اینجا نیز این است که با ذخیره یک صورت نوشتاری امکان دسترسی به سایر صورت‌ها نیست.

ت- تغییر نام خاص افراد

نکته دیگر این که ساماندهی و بازیابی نام‌های خاص متأثر از هر تغییر است که در مورد نام‌های خاص ممکن است روی دهد. نام و نام فامیلی افراد ممکن است دچار مسائل متعدد شود به عنوان مثال نام کوچک، نام خانوادگی فرد تغییر کند، بخشی از آن حذف شود و یا اسم مستعار بکار برد یا آن را تغییر دهد. مشکل نمایه‌سازی این موارد شامل یافتن مرجع اسم و تغییرات مذکور است.

۲-۳-۳. چکیده و عنوان

- خطا در درج علائم و فرمول‌های ریاضی و شیمی

نمایه‌سازی «چکیده، عنوان و یا کلیدواژه‌های» دارای فرمول‌ها و علائم لاتین در حوزه‌هایی مانند ریاضی و شیمی یکی از مشکلات ساماندهی مدارک است.

(0) (0)

نتایج هندسی از خاصیت بورلی ثابت ایده‌آهای مونومیال

بابانامه . دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه صنعتی خواجه نصیرالدین طوسی - دانشکده ریاضی و کامپیوتر . 1393 . کارشناسی ارشد

موضوع: علوم و علوم کاربردی < ریاضیات

استاد راهنما: حسن حقیقی | استاد مشاور: فرشته ملک | پدیدآور: عبدالباسط بیابانی

دسترسی به فایل تمام‌متن پیشنهادیهایی (رکوردهایی) که نشانه «پی.دی.اف.» ندارند، شدنی نیست.

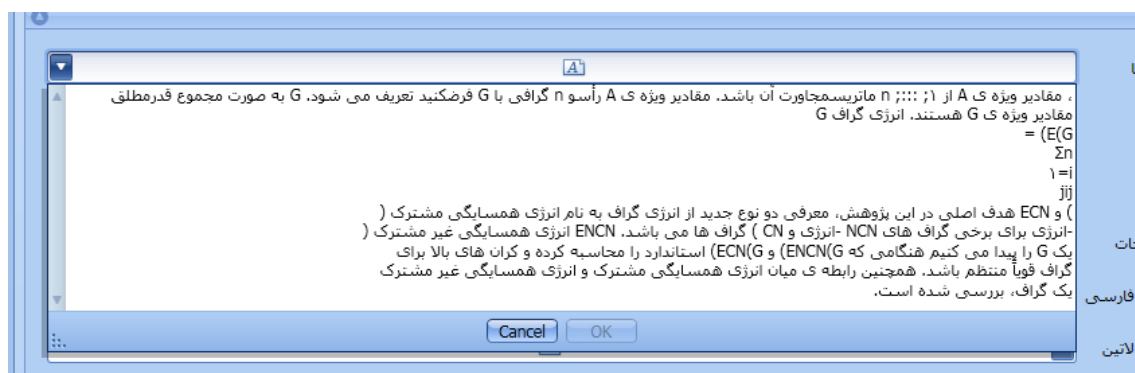
ایده آل آغازی یک ایده آل I در حلقه $R = k[x_1, \dots, x_n]$ ، که k یک میدان است، به شکلی ساده و قابل محاسبه‌ای اطلاعات هندسه‌جبری بسیار مفیدی را درباره حلقه خارج‌قسمتی $\frac{R}{I}$ بدست می‌دهد. اما این ایده آل آغازی بستگی به انتخاب دستگاه مختصات دارد و ممکن است تحت تعویض متغیر خطی تغییر کند. در نتیجه خواص استنتاج شده از ایده آل آغازی ایده آل I ممکن است تحت هر تغییر متغیر خطی پایدار نماند. برای رفع این مشکل، عمل یک زیرگروه معین از گروه تبدیلات خطی وارون‌پذیر روی I موسوم به تبدیلات بورلی را در نظر می‌گیریم و ایده آل آغازی ثبات این عمل را با $\mathfrak{g}(I)$ نشان داده و آن را ایده آل آغازین عام می‌نامیم. به‌خصوص اگر I ایده آل تعریف‌کننده یک وارینه تصویری Γ باشد، $\mathfrak{g}(I)$ تحت ترتیب القابای وارون، اطلاعات هندسی بسیار مفیدی درباره این وارینه در اختیار می‌گذارد. در این پایان‌نامه قصد داریم برخی از اطلاعات فوق را بطور مبسوط مورد بررسی قرار دهیم.

نمایه‌ها:

مجموعه بورلی | ایده‌آل تک جمله‌ای | توابع هیلبرت | هندسه جبری | ایده‌آل آغازین

شکل ۲-۱۵: مسائل درج فرمولها و علائم ریاضی و شیمی

مدرک بالا نمونه‌ای از یک چکیده نمایه‌سازی شده است. فرمول‌ها و علائم ریاضی به درستی ساماندهی نشده است. در واقع با متنی مواجه هستیم که فارسی و لاتین است. مسائل مربوط به درج کاراکترها، علائم و فرمول‌ها یک مسئله زبانشناختی محسوب نمی‌شود و در حوزه مسائل فنی ساماندهی قرار می‌گیرد. شکل زیر نمونه دیگری از چکیده حاوی علائم و فرمول که توسط کاربر در سامانه وارد شده است را نشان می‌دهد.



شکل ۲-۱۶: مسائل درج فرمولها و علائم ریاضی و شیمی

چکیده اصلی آن به صورت زیر است

چکیده

فرض کنید G گرافی با n رأس و A ماتریس مجاورت آن باشد. مقادیر ویژه $\lambda_1, \dots, \lambda_n$ از A ، مقادیر ویژه G هستند. انرژی گراف G به صورت مجموع قدرمطلق مقادیر ویژه G تعریف می‌شود.

$$E(G) = \sum_{i=1}^n |\lambda_i|$$

هدف اصلی در این پژوهش، معرفی دو نوع جدید از انرژی گراف به نام انرژی همسایگی مشترک (ECN) و انرژی همسایگی غیر مشترک ($ENCN$) گراف‌ها می‌باشد. ECN -انرژی و $ENCN$ -انرژی برای برخی گراف‌های استاندارد را محاسبه کرده و کران‌های بالا برای $ECN(G)$ و $ENCN(G)$ را پیدا می‌کنیم هنگامی که G یک گراف قویاً منتظم باشد. همچنین رابطه‌ی میان انرژی همسایگی مشترک و انرژی همسایگی غیر مشترک یک گراف، بررسی شده است.

واژه‌های کلیدی: نظریه‌ی طیفی گراف، انرژی گراف، انرژی برچسب، انرژی همسایگی مشترک، انرژی همسایگی غیر مشترک

همان‌طور که ملاحظه می‌شود ساختار زبانی این چکیده فارسی و لاتین است. به نظر می‌رسد ابتدا لازم است مسائل مربوط به درج فرمولها به لحاظ فنی حل شده سپس ساختار زبانی متن مورد بررسی قرار گیرد.

• اشتباهات تایپی و نگارشی

اشتباهات تایپی و املائی در همه هفت قالب نمایه‌سازی یعنی چیکده (فارسی - انگلیسی)، عنوان (فارسی - انگلیسی)، کلیدواژه (فارسی - انگلیسی)، پدیدآوران شخص (پدیدآور، استاد راهنما، استاد مشاور)، پدیدآوران سازمان (دانشگاه، دانشکده)، اطلاعات تخصصی پایان‌نامه (رشته/گرایش، مکان تولید) و حوزه موضوعی (در دو سطح) ممکن است روی دهد. اما درصد این اشتباهات در چکیده و عنوان و کلیدواژه‌ها بیشتر است. مثال زیر یک چکیده نادرست را نشان می‌دهد.

(0)

ارزیابی زیرساخت های توسعه گردشگری شهر یزد (با تأکید بر سیستم حمل و نقل)

دریافت فایل

بایان‌نامه . دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه علم و هنر - دانشکده علوم انسانی . 1394 .
 کارشناسی ارشد

موضوع: علوم انسانی <

استاد راهنما: سیدمحمود زنجیری | استاد مشاور: حامد فلاح | دانشجو: حمیدرضا کلانتری

صنعت گردشگری یکی از پزیر تر ین و متنو تر ین صن ابع در دنیا م ی ک باشند . یکی از مهم تر ین زیرساخت های این صنعت حمن و نقن بنوده به نحتو ی کنه جهناترردی را به ع نئوان پرئنده ای م ی ک شناسئند کنه حمن و نقن بال های آن بوده و ضئع و بنا بنئود آن به منزله عئدم ئوانئایی بنرواز آن پرئنده م ی ک باشند . در همنین ارتباط پژوهشئر استان یزد را بنا ئوخته به جا به هئای نئنی بیئعی، ئاری ک و یادهانئدن ک، منههیی فرهئرئی به عئنوان منورء ملالئعئائی انت ناب کنرءه و در سئند ارزیئابی زیرساخت هئای ئوسئعه گردشئرئی بنا ئاکئید بنر سئستم حمن و نقن م ی ک باشند و بنا شناسئایی و معرفئی مهم تر ین عوامن مئئرر بنر مئئی داخلى و خئارجی صنعت گردشئرئی اسئنتان و بنا اسئنتءاءه از ئحلئین منئاری سئئواء، ئحلئین شئنگاو و ئربئالئرئی فنئاری اسئئرئزئی هئا، راهبرءهئای ئوسئعه گردشئرئی استان را ارائه مئی ءهئء. جامعئه آمئاری ئءقیق را ؟ ئءر از خبرگان صنعت گردشئرئی شنام م ی ک شئوء کئ بنا اسئءءاءه از م صئاحیه و پرسشئنامه به الوئئ بنئدی ابءءاء چهارگانئه ئکئین سئئواء پرءاخئئنه شئئءه اسئئئ

نمائه ها:

یزء ۹۹۹ |

شکل ۲-۱۷: اشتباهات تایپی و نگارشی در چکیده

اشتباهات تایپی و املایی را می‌توان به دسته‌های مختلف تقسیم نمود.

الف- تایپ نادرست صورت واژه

با یک بررسی اولیه به نظر می‌رسد درصد اشتباهات تایپی به ویژه در چکیده‌ها بالاست. تصحیح دستی این خطاها از سوی نمایه‌ساز هم زمان‌بر است و هم وابسته است به میزان دقت نمایه‌ساز در تشخیص خطاها، تسلط او به صورت‌های نوشتاری زبان فارسی و لاتین، تسلط او به صورت‌های نوشتاری اصطلاحات علمی و نیز به عوامل جسمی مانند خستگی و غیره. به عنوان مثال کلمه «پژوهش» با اشتباهات تایپی متفاوت به صورت زیر دیده می‌شود.

پژوهش	بژوهش	پژهش	پژوش	پژژوهش	پژووهش	پژوهشش	پپژوهش
۵۳۸	۱۳	۱۳۶	۳۷	۱	۶	۲	۳

نمونه‌ای از این خطای تایپی در مدرک نمایه‌شده زیر مشاهده می‌شود.

بررسی رابطه دینداری و نگرش نسبت به دموکراسی؛ مطالعه موردی: دانشجویان دانشگاه تربیت معلم

پایان نامه . دولتی - دانشگاه پیام نور (وزارت علوم، تحقیقات، و فناوری) - دانشگاه پیام نور استان تهران - مرکز پیام نور تهران - دانشکده علوم اجتماعی . 1388 . کارشناسی ارشد

موضوع: علوم اجتماعی < جامعه شناسی

استاد راهنما: محمدجواد زاهدی مازندرانی | استاد مشاور: پروانه دانش | پدیدآور: مهران صولتی

دسترسی به فایل تمام متن پیشینه‌های (رکورد‌هایی) که نشانه «پی.دی.اف.» ندارند، شدنی نیست.

شکل ۲-۱۸: اشتباهات تایپی و نگارشی در چکیده

کلمه «بررسی» با اشتباهات تایپی زیر دیده می شود.

تعداد	اشتباه در تایپ کلمه «پژوهش»
۱۵۸۰	برسی
۳۶۵	بررسی
۵	ترسی
۱۹	بررسی
۱۰	بیرسی
۲۹	بررسی

نمونه زیر صورت نادرست این واژه را در عنوان (۳۷۷ مورد) نشان می‌دهد.

بررسی ارزشهای خانوادگی و تغییرات آن در ایران (1379-1385)

دریافت فایل

بایان‌نامه ، دولتی - دانشگاه پیام نور (وزارت علوم، تحقیقات، و فناوری) - دانشگاه پیام نور استان تهران - مرکز پیام نور تهران - دانشکده ادبیات و علوم انسانی ، 1390 ، کارشناسی ارشد

موضوع: علوم اجتماعی < جامعه‌شناسی

استاد راهنما: رحمت اله صدیق سروسستانی | استاد مشاور: پروانه دانش | پدیدآور: لیلا روستایی

چکیده امروزه در تمام جوامع از جمله ایران ، اولویتهای ارزشی در عرصه ی روابط خانوادگی ، نقش های جنسی و جنسیتی تا حد زیادی دستخوش تغییر شده اند ، این امر اذهان بسیاری از متفکران حوزه جامعه شناسی و سیاستگذاران را متوجه این سوال کرده است که "وضعیت ارزش های خانوادگی در جامعه ایرانی در حال حاضر چگونه است ؟ علت طرح این سؤال وقوع تحولات اقتصادی، اجتماعی و سیاسی است که پیامدهای فرهنگی ارزشی شگرفی را به دنبال داشته است، در سالهای اخیر همواره با دگرگونی های ساختی و اجتماعی چون شهرنش ...

نمایه ها:
 ارزش | اینگلهارت، رونالد | جامعه سنتی | خانواده | نظریه نوسازی

شکل ۲-۱۹: اشتباهات تایپی و نگارشی در چکیده

ب) همگراف^۱

تایپ نادرست واژه‌ها نتایج متفاوتی دارد. به عنوان مثال تایپ نادرست « پژوهش» به صورت‌هایی که به آنها اشاره شد منجر به تولید هیچ واژه معنادار دیگری در فارسی نمی‌شود. تشخیص این خطاهای املائی ممکن است برای برنامه‌های تصحیح خطاهای تایپی و املائی آسان باشد. اما گاهی تایپ نادرست یک واژه مثل «پر» به صورت «بر» منجر به تولید واژه معنادار دیگر یعنی «بر» می‌شود. در این حالت تشخیص اشتباه تایپی برای این برنامه‌ها و حتی گاهی انسان دشوار است.

به عنوان مثال اشتباه تایپی در نوشتن واژه «مقایسه» به صورت «مقیسه» منجر به بازیابی ۵۳ رکورد شد. این تعداد رکورد هم شامل رکوردهای دارای اشتباه تایپی است و هم رکوردهایی است که مربوط به نام جغرافیایی و نیز نام خانوادگی «مقیسه» است.

اشتباه تایپی «مقایسه»	«مقیسه»
نام جغرافیایی	«مقیسه»
نام خانوادگی	«مقیسه»
تعداد کل رکورد	۵۳

رکورد زیر مثال اشتباه تایپی را نشان می‌دهد.

^۱ . Homograph

(0) (0)     

بررسی و مقایسه عملکرد هیبریدهای جدید آفتابگردان (نمیدی)

طرح پژوهشی . موسسه تحقیقات اصلاح و تهیه نهال و بذر . 1369

موضوع: مهندسی کشاورزی < اصلاح نباتات

نقش مشخص نیست: محمدجواد فقیه

دسترسی به فایل تمام‌متن پیشینه‌هایی (رکوردهایی) که نشانه «پی.دی.اف.» ندارند، شدنی نیست.

به منظور بررسی و مقایسه عملکرد هیبریدهای جدید آفتابگردان، 20 هیبرید جدید در قالب طرح بلوکهای کامل تصادفی در 4 تکرار مورد ارزیابی قرار گرفت . نتایج نشان داد که هیبرید CMS15 x R28 با عملکرد 3300 کیلوگرم در هکتار و هیبرید CMS31 x R43 با عملکرد 3231 کیلوگرم در هکتار نسبت به سایر تیمارها برتری داشتند. ...

نمایه‌ها:

آفتابگردان دورگه cms15 x r28 | آفتابگردان دورگه cms 31x r43 | دورگه | آفتابگردان | مقایسه عملکرد |

شکل ۲-۲۰: اشتباهات تایپی و نگارشی در چکیده

مقیسه: نام جغرافیایی

(0) (0)     

زمین شناسی، پترولوژی و ژئوشیمی گنبدهای ساب ولکانیک منطقه مقیسه (جنوب غرب سبزوار)

بایان‌نامه . دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه صنعتی شاهرود - دانشکده علوم زمین . 1393 دریافت فایل  . کارشناسی ارشد

موضوع: علوم و علوم کاربردی < زمین شناسی

استاد راهنما: حبیب‌الله قاسمی | استاد مشاور: محمود صادقیان | پدیدآور: فاطمه فتح آبادی

در منطقه مقیسه (جنوب غرب سبزوار)، گنبد‌های ساب‌ولکانیک مرتبط با فروانش با ترکیب داسیت-ریوداسیت، مجموعه آتشفشانی - رسوبی آتوسن را قطع کرده‌اند. قطعاتی از سنگ میزبان آتوسن و حتی قطعات مارنی الیگومیوسن درون سنگ‌های مورد مطالعه یافت شده است. اصلی‌ترین کانی‌های سازنده این سنگ‌ها شامل فنوکریست‌های پلازیوکلاز با منطقه بندی ترکیبی، آمفیبول با حاشیه سوخته، کوارتز و سانیدین در زمینه‌ای دانه ریز تا شیشه‌ای از همین کانی‌ها است، که بافت‌های متنوع مانند پورفیری، گلومروپورفیری و جریان‌ی را ب ...

نمایه‌ها:

زمین‌شناسی | سنگ‌شناسی | زمین‌شیمی | آداکیت | منطقه جغرافیایی | سنگ‌زایی | زیرآتشفشانی | مقیسه (روستا) | سبزوار (جنوب غرب) |

شکل ۲-۲۱: اشتباهات تایپی و نگارشی و همگراف

مقیسه: نام خانوادگی

اثر مقادیر متفاوت شوری بر فعالیت آنزیم‌های کاتالاز، پراکسیداز و نیترات ردوکتاز در ریشه و برگ‌های ارقام کلزا (Brassica napus)

مقاله‌های محله‌های علمی، پژوهش و سازندگی (در زراعت و باغبانی)، ۱۶، ۵۸، بهار ۱۳۸۲، ۳۹ تا ۴۳

موضوع: پژوهش و توسعه (کشاورزی) < کشاورزی

نویسنده: آرین ساطعی، الهام مقیسه، مه‌لقا قربانلی

دسترسی به فایل تمام‌متن پیشینه‌هایی (رکورد‌هایی) که نشانه «پی.دی.اف.» ندارند، شدنی نیست.

در پژوهش حاضر تاثیر محیط‌های خاک‌ی غیرشور، با شوری کم، متوسط و زیاد (هدایت الکتریکی ۱ و ۷ و ۱۰ و ۱۸ دسی‌زیمنس بر متر) بر فعالیت کاتالاز، پراکسیداز و نیترات ردوکتاز در اندام‌های مختلف کلزا Brassica napus L. رقم‌های Hyola 401 و PF 7045.91 مورد توجه قرار گرفته است. همچنین ارزش مقایسه تغییرات فعالیت این آنزیم‌ها به عنوان معیاری در برآورد تحمل هر یک از این دو رقم به شوری مورد بررسی بوده است. گیاهان دوره کامل رشد و نمو را در کشتهای گل‌دانی طی کرده و پس از انتخاب تصادفی، در ماه چهارم ...

نمایه‌ها:

پراسیکا ناپوس | پراکسیداز | نیترات ردوکتاز | آنزیم | فعالیت | اثر | کلزا | ریشه | برگ | شوری

شکل ۲-۲۲: اشتباهات تایپی و نگارشی و همگراف

مثال دیگر صورت تایپی نادرست کلمه «مسئله» است. اگر در تایپ این کلمه به جای «همزه» حرف «ی» گذاشته شود، صورت نادرست تایپی با کلمه «مسيله» که یک نام جغرافیایی است یکسان خواهد شد. با بازیابی واژه «مسيله» ۳۱۲ رکورد بدست آمد که هم شامل اشتباه تایپی و هم نام جغرافیایی است. در سند نمایه‌شده زیر اشتباه تایپی «مسئله» ملاحظه می‌شود.

رویکردی همه‌جانبه در مورد ارزشیابی بویا: شناسایی بخش‌های دشوار با استفاده از روش گراشی تأمل‌گرایانه

پایان‌نامه، دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه علامه طباطبائی - دانشکده زبانهای خارجی، ۱۳۹۰، کارشناسی ارشد

موضوع: هنر و علوم انسانی < ادبیات انگلیسی

استاد راهنما: فهیمه معرفت | استاد مشاور: شهین واعظی | پدیدآور: اله فضل‌اللهی

الف. موضوع و طرح مسئله (اهمیت موضوع و هدف): ارزشیابی بویا براساس نظریه ی وابستگی و توجه به اهمیت موضوع زی بی دی است، که در آن توجه به پتانسیل افراد در یادگیری بیشتر یا استفاده از کمک خارجی شده است. در واقع ارزشیابی بویا بر پایه ی این نظریه که تدریس و ارزشیابی نباید به صورت جدا در نظر گرفته شوند، تکیه دارد. اغلب دانشجویان در بخش نگارش زبان دچار مشکل اند و ارزشیابی های سنتی تنها به اظهارنظر راجع به درستی یا نادرستی نگارش افراد بدون در نظر گرفتن ابعاد دیگر چون طرز تفکر و علاقه ...

نمایه‌ها:

ارزشیابی بویا | نوشتار | تفکر بازتابگر | زبان انگلیسی | تأمل | نگارش توصیفی | روش‌های تدریس | ارزشیابی | آموزش سنتی

شکل ۲-۲۳: اشتباهات تایپی و نگارشی در چکیده

مسئله: نام جغرافیایی

(0)
 (0)

پایش تغییرات پوشش گیاهی و برخی ویژگی‌های فیزیکی و شیمیایی خاک در منطقه

مسئله قم

پایان‌نامه . دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه صنعتی اصفهان - دانشکده مرتع و آبخیزداری و شیلات و محیط زیست . 1391 . کارشناسی ارشد

موضوع: علوم و علوم کاربردی < کشاورزی، عمومی و میان رشته ای

استاد راهنما: حسین بشری | استاد مشاور: محمدرضا مصدقی . سیدمهدی ادنایی | بدی‌آور: صغری بودینه

سالیان متمادی، بوم‌شناسان سعی داشته اند عواملی را که ترکیب گونه ای وضعیت خاک و تغییرات پوشش گیاهی را کنترل می کنند مشخص سازند. اهمیت مراتع، لزوم ارزیابی های سالانه و میان مدت این اکوسیستم های طبیعی را جهت تعیین روند تغییرات ضروری می نماید. پایین مراتع از اساسی ترین اقدامات مورد نیاز در جهت شناخت و مدیریت مناسب در این عرصه ها می باشد. مناطق دشت **مسئله** و حوض سلطان در استان قم از جمله مناطقی هستند که دستخوش تغییرات طبیعی و مدیریتی زیادی شده اند. پوشش گیاهی و خاک این منطقه طی طرح ...

نمایه ها:

پوشش گیاهی | دریاچه مسئله | خواص شیمیایی-فیزیکی خاک | **پایش** | ناماریکس | قم (استان) | تحلیل چندمتغیره | دریاچه حوض سلطان |

شکل ۲-۲۴: اشتباهات تاپی و نگارشی و همگراف

ج- به هم چسبیدن و اتصال حروف به دلیل حذف فاصله یا نیم فاصله

در بازیابی رکوردهایی دیده می شود که با حذف فاصله یا نیم فاصله، واژه‌ها در عنوان، چکیده و کلیدواژه‌ها به هم چسبیده است.

(0)
 (0)

تحلیلی بر متریک فعالیت توسعه‌دهنده درشناسایی آسیبپذیری نرم‌افزار با استفاده از کاوش مختارت نرم‌افزاری

مقاله‌های همایش‌های ایران . دوازدهمین کنفرانس ملی سیستم‌های هوشمند . انجمن سیستم‌های هوشمند ایران . 1392

نویسنده: جمیله شفیعی . اشکان سامی

شناسایی آسیبپذیری نرم‌افزار، بعد از انتشار، نه تنها موجب تحمیل هزینه‌های اضافی جهت رفع آن میشود، بلکه باعث از بین رفتن اعتبار شرکت نرم‌افزاری نیز میشود. بنابراین هر سازمانی باید شناسایی آسیبپذیری را در اولویت‌های خود قرار داده و از رخداد آن، جلوگیری نماید. تکنیک‌هایی که برای جلوگیری و شناسایی آسیبپذیری نرم‌افزار قبل از انتشار وجود دارند، تکنیک‌های هستند که به پایش چشمی کد می‌پردازند، بنابراین دشوار و زمانبر هستند. علاوه بر این، پایش چشمی همگی قسمت‌های کد نیز کاری پرهزینه است و با توجه ب ...

نمایه ها:

متریک فعالیت توسعه‌دهنده | آسیب پذیری نرم افزار | پیشبینی آسیب پذیری | پیش بینی خطا |

شکل ۲-۲۵: اشتباهات تایپی و نگارشی در چکیده

د- ایجاد فاصله بین حروف یک واژه

از جمله مواردی که سبب بروز اشتباهات تایپی در چکیده یا عنوان می‌شود، درج فاصله یا نیم فاصله به طور نادرست بین حروف یک کلمه است. به عنوان مثال درج فاصله بین حروف در کلمات «ناهم‌خوانی»، «رم‌قیاس- را بطة، می توان-ن-توا نمندیها» سبب بروز خطای تایپی در این کلمات شده است.

ه- خطا در درج علائم سجاوندی و قواعد نگارشی

عدم رعایت فاصله بین علائم نگارشی و کلمات به عنوان مثال فاصله نقطه، پرانتز یا کاما با کلمه قبل و بعد، کاربرد فاصله به جای نیم فاصله، درج خط فاصله به جای نیم فاصله و غیره مواردی هستند از عدم توجه به قواعد نگارشی در هنگام ورود اطلاعات به سامانه هستند. شکل‌های زیر نمونه ای از این موارد را نشان می‌دهند.

تأثیر مشاوره گروهی با تأکید بر حمایت درمانی به روش روابط بین فردی (IPT) بر میزان افسردگی سالمندان مرد

(0) (0) (0) (0) (0)

بایان‌نامه . دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه علامه طباطبائی . 1390 . کارشناسی ارشد

دریافت فایل

موضوع: علوم اجتماعی > روانشناسی تربیتی

استاد راهنما: احمد اعتمادی | استاد مشاور: محمد عسگری | پدیدآور: شهره جعفری

این پژوهش با هدف بررسی تأثیر مشاوره گروهی با تأکید بر حمایت درمانی به روش درمان بین فردی (IPT) بر میزان افسردگی سالمندان مرد انجام شد. نمونه آماری این پژوهش را سالمندان مرد افسرده ای تشکیل دادند که در آزمون مقیاس افسردگی سالمندی (GDS، بیش آزمون) نمره 8 و بالاتر را کسب کرده بودند. نمونه شامل 16 نفر برای ورود به تحقیق مناسب تشخیص داده شده بودند که این 16 نفر به طور تصادفی در گروه های آزمایش (8 نفر) و کنترل (8 نفر) قرار گرفتند. گروه آزمایش در 12 جلسه مشاوره گروهی با رویکرد درمان بین فردی که هر جلسه به مدت 60 دقیقه بود شرکت نمودند. در پایان جلسات پرسش نامه مقیاس افسردگی سالمندان به عنوان پس آزمون اجرا گردید. فرض پژوهش که عبارت بود از مشاوره گروهی با تأکید بر حمایت درمانی به روش درمان بین فردی بر میزان افسردگی سالمندان مرد موثر است، با استفاده از روش تحلیل کو واریانس و با توجه به سطح معنی داری ($P < 0/01$) تأیید می گردد. یعنی کاهش افسردگی آزمودنی های شرکت کننده در گروه آزمایش را می توان به درمان بین فردی نسبت داد.

نمابه ها:

افسردگی (روانشناسی) | گروه درمانی | روان درمانی میان فردی | اتریشی | مشاوره گروهی |

روان درمانی حمایتی | سالخورده | مرد |

شکل ۲-۲۶: اشتباهات تایپی و نگارشی در چکیده

برنامه ریزی توسعه فضایی گردشگری در شهر یزد با رویکرد توسعه پایدار گردشگری

پایان‌نامه . دانشگاه بین المللی امام خمینی (ره) - قزوین - دانشکده معماری و شهرسازی . 1389 . دریافت فایل

کارشناسی ارشد

موضوع: هنر و علوم انسانی < شهرسازی

استاد راهنما: ناصر براتی | استاد مشاور: بهرام امین‌زاده | پدیدآور: مرجان منتظری

برنامه‌ریزی بخش گردشگری به دلیل دارا بودن ابعاد مختلف اقتصادی، اجتماعی، فرهنگی و سیاسی از پیچیده‌ترین برنامه‌ریزی‌ها به شمار می‌رود و توسعه پایدار گردشگری نیازمند شناخت و توجه به تمامی ابعاد این فعالیت گسترده است. به دلیل چند بعدی بودن صنعت گردشگری، در صورت برنامه‌ریزی صحیح و اصولی، این صنعت قادر به ایجاد منافع مستقیم و غیرمستقیم متعدد در زمینه‌های گوناگون می‌باشد. در این میان برنامه‌ریزی راهبردی در توسعه گردشگری با توجه به شرایط عدم ثبات موجود در این بخش و ماهیت این گونه برنامه‌ریزی می‌تواند کاربرد داشته باشد. از آنجا که شهر یزد با برخورداری از ویژگی‌های متنوع طبیعی و تاریخی فراوان و میراث غنی فرهنگی و مصنوع و موقعیت نسبی، فرصت بی‌نظیری برای توسعه گردشگری به حساب می‌آید که تاکنون چندان مورد استفاده قرار نگرفته است؛ لذا رساله حاضر سعی بر آن دارد تا با بهره‌گیری از روش مطالعه موردی، فرصت بهره‌مندی از مزایای متعدد گردشگری در ابعاد مختلف را از طریق برنامه‌ریزی استراتژیک با بهره‌گیری از اصول توسعه پایدار گردشگری در نمونه موردی شهر یزد تبیین نماید و به ارتقاء صنعت توریسم در شهر یزد و تقویت نقش و جایگاه این شهر به عنوان یک مقصد گردشگری در سطح ملی و فراملی دست یابد. در تحقیق حاضر ابتدا به طرح مباحث نظری و پایه‌ای با هدف آشنایی با مبانی، تعاریف و اصول صنعت گردشگری در حوزه‌های مختلف مرتبط با شهرسازی پرداخته شده است. همچنین به منظور آشنایی با شهرهای داخلی و خارجی موفق در زمینه گردشگری به بررسی تجارب آنها در زمینه تدوین و اجرای برنامه‌های گردشگری پرداخته و سعی شده است نقاط قوت تجربیات آنها در برنامه‌ریزی نمونه موردی این رساله به کار گرفته شود. در مرحله بعد مطالعات صورت گرفته در ارتباط با نمونه موردی - شهر یزد- در قالب مطالعات تاریخی، محیطی، اقتصادی، اجتماعی و جمعیتی، فرهنگی و کالبدی مطرح شده‌اند و سپس مطالعات گردشگری با بررسی شاخص‌های مطرح در زمینه گردشگری و منابع و جاذبه‌های گردشگری شهر یزد اعم از جاذبه‌های تاریخی، فرهنگی و طبیعی صورت گرفته است. در نهایت با توجه به این نکته که یکی از اهداف انجام این پژوهش، تعیین نحوه به کارگیری فرایند برنامه‌ریزی راهبردی در برنامه‌ریزی گردشگری در مقیاس شهری است، لذا برنامه‌ریزی راهبردی توسعه گردشگری شهر یزد در چند مرحله تحقیقاتی صورت گرفته و نهایتاً ارائه راهبردهای قابل اجرا با ذکر اولویت و در سطوح بعدی تعیین برنامه‌ها و فعالیت‌های اجرایی از نتایج کاربردی پایانی این رساله می‌باشد. به طور کلی نتایج از این پژوهش نشان می‌دهد می‌توان با اتکاء به نظام برنامه‌ریزی راهبردی با در نظر داشتن ابعاد مختلف صنعت گردشگری، برنامه‌ریزی کالبدی- فضایی مناسب و همچنین با بهره‌گیری از پتانسیل‌های بافت قدیم شهر یزد، بسیاری از نقصان‌های کنونی در پاسخگویی به نیازهای گردشگران را برطرف نمود و فضای توریسم شهری یزد را بهینه کرد.

شکل ۲-۲۷: اشتباهات تایپی و نگارشی در چکیده

و) خطاهای املائی و نگارشی در صورت‌های لاتین کلیدواژه‌ها، چکیده، عنوان

خطاهای املائی و نگارشی به میزان زیادی قابل ملاحظه است.

Reserch (۸) – resarch (۲۰) - reseach (۸) - research (۲)

• درج مطالب نامربوط در چکیده

گاهی کاربر به اشتباه سایر اطلاعات کتابشناختی و یا عنوانهایی مانند "چکیده" و یا کلیدواژه‌ها را در قسمت چکیده وارد می‌کند و بنابراین لازم است در مرحله ساماندهی این مطالب از چکیده حذف شود.

شکل زیر نمونه‌ای از درج مطالب نامربوط را در چکیده نشان می‌دهد.

[illegible]

۲-۴. نتیجه گیری

۲-۴-۱. مرحله تکمیل اطلاعات کتابشناختی و فایلهای ضمیمه

ب- مدار کی است که همچنان ناقص باقی می ماند یعنی اطلاعات کتابشناختی یا فایل های ضمیمه آنها با جستجوی نمایه ساز نیز پیدا نمی شوند و ناقص باقی می مانند. به نظر می رسد این مدارک به همان صورت ناقص وارد پایگاه اطلاعات گنج می شوند. از آنجا که ورود مدارک ناقص از میزان کیفیت بازاریابی و در نهایت از میزان اعتماد کاربران به سامانه گنج می کاهد بهتر است تدبیری اتخاذ شود تا از ورود مدارک مشکوک یا ناقص به سامانه جلوگیری شود.

ت-موردی که به نظر نمی‌رسد در حیطه کار نمایه ساز قرار بگیرد. به عنوان مثال مورد زیر نمونه‌ای از مدرکی است که اطلاعات در فایل ضمیمه نادرست است.



ارزیابی نقش طرح های هادی در توسعه ی کالبدی سکونتگاه های روستایی مطالعه موردی دهستان دشت آب ، شهرستان بافت ، استان کرمان

دریافت فایل

پایان نامه . دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه سیستان و بلوچستان . 1395 . کارشناسی ارشد

موضوع: جغرافیا < برنامه‌ریزی شهری و روستایی

استاد راهنما: جواد بذرافشان | دانشجو: مریم احمدی‌زاده

. چکیده از دو دهه قبل، طرح های هادی روستایی از جمله گسترده ترین طرح های کالبدی روستایی در ایران هستند که عمدتاً جنبه های کالبدی بافت فیزیکی روستاها را مورد توجه قرار می دهند. با ارزیابی اثرات کالبدی این طرح ها می توان با دیدگاه جامع تری امکان برنامه ریزی را در آینده فراهم آورد. مهمترین هدف ارزیابی طرح‌های هادی فراهم ساختن ابزار مدیریت لازم برای نظارت، کنترل، جهت‌دهی و بهبود کارایی آنهاست که بر اساس آن بتوان تجدید حیات و هدایت روستا را به لحاظ ابعاد اجتماعی، اقتصادی و ف ...

نمایه ها:

طرح های هادی روستایی | توسعه کالبدی | دهستان دشتاب | شهرستان بافت | استان کرمان |

شکل ۲-۲۹: عدم انطباق عنوان لاتین با فارسی مدرک

مدرک بالا همراه با یک فایل پی‌دی‌اف است. اما در فایل پی‌دی‌اف مذکور عنوان انگلیسی پایان‌نامه با عنوان فارسی همخوانی ندارد.

پایان‌نامه جهت دریافت درجه کارشناسی ارشد رشته جغرافیا و برنامه‌ریزی روستایی

ارزیابی نقش طرح های هادی در توسعه ی کالبدی سکونتگاه های روستایی

مطالعه موردی دهستان دشت آب ، شهرستان بافت ، استان کرمان

Investigation the effect of mass media in promoting cultural development on rural women(Case Study: Kyare gharbi Village, Chaharmahal and Bakhtiari Province)

استاد راهنما:

دکتر جواد بذرافشان

نگارش:

تشخیص این موارد بسیار وقت گیر است و در حیطه کار نمایه‌ساز قرار نمی‌گیرد.

ث- ترجمه به فارسی یا انگلیسی اطلاعات کتابشناختی نمونه دیگری از مسائل زبان‌شناختی در مرحله تکمیل اطلاعات است. ترجمه از یک زبان به زبان دیگر مستلزم دانش زبانی در زمینه ساختارهای زبانی هر یک از زبان‌های مبدأ و مقصد است. ترجمه اشتباه یا غیردقیق سبب ایجاد خطا در ساماندهی و در نتیجه کاهش دقت بازیابی می‌شود.

۲-۴-۲. مرحله نمایه‌سازی

الف- استفاده از معادل‌های متعدد برای یک واژه لاتین: به نظر می‌رسد ساماندهی این موارد باید به گونه‌ای صورت گیرد که معادل‌های متعدد در کنار هم آورده شود. بدین صورت که معادل‌های متعددی که برای یک واژه بکار می‌رود، به نحوی در اختیار کاربر قرار گیرد و به او پیشنهاد شود. در این صورت بازیابی آنها نیز بهتر صورت می‌گیرد.

ب- نوشتن صورت‌های مخفف لاتین به اشکال مختلف: مثلاً با علائم نگارشی نظیر نقطه، خط فاصله و رقم (ا.چ. وی- سی ۲ یا ا.چ. وی سی نوع ۲ و غیره). این تفاوت‌های نگارشی رکوردهای بازیابی متفاوتی را بدست می‌دهد.

ت- تعیین اختصار فارسی برای یک واژه لاتین: به عنوان مثال اصطلاح voluntary counselling and testing دارای صورت اختصاری (VCT) است اما صورت مختصر فارسی آن (ماد) برای معادل «مشاوره و آموزش اختیاری» تعیین شده است. بازیابی با این صورت‌های مختلف منجر به تولید رکوردهای مختلف می‌شود. لازم است در بازیابی همه این صورت‌ها با هم دیده شود.

ث- تعدد صورت‌های نوشتاری اسامی خاص افراد (پدیدآوران): اول نام پدیدآوران ممکن است به چند صورت ملاحظه شود مثلاً شمس لیالستانی یا فیض شمس لیالستانی. دوم اینکه یک نام ممکن است به صورت‌های مختلفی نوشته شود. (یدالله یا یداله یدا...) به دلیل صورت‌های مختلف نوشتاری، وجود فاصله یا نیم فاصله و غیره و سوم این که نام افراد (مانند نام سازمانها) ممکن است در طی زمان تغییر کند و یا صورت‌های دیگری مثل اسم مستعار بکار رود که در این صورت لازم است همه صورت‌ها در بازیابی مشاهده شود.

ج- انواع خطاهای املایی و نگارشی در کلیدواژه‌ها، چکیده و پدیدآوران شخص در هر دو سطح فارسی و لاتین

ح- درج مطالب نامربوط در چکیده: به عنوان مثال درج کلمه چکیده یا گذاشتن کلیدواژه‌ها در چکیده

شناسایی مسائل و ارائه راهکارهای زبانشناختی سازمان‌دهی و بازیابی اطلاعات در گنج □ ۳۸

خ- یافتن کلیدواژه: الف) تعیین معادل فارسی کلیدواژه‌های لاتین و ب) تعیین صورت‌های اختصاری کلیدواژه‌ها

د- خطاهای درج علائم ریاضی و فرمولهای شیمی در چکیده و عنوان

فصل سوم

مسائل زبانشناختی بازیابی اطلاعات

۳. مقدمه

نتایج گزارش‌های متعدد تدوین شده از سوی مرکز اطلاعات و مدارک علمی ایران به عنوان مثال «گزارش بررسی پایگاه اطلاعات گنج و ارائه راه‌حل‌ها» و نیز گزارش دفتر تحقیق و توسعه (یکپارچه‌سازی کاراکترها و شناسایی اشکال‌های فراداده مدارک ویرایش شده در نسخه تازه پایگاه اطلاعات علمی ایران گنج، اسفند ۹۵) نشان می‌دهد که بازیابی اطلاعات در سامانه گنج از کیفیت و دقت لازم برخوردار نیست. به نظر می‌رسد بخشی از مشکلات ریشه در ساختار زبانی متن سامان‌دهی شده دارد و برخی دیگر ناشی از مسائل فنی مربوط به سامان‌دهی و بازیابی اطلاعات است.

کیفیت و دقت بازیابی در یک سامانه با معیارهای متعددی اندازه‌گیری می‌شود. اثربخشی^۲ و کارایی^۳ دو پارامتر اصلی برای ارزیابی عملکرد سیستم‌ها محسوب می‌شوند. این پارامترها با معیارهای متعددی اندازه‌گیری می‌شوند. لنکستر (۱۹۷۹) پنج معیار برای ارزیابی سیستم‌های بازیابی اطلاعات ارائه می‌کند:

۱- میزان پوشش سیستم،

۲- توانایی سیستم برای بازیابی اطلاعات خواسته (بازخوانی)^۴،

۳- توانایی سیستم برای پرهیز از بازیابی اطلاعات ناخواسته (دقت)^۵،

۴- مدت‌زمان پاسخ‌دهی سیستم

۵- میزان تلاش کاربر برای دریافت اطلاعات.

براساس معیارهای گفته شد، می‌توان دقت بازیابی را سنجید. مسائلی که در این طرح به آنها پرداخته می‌شود، بنابراین محدود به مسائل زبانی و زبانشناختی متن است.

۳-۱. بررسی لاگ‌های کاربران

یکی از معیارهایی که لنکستر (۱۹۷۹) به آن اشاره می‌کند، میزان توانایی سیستم برای بازیابی اطلاعات خواسته یا بازخوانی است. به لحاظ زبانشناختی این سوال مطرح می‌شود که ساختار زبانی اطلاعات خواسته چگونه است. به عبارتی جستجوی کاربر در قالب چه ساختارهای زبانی صورت می‌گیرد. برای پاسخ به این سوال لاگ‌های جستجوی کاربران در یک بازه سه ماهه در پایگاه گنج برای سال ۱۳۹۶

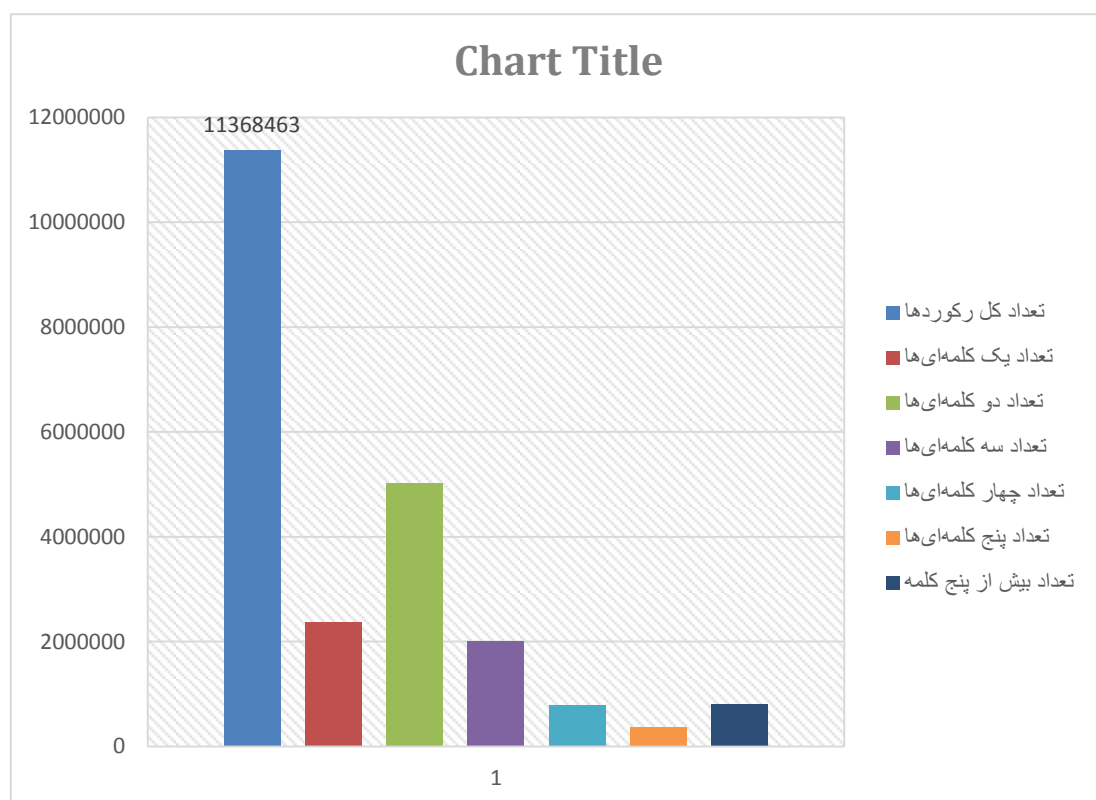
^۲ . effectiveness

^۳ . efficiency

^۴ . recall

^۵ . precision

بررسی شد. این لاگها توسط دفتر تحقیق و توسعه تهیه و برای بررسی مورد استفاده قرار گرفت. لاگ‌های کاربران در بازه سه ماهه در پایگاه گنج نشان می‌دهد که بیشتر میزان بازیابی براساس زنجیره‌ای از دو کلمه صورت گرفته است. پس از دو کلمه‌ای‌ها بیشتر میزان بازیابی به ترتیب با یک کلمه و سپس سه کلمه‌ای صورت گرفته است (نمودار زیر).



نمودار ۳-۱: ساختار زبانی عبارت جستجو

بر این اساس نقش کلیدواژه‌های دو یا یک کلمه‌ای در بازیابی روشن می‌شود. بنابراین سیستم باید به گونه‌ای باشد که بتواند فیلدهای متفاوت براساس یک کلمه یا دو کلمه را به صورت دسته‌بندی در اختیار کاربر قرار دهد به نحوی که کاربر در ادامه بتواند به جستجوی مورد نظر برسد.

بررسی لاگ‌های کاربران به لحاظ ساخت زبانی روشن می‌کند که چه ساخت‌های زبانی بیشترین کاربرد را دارد. بررسی هزار نمونه از این لاگها نشان می‌دهد که:

- ۱- بیشتر جستجوها براساس ساخت زبانی اسم + اسم است.
- ۲- تنها یک مورد جستجو بر اساس نام پدیدآور در هزار مورد قابل مشاهده است.
- ۳- در حدود ۵۰ مورد جستجوی لاتین در بین ۱۰۰۰ مورد قابل مشاهده است.

این نتایج نشان می‌دهد که جستجو براساس نام پدیدآور به ندرت صورت می‌گیرد. و دیگر این که از آنجا که بازیابی به لاتین هم صورت می‌گیرد شناسایی مسائل مدارکی که به زبان لاتین هستند، نیز ضروری است.

۲-۳. مسائل زبان‌شناختی بازیابی

در فصل پیش مجموعه‌ای از مسائلی که در مرحله نمایه‌سازی متون فارسی به لحاظ زبان‌شناختی قابل مشاهده است، ارائه شد. در این بخش به مسائل زبانی توجه می‌کنیم که حین بازیابی رکوردها دیده می‌شود.

۱-۲-۳. اشتباه و اشکال در صورت زبانی متن

با جستجوی یک کلیدواژه در پایگاه متن، مجموعه‌ای از رکوردها در اختیار کاربر قرار می‌گیرد. با یک نگاه اولیه در متن این رکوردها اشتباهات و اشکالات متعددی قابل تشخیص است یعنی به عبارت دیگر رکوردی که در اختیار کاربر قرار می‌گیرد تمیز نیست و اشتباهات و اشکالات متعدد در آن قابل مشاهده است هرچند در رکوردهای نمایه‌سازی شده این اشتباهات و خطاها بسیار کمتر است. این اشتباهات در صورت زبانی متن وجود دارند یعنی اشتباهات و اشکالاتی که به لحاظ دستور زبان فارسی و ساختار زبان فارسی خطا محسوب می‌شود. این خطاها همان مسائلی هستند که لازم است در مرحله ساماندهی و نمایه‌سازی اصلاح شوند و در فصل قبل مثالهایی از آنها آورده شد. عدم اصلاح این موارد سبب می‌شود تا متن سند با مجموعه‌ای از خطاهای زبانی در اختیار کاربر قرار گیرد. این موارد به طور خلاصه شامل موارد زیر است:

- ۱- کاربرد اشتباه علائم نگارشی
- ۲- جملات تمام نشده
- ۳- فرمول‌های نادرست ریاضی، شیمی، علائم نادرست و غیره
- ۴- اشتباهات تایپی و املایی: به عنوان مثال با بازیابی صورت املایی نادرستی مانند «شناسایی» ۱۹ رکورد (صورت درست: شناسایی) بدست می‌آید.
- ۵- درج فاصله به جای نیم‌فاصله (در چکیده یا عنوان) و یا پراوتز

۳-۲-۲. صورت‌های متفاوت نوشتاری

وقتی کلمات در متن (چکیده یا عنوان و غیره) با صورتهای متفاوت نوشتاری ذخیره می‌شوند. در بازیابی همه رکوردها یکسان بدست نمی‌آید. یعنی به عنوان مثال تعداد رکوردهایی که با بازیابی صورت نوشتاری «مؤثر» بدست می‌آید در حدود ۱۵۲۳ مورد است درحالی که با بازیابی «موثر» تعداد رکوردها به ۶۱۹۴۷ مورد می‌رسد. این مسائل که بر دقت بازیابی تأثیر می‌گذارند، ناشی از صورتهای متفاوت نوشتاری هستند. نکته ای که به لحاظ زبانی در اینجا مطرح می‌شود این است که این صورتهای نوشتاری متفاوت دو نوع هستند: ۱- صورتهایی که در نوشتار فارسی مجاز محسوب می‌شوند و مرسوم هستند، مانند صورت «جزیی» با ۳۰۴۶ رکورد و «جزئی» با ۸۰۲۴ رکورد و ۲- صورتهایی که اشتباه املائی محسوب می‌شوند. در مورد صورتهایی که اشتباه املائی محسوب می‌شوند لازم است در مرحله نمایه سازی یا پیش از آن اصلاح شوند. اما در مورد صورتهایی که هر دو صورت در زبان رایج است لازم است بازیابی به گونه ای صورت گیرد که شامل هر دو مورد تفاوت نوشتاری شود. این صورتهای مختلف نوشتاری را می‌توان به دسته های زیر تقسیم کرد

الف- کاربرد علامتهای مختلف نوشتاری: تشدید، مد، همزه، تنوین، آوا، نقطه، خط تیره

مثال‌های زیر نمونه ای از تفاوت در رکورد بازیابی بر اساس تفاوت در صورت نوشتاری را نشان می‌دهد. از آنجا که درج صورتهای دارای علائمی مثل همزه، تشدید، مد و غیره نیازمند کاربرد همزمان دو کلید بر روی کیبورد است، معمولاً این علائم کمتر درج می‌شوند و همانطور که مثالهای زیر نشان می‌دهد معمولاً صورت ساده تر در بازیابی بیشتر مشاهده می‌شود.

نکته قابل ذکر در اینجا این است که صورتهای نادرست لازم است در مرحله نمایه سازی مشخص و اصلاح شوند. از آنجا که بازیابی بیشتر براساس دو یا یک واژه است بنابراین وجود این تفاوتها بر بازیابی تأثیر بسزایی دارد.

جدول ۳-۱: صورت‌های متعدد نوشتاری در بازیابی

واژه (تعداد رکورد بازیابی)	واژه (تعداد رکورد بازیابی)	واژه (تعداد رکورد بازیابی)
۲۸۶ باکتریائی ۳۲۸۷- باکتریائی	۲۴۰۶- مسؤل ۲۲۵- مسؤل	۲۷۹۱۴- فرآیند ۲۵۴۹۰- فرآیند
۸۰۲۴- جزئی ۳۰۴۶- جزئی	۲۹۶۷۲- مسائل ۵۷۶۷- مسائل	۴۲۲۸- فرآورده ۲۳۹۲- فرآورده
۲۲۶۵- غذائی ۲۰۱۰۷- غذائی	۷۱۱۲- مسأله ۱۳۰۶۶- مسأله	۱۹۳۴۴- برآورد ۶۳۹- برآورد

	مسئله ۲۴۸۹۷-	
برایند ۱۰۱- برآیند ۷۲۹	سوال ۱۰۹۲۰- سؤال ۳۰۴-	اجرائی ۸۹۰۱- اجرائی ۱۴۴۱
	آرا ۱۴۰۲- آراء ۳۰۹۳	جدایی ۲۳۳۴- جدائی ۱۶۶
		علایم ۲۵۸۶- علائم ۶۹۱۰

ب- صورت‌های متفاوت کسره اضافه

کسره اضافه به هر دو صورت همزه یا «ی» در ترکیبات اسمی مشاهده می‌شود. از آنجا که بازیابی بیشتر به صورت دو کلمه ای و با ساخت اسم + اسم است بنابراین نقش کسره اضافه در میزان دقت بازیابی اهمیت دارد. به طور کلی به نظر نمی‌رسد کاربران برای بازیابی صورتهایی که دارای ساخت اسم + اسم است، کسره اضافه را درج کنند. اما در بازیابی درج کسره اضافه به صورت های متفاوت منجر به بازیابی متفاوت می‌شود، به عنوان مثال بازیابی «سرمایه ی اجتماعی» ۲۲۶۳ رکورد را نشان می‌دهد. حال آنکه جستجوی «سرمایه اجتماعی» ۹۲۲۱ رکورد و جستجوی «سرمایه اجتماعی» با کیورد نگارنده تنها ۲ رکورد را نشان می‌دهد. شکل زیر را نشان می‌دهد که درج کسره اضافه با «ی» به صورت های متفاوت با نیمه فاصله یا فاصله کامل صورت گرفته است.

تحلیل بر ویژگی‌های روان‌سنجی پرسشنامه تیب شخصیتی A و B بورتیر براساس نظریه (0) (0) سوال - پاسخ و نظریه کلاسیک اندازه‌گیری

پایان‌نامه . دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه علامه طباطبائی - دانشکده علوم تربیتی و روانشناسی . 1388 . کارشناسی ارشد

موضوع: علوم و علوم کاربردی < روانشناسی

استاد راهنما: علی دلاور | استاد مشاور: غلامرضا صرامی | پدیدآور: مهین شمسی

به منظور بررسی خصوصیات روان‌سنجی پرسشنامه ی تیب شخصیتی B و A بورتیر، در چارچوب نظریه ی سوال- پاسخ، سوال های این پرسشنامه با استفاده از یک نمونه ی 200 نفری از کارمندان آموزش و پرورش مورد تجزیه و تحلیل قرار گرفت. با توجه به نتایج حاصل از بررسی فرض‌های نظریه ی سوال - پاسخ و قابل قبول بودن نسبی آن ها با روش‌ها و ابزارهای موجود، در ادامه پارامترهای سوال ها با استفاده از مدل مقیاس رتبه‌ای که به وسیله ی نرم افزار PARSCALE4 قابل اجرا است، گزارش شده اند. این مدل که از بسط مدل پاسخ مدرج سمی جیما به دست آمده برای سوال‌های طیف لیکرت یک پارامتر جایگاه و یک شیب ارایه داده و برای هر یک از مقوله‌های پاسخ نیز یک پارامتر ارایه می‌دهد سوال اول این پژوهش این است که "آیا بین نظریه های کلاسیک اندازه گیری و سوال- پاسخ از نظر دقت برآورد تفاوت وجود دارد؟" سوال دوم این پژوهش نیز این بوده است که "آیا بین نظریه های کلاسیک اندازه گیری و سوال- پاسخ از لحاظ تعداد سوالات مناسب در سوالات پرسشنامه ی بورتیر تفاوت وجود دارد؟" پس از انجام تحلیل های کلاسیک و بررسی فرض های زیر بنایی IRT، با استفاده از نرم افزار PARSCALE4 به برآورد پارامترهای سوال پرداخته شد. نتایج نشان داد که در کلیه ی سوال ها، فرض های زیر بنایی IRT برقرار و هیچ کدام از سوال ها، فرض های مورد نظر را نقض نکردند. تجزیه و تحلیل سوال های پژوهش نیز نشان داد که نظریه ی سوال- پاسخ در سطوح مختلف توانایی، خطاهای مختلف را برآورد می‌کند، اما نظریه ی کلاسیک برای تمام سطوح صفت مورد اندازه‌گیری تنها یک برآورد از خطای استاندارد اندازه‌گیری ارائه می‌دهد. از آنجایی که تابع آگاهی هم در مورد سوال‌ها و هم در مورد آزمون پارامتر آگاهی فراهم می‌کند، پس با نظریه ی سوال- پاسخ بهتر می‌توان قضاوت کرد، اما نظریه ی کلاسیک فاقد این مزیت می‌باشد، بنابراین می‌توان نتیجه گرفت که دقت برآورد و کارآمدی نظریه ی سوال- پاسخ بیشتر از نظریه ی کلاسیک است. بررسی سوال دوم این پژوهش نیز نشان داد که به جز دو سوال 1 و 6 کلیه ی سوال‌ها همبستگی نسبتاً بالایی با آزمون داشتند. در نظریه ی سوال- پاسخ نیز سوال های 1، 2، 3 و 6 دارای شیب بسیار پایینی هستند. با این حال سایر سوال ها نیز از شیب مطلوبی برخوردار نیستند.

شکل ۳-۱: صورتهای متعدد نوشتاری کسره اضافه

شکل زیر نمونه ای از اشتباهات تایپی ناشی از درج کسره اضافه با «ی» را نشان می دهد.

بررسی هنرشناختی آثار معماری و نقاشیهای دیواری منطقه جلفای اصفهان (مکتب اصفهان و دوره ی شاه عباس صفوی)

پایان‌نامه . دولتی - دانشگاه پیام نور (وزارت علوم، تحقیقات، و فناوری) - دانشگاه پیام نور استان تهران - مرکز پیام نور تهران - دانشکده هنر و معماری . 1393 . کارشناسی ارشد

موضوع: هنر و علوم انسانی < پژوهش در هنر

استاد راهنما: آناهیتا مقبلی | استاد مشاور: رحمت امینی | پدیدآور: سایه آرین

پژوهش حاضر به منظور بررسی هنرشناختی آثار معماری و نقاشیهای دیواری منطقه ی جلفای اصفهان (مکتب اصفهان و دوره ی شاه عباس صفوی) انجام شد. بدین منظور از روش توصیفی - تحلیلی بهره گرفته شد. این تحقیق از نوع کیفی و کاربردی است. اطلاعات ضروری تحقیق با استفاده از روشهای کتابخانه‌ای (فیشبرداری) (میدانی) (گفتگو با کارشناسان) (گردآوری شد). مهمترین یافته‌ها و نتایجی که به دست آمد، عبارتند از این که مکتب اصفهان و هنر معماری و نقاشی منطقه ی جلفای اصفهان و کلیسای وانک بر یکدیگر تأثیرگذار بوده‌اند. افزون بر این، هنرمندان مکتب اصفهان و منطقه ی جلفا در دوران شاه عباس صفوی با هنر و هنرمندانی از اروپای آن روزگار در ارتباط بوده‌اند و با یکدیگر داد و ستد فرهنگی - هنری داشته‌اند. نتایج حاکی از آن است که بررسی دقیق آثار هنری متعلق به مکتب اصفهان و منطقه ی جلفا و چگونگی آفرینش این آثار، میتواند ایده‌پرداز آفرینش آثار هنری ارزشمند و ماندگار در دوران اکنون و نیز تعامل و همافزایی هنرمندان روزگار اکنون ایران با هنر روز دنیا باشد.

نمایه ها:

نقاشی دیواری | معماری | مکتب اصفهان (نقاشی) | اصفویان | عباس صفوی اول، شاه ایران | انگارگری |

شکل ۳-۲: اشتباهات تایپی در چکیده

ت-صورت‌های مختلف جمع

صورت‌های مختلف جمع «ها، ان، جمع مکسر عربی و غیره» منجر به بازیابی میزان رکوردهای متفاوت می‌شود. به عنوان مثال بازیابی «بنگاه اقتصادی» ۱۷۴۴ رکورد اما «بنگاه‌های اقتصادی» ۱۶۹۳ رکورد را نشان می‌دهد. بازیابی «حسابدار رسمی» ۱۷ رکورد و «حسابداران رسمی» ۱۱۳ رکورد و «عامل اقتصادی» ۶۱۸۶ و «عوامل اقتصادی» ۱۵۸۶۲ رکورد را نشان می‌دهد. بازیابی «سؤال» ۳۹۳ رکورد، «سؤالها» ۱ رکورد، «سؤالهای» ۱۰ رکورد و «سؤالات» ۲۲۶ رکورد را نشان می‌دهد. بازیابی «میوه‌ها» ۴۱۰۵ رکورد و «میوه جات» ۱۶۷ رکورد را نشان می‌دهد.

جستجوی گروه اسمی «استاد دانشگاه» ۱۲۱۷ رکورد، «اساتید دانشگاه» ۲۳۱۴ رکورد و «استادان دانشگاه» ۵۲۲ رکورد را نشان می‌دهد.

ث-صورت‌های مختلف اختصارات

بازیابی اختصارات یکی از مواردی که به نظر می‌رسد از دقت بسیار کمی برخوردار است. به عنوان مثال جستجوی «وی پی ان» منجر به یافتن ۹۲ رکورد شد. با نگاهی به رکوردهای بدست آمده در صفحه اول می‌توان دید که هیچ یک از رکوردها راجع به «وی پی ان» نمی‌باشند بلکه رکوردهایی هستند که در آنها کلمه‌هایی مثال «پی در پی»، «وی (او)»، «پی وی سی» و غیره بکار رفته است.

با بازیابی رکوردهای «ام‌اس-ام‌اس-اس‌ام‌اس و اس‌ام‌اس» تعداد یکسانی رکورد ۱۳۷۹ مورد بدست می‌آید. نکته قابل توجه این است که با بازیابی «اس‌ام‌اس» در صفحه اول تنها یک رکورد مربوط به «اس‌ام‌اس» است و سایر رکوردها بیشتر شامل «ام‌اس» است.

بازیابی «دی ان‌ای» و «دی.ان‌ای» ۱۵۷۴ رکورد اما «دی ان آ» ۱۵۸۴ رکورد را نشان می‌دهد.

با جستجوی «ای آر پی» و «ای.آر.پی» ۱۰۴۷ رکورد بدست می‌آید که هیچکدام در صفحه اول مربوط به این کلمه اختصاری نیستند بلکه قسمتی از این کلمه اختصاری در آنها بکار رفته است.

۳-۲-۳. مسائل معنایی-صرفی-نحوی

الف- کلمات چندمعنا با املای یکسان (اما صورت آوایی یکسان یا متفاوت)

از آنجا که نتایج نشان داد جستجوی کاربران بیشتر به صورت دوکلمه‌ای است، بنابراین بررسی تأثیر کلمات چند معنا بر میزان دقت بازیابی را در ترکیبات بررسی می‌کنیم. به عنوان مثال با جستجوی عبارت «دستگاه شور» ۲۰۲ رکورد را نشان می‌دهد که شامل رکوردهای مربوط به «آب شور، خاک شور، اراضی شور» و غیره است. جستجوی «آب شور» ۲۲۱۳ رکورد و «خاک شور» ۱۶۰۱ رکورد را نشان می‌دهد.

جستجوی عبارت‌هایی مانند «قرص مسکن، داروی مسکن یا مسکن‌های دارویی» هیچ تداخلی با کلمه «مسکن یا خانه» نشان نداد. اما رکوردهای بدست آمده از جستجوی «مسکن‌های قوی» تقریباً همه رکوردهای مربوط به «ساختمان و مسکن» هستند. جستجوی «مسکن‌های درمانی» شامل هم رکوردهای «ساختمان و مسکن» و هم رکوردهای مربوط به «مسکن در حوزه دارویی» است. جالب توجه است که در صورت متصل کردن «های جمع» به کلمه «مسکن» به صورت «مسکنهای دارویی و یا مسکنهای درمانی» بازیابی هیچ رکوردی را نشان نمی‌دهد و بازیابی به صورت «جستجو نتیجه‌ای در بر نداشت» دیده می‌شود.

ب- صورت‌های متفاوت واژگانی با معنای یکسان یا مرتبط

صورت‌های متفاوت واژگانی که معنای یکسان یا مرتبط به هم دارند حجم وسیعی به ویژه در حوزه اصطلاحات علمی را شامل می‌شود. به عنوان مثال چهار اصلاح «متانول - متیل الکل - الکل متیلیک - الکل چوب» دارای معنای یکسان هستند. بنابراین انتظار می‌رود در بازیابی رکوردهای هر چهار اصطلاح دیده شود. جستجوی این ترکیبات به ترتیب ۲۳۲۲، ۴۱۵، ۱۸ و ۴۷ رکورد را نشان می‌دهد.

بررسی صورت‌های متفاوت «دریای خزر» ۲۵۶۹ رکورد، «دریای کاسپین» ۳۷ رکورد، «دریای مازندران» ۴۸۱ رکورد و «بحر خزر» ۲۶ رکورد را نشان می‌دهد.

مورد دیگر معادلهای متفاوت اصطلاحات علمی است. به عنوان مثال برای اصطلاح lubrication معادلهای متفاوتی مانند روغن کاری، نرم کاری، روانسازی و لیزاندن بکار رفته و به عنوان کلیدواژه ذخیره شده است. لازم است که برای افزایش دقت بازیابی همه این کلیدواژه‌ها دیده شود. مثال‌های دیگر مانند معادلهای «صوت‌شیمی، آواشیمی، سونوشیمی» که برای اصطلاح soundchemistry بکاررفته است.

ت- صورت‌های کوتاه / کامل اسامی یا سازمان‌ها

جستجو برحسب فیلد پدیدآور نیز بر سرعت و دقت بازیابی تأثیر بسزایی دارد. جستجو برحسب نام پدیدآور «محمد احمدی» در پایگاه گنج بتا ۵۶۲ رکورد را نشان می‌دهد که در صفحه اول تنها یک مورد ترکیب مذکور را نشان می‌دهد و بقیه موارد یا برحسب نام اول «محمد» یا نام دوم «احمدی» منحصر است. در مورد اسامی چند قسمتی مشکلات دیگری نظیر متصل یا جدا بودن اسمها یا فاصله و نیم فاصله و سایر مواردی که در فصل قبل به آنها اشاره شد نیز وجود دارد. به عنوان مثال جستجوی «میرفیض فلاح شمس لیالستانی» ۵۵ رکورد را نشان می‌دهد اما «فیض فلاح شمس لیالستانی» هیچ رکوردی را نشان نمی‌دهد. در مورد ترکیب «یدالله مهرعلی زاده» تنها ۱ رکورد در گنج دیده می‌شود اما در مورد «یدالله مهرعلیزاده» ۳ رکورد بدست می‌آید. در مورد «شمس الله شیرین بخش ماسوله» تعداد رکوردها ۶ مورد، «شمس الله شیرین بخش» ۱۲ رکورد و «شیرین بخش ماسوله» ۱۰ رکورد را نشان می‌دهد.

۳-۲-۴. مسائل فنی

علاوه بر مسائل زبانی که در بالا ذکر شد، برخی مسائل فنی نیز قابل ذکر است:

الف- اشکالات تایپی (کارکترهای نوشتاری نادرست یا بهم چسبیدن کلمات) در چکیده و غیره ناشی از کدهای نوشتاری متفاوت (یونیکد نیست)

چمدون قصار، ابوحنص حداد و جایگاه آنها در مکتب ملائمه

پایان‌نامه . دولتی - وزارت علوم، تحقیقات، و فناوری - دانشگاه سمنان ، 1390 . کارشناسی ارشد

موضوع: هنر و علوم انسانی < ادبیات فارسی

استاد راهنما: ناصر رحیمی | **استاد مشاور:** عصمت اسماعیلی | **پدیدآور:** رحیمه ادهم

دریافت فایل

ابوحنص حداد و چمدون قصار دو تن از بنیان‌گذاران و رواج‌دهندگان **طریق**؟ ملامت در نیشابور قرن سوم بوده‌اند. چون ملائمه اعتقادی به نوشتن نداشته‌اند، اقوال و حکایت‌های اندکی از این دو تن به جا مانده که دیگران از آنها نقل کرده‌اند. با بررسی آنچه در منابع کهن و کتاب‌های عرفانی و تاریخی و ادبی از این دو پیر ملامت آمده، آگاهی‌هایی **دربار**؟ زندگی و **خانواد**؟ آنها به دست داده‌ایم. با تأمل در سخنان و حکایت‌هایی که از آنان نقل شده برخی از زوایای **اندیش**؟ آنها را نیز روشن ساخته‌ایم که از طرفی بیانگر خطوط اصلی تفکر ایشان و از سوی دیگر نشان **دهند**؟ اصول ملائمه است. همچنین توجه آنها به فتوت و رابطه‌ای که با اهل فتوت داشته‌اند تا حدودی تأثیر گذاری‌ها و تأثیر پذیری‌های این دو جریان را آشکار می‌سازد. از آنجا که تصوف در ادبیات فارسی تأثیری قوی نهاده و اقوال و احوال و سلوک و **اندیش**؟ عارفان بر ذهن و زبان بسیاری از شاعران اثر گذاشته، این تحقیق چند نمونه از بازتاب‌های سخنان و حکایت‌های چمدون قصار را در شعر شاعران بزرگ و مشهور، به ویژه در **حور**؟ ادب عرفانی و تعلیمی نشان می‌دهد و برای یافتن نمونه‌هایی بیشتر از تأثیر تفکر ملائمتی و اندیشه و اقوال این دو شیخ بر شعر شاعران این حوزه سودمند خواهد بود.

نمایه‌ها:

ادبیات تعلیمی | ادبیات عرفانی | تصوف | جوانمردی | حداد، ابوحنص | شعر | قصار، چمدون | ملائمه

بررسی عوامل مؤثر بر تناسب الزامات سیستمهای اطلاعاتی حسابداری با ظرفیت آنها

پایان‌نامه . غیر دولتی - دانشگاه آزاد اسلامی - دانشگاه آزاد اسلامی واحد بزد - دانشکده اقتصاد و حسابداری

1391 . کارشناسی ارشد

موضوع: علوم اجتماعی < اقتصاد

استاد راهنما: محمود معین‌الدین | **استاد مشاور:** شهناز نایب‌زاده | **پدیدآور:** زهرا معینی

دسترسی به فایل تمام‌متن پیشینه‌هایی (رکورددهایی) که نشانه «پی.دی.اف.» ندارند، شذنی نیست.

یکی از وظایف اساسی حسابداران، تأمین اطلاعات صحیح و به‌موقع است تا مدیران تصمیم‌های خود را بر اساس این اطلاعات اتخاذ نمایند. در راه نیل به این اهداف، سیستمهای اطلاعاتی حسابداری رایانه ای، بعنوان منبع اولیه، با ارائه اطلاعات بموقع و مربوط به سازمانها در مدیریت تجارتشان کمک می‌کند. بدین منظور پژوهش حاضر به بررسی عوامل مؤثر بر تناسب الزامات سیستمهای اطلاعاتی حسابداری و ظرفیت آنها می‌پردازد. مدل تحقیق، با استفاده از متغیرهای خبرگی فنی، خبرگی اطلاعاتی، دانش مدیر، تعهد مدیر، حضور متخصصان داخلی و حضور متخصصان خارجی به عنوان متغیر مستقل، و عدم تناسب الزامات سیستمهای اطلاعاتی حسابداری و ظرفیت آنها به عنوان متغیر وابسته، شکل گرفته است. روش تحقیق حاضر، از نظر هدف، کاربردی بوده و بر حسب روش تحقیقی توصیفی-پیمایشی است که با بکارگیری ابزار پرسشنامه و با روش پیمایشی، داده‌های مورد نیاز جمع‌آوری شده است. جامعه آماری این پژوهش، مدیران مالی شرکتهای تولیدی استان کرمان می‌باشد و از روش نمونه‌گیری به روش گرجسی و مورگان استفاده گردیده و تعداد 117 نمونه قابل قبول جمع‌آوری گردید. تجزیه و تحلیل داده‌ها به روش رگرسیون چند متغیره و با استفاده از نرم افزارهای Excel , spss انجام شده است. نتایج پژوهش نشان می‌دهد که بین الزامات سیستمهای اطلاعاتی حسابداری و ظرفیت آنها رابطه معناداری وجود دارد و از طرف دیگر، در شرکت های تولیدی مورد بررسی، بین الزامات سیستمهای اطلاعاتی حسابداری و ظرفیت آنها عدم تناسب وجود دارد که این منجر به کاهش سطح اهمیت و وسعت استفاده از AIS در این شرکت ها و نهایتاً منجر به ضعف عملکرد آنها می‌شود. مطالعه عوامل شش گانه مؤثر بر تناسب الزامات سیستم های اطلاعات حسابداری و ظرفیت آنها نشان دهنده این است که از بین این عوامل: "خبرگی فنی" و "حضور متخصصان داخلی" با عدم تناسب بین الزامات سیستمهای اطلاعاتی حسابداری و ظرفیت آنها رابطه معنادار و

شکل ۳-۳: اشتباهات ناشی از تفاوت در کدهای نوشتاری

ب- بازیابی رکوردهای دارای اطلاعات ناقص در همه موارد

طرح تحقیقاتی یافتن ضرایب ناشی از تقسیم دو سری توانی با استفاده از یک درمیان

طرح پژوهشی . 1390

موضوع: علوم پایه <

مجری: محمد غلامی بابدگانی

ناقص

دسترسی به فایل تمام متن پیشینه‌هایی (رکوردهایی) که نشانه «پی.دی.اف.» ندارند، شدنی نیست.

طرح تحقیقاتی یافتن ضرایب ناشی از تقسیم دو سری توانی با استفاده از یک درمیان ...

نمایه ها:

طرح تحقیقاتی یافتن ضرایب ناشی از تقسیم دو سری توانی |

شکل ۳-۴: رکوردهای دارای اطلاعات ناقص

ت-بازیابی رکوردهای فاقد چکیده و رکوردهای دارای چکیده نادرست

بیش بینی کیفیت روابط جنسی بر اساس هوش هیجانی و اهمال کاری زناشویی

دریافت فایل

پایان نامه . غیر دولتی - دانشگاه آزاد اسلامی - دانشگاه آزاد اسلامی واحد مرودشت - دانشکده علوم تربیتی و روانشناسی . 1393 . کارشناسی ارشد

موضوع: علوم انسانی <

استاد راهنما: نادره سهرابی شگفتی | استاد مشاور: سیامک سامانی | دانشجو: هاجر جاهدیری

چکیده نادرست

ا دانشگاه آزاد اسلامی واحد مرودشت دانشکده علوم تربیتی و روان شناسی پایان نامه برای دریافت کارشناسی ارشد (M.A) در رشته روانشناسی عمومی عنوان: پیش بینی کیفیت روابط جنسی بر اساس هوش هیجانی و اهمال کاری زناشویی استاد راهنما: دکتر نادره سهرابی استاد مشاور: دکتر سیامک سامانی نگارنده: هاجر جاهدیری 2931-39 آ ب تقدیم به : همسر که در تمام مراحل چون کوه بر او تکیه کردم و با سپاس از حمایت ها و صبوری هایش، و تقدیم به نوگل باغ زندگیم بهینا ستاره زندگیتان بر فروغ . وبا تقدیم فراوان به: پدر و مادرم که همیشه عاشقانه چراغ راه زندگی ام بوده اند و بوسه میزنم بردستان پر از مهرشان ، آنان که هیچ گاه حمایت های خویش را از من دریغ نکردند، سایه شما مستدام، ج با تشکر و قدر دانی فراوان: از زحمت ها و تلاش های فراوان استادان گرانقدرم : سرکار خانم دکتر نادره سهرابی و جناب آقای دکتر سیامک سامانی که با راهنمایی های خود و صبر و شکیبایی بسیار، دلسوزانه و در سخاوت علمی تمام بنده را در به ثمر رساندن این پدیده علمی یاری دادند . د فهرست مطالب عنوان صفحه فصل اول کلیات پژوهش چکیده

مقدمه	2
مسئله	9
بیان	
اهمیت	5
پژوهش	
8	
پژوهش	
21	
تعریف نظری و تعریف عملیاتی	
22	
فصل دوم	
مبانی	
نظری	
کیفیت	
روابط	
جنسی	
21	
نظریه های	
جنسی	
21	
فروید	

شکل ۳-۵: رکورد دارای چکیده نادرست

بازیابی رکوردهای دارای عنوان نادرست

زمینه و هدف: امروزه در مقایسه با گذشته شاهد افزایش روز افزون شاعل بودن زنان و تغییر نقش های سنتی آنان هستیم. این تغییر نقش ها می تواند رضایت زناشویی زوجین و همچنین جنبه های مختلف زندگی و فرزندان را تحت تاثیر قرار دهد و از آن می توان به عنوان عنصری موثر در تعارضات زناشویی یاد کرد. این پژوهش باهدف مقایسه خصیصه سلطه گری در زنان شاعل و غیر شاعل و رابطه آن با رضایت زناشویی آنان انجام گرفت. روش تحقیق: این پژوهش از نوع توصیفی - پیمایشی بوده و به روش همبستگی مورد ارزیابی قرار گرفته است. روش نمونه گیری بصورت تصادفی بوده و تعداد 120 نفر از زنان مناهل که 60 نفر شاعل و 60 نفر خانه دار در بازه سنی 35-45 سال انتخاب و بررسی شدند. برای سنجش خصیصه سلطه گری از پرسشنامه برن روینر و میزان رضایت زناشویی از پرسش نامه انبرج استفاده شد. داده های حاصل از این پژوهش از طریق آزمون t برای دو گروه مستقل در سطح $p \leq 0.001$ و ضریب همبستگی پیرسون مورد تجزیه و تحلیل قرار گرفتند. یافته ها: میانگین نمره سلطه گری در زنان شاعل 62/13 و در زنان غیرشاعل 07/5- بدست آمده اختلاف میان آنها معنی دار بود. همچنین میانگین نمره رضایت زناشویی در زنان شاعل 28/144 و در زنان غیر شاعل 43/174 بدست آمد که این اختلاف نیز معنی دار بود. بین سلطه گری زنان شاعل و رضایت زناشویی و همچنین سلطه گری زنان غیرشاعل و رضایت زناشویی همبستگی بدست نیامد. نتیجه گیری: بر اساس یافته های این مطالعه، به نظر می رسد زنان شاعل سلطه گر تر از غیر شاعلین و همچنین رضایت زناشویی آنان کمتر از غیرشاعلین می باشد. کلید واژه ها: سلطه گری، رضایت زناشویی، زنان شاعل، زنان غیرشاعل

پایان نامه، غیر دولتی - دانشگاه آزاد اسلامی - دانشگاه آزاد اسلامی واحد شاهرود - دانشکده ادبیات و علوم انسانی، 1394، کارشناسی ارشد

موضوع: علوم انسانی >

استاد راهنما: ابوالفضل کرمی | دانشجو: محمدحسن صباغ

دریافت فایل

زمینه و هدف: امروزه در مقایسه با گذشته شاهد افزایش روز افزون شاعل بودن زنان و تغییر نقش های سنتی آنان هستیم. این تغییر نقش ها می تواند رضایت زناشویی زوجین و همچنین جنبه های مختلف زندگی و فرزندان را تحت تاثیر قرار دهد و از آن می توان به عنوان عنصری موثر در تعارضات زناشویی یاد کرد. این پژوهش باهدف مقایسه خصیصه سلطه گری در زنان شاعل و غیر شاعل و رابطه آن با رضایت زناشویی آنان انجام گرفت. روش تحقیق: این پژوهش از نوع توصیفی - پیمایشی بوده و به روش همبستگی مورد ارزیابی قرار گرفته اس ...

نمایه ها:

سلطه گری 365 | رضایت زناشویی 365 | شاعل 365 | غیرشاعل 365 |

شکل ۳-۶: عنوان نادرست

ث- بازیابی رکوردهای فاقد پیوست

ج- درج عدد در نام نویسندگان

ح- درج ایمیل یا نام کاربری به جای نام نویسندگان

خ- کلیدواژه های نادرست

د- بازیابی رکوردهایی که اطلاعات آنها با پیوست متفاوت است.

۳-۳. نتیجه‌گیری

لاگ‌های کاربران در بازه سه ماهه در پایگاه گنج نشان می‌دهد که بیشتر میزان بازیابی براساس زنجیره‌ای از دو کلمه صورت گرفته است. پس از دو کلمه‌ای‌ها بیشتر میزان بازیابی به ترتیب با یک کلمه و سپس سه کلمه‌ای صورت گرفته است. بر این اساس نقش کلیدواژه‌های دو یا یک کلمه‌ای در بازیابی روشن می‌شود. بنابراین سیستم باید به گونه‌ای باشد که بتواند فیلدهای متفاوت براساس یک کلمه یا دو کلمه را به صورت دسته‌بندی در اختیار کاربر قرار دهد به نحوی که کاربر در ادامه بتواند به جستجوی مورد نظر برسد.

بررسی لاگ‌های کاربران به لحاظ ساخت زبانی روشن می‌کند که چه ساخت‌های زبانی بیشترین کاربرد را دارد. بررسی هزار نمونه از این لاگ‌ها نشان می‌دهد که:

- ۱- بیشتر جستجوها براساس ساخت زبانی اسم + اسم است.
- ۲- تنها یک مورد جستجو براساس نام پدیدآور در هزار مورد قابل مشاهده است.
- ۳- در حدود ۵۰ مورد جستجوی لاتین در بین ۱۰۰۰ مورد قابل مشاهده است.

این نتایج نشان می‌دهد که جستجو براساس نام پدیدآور به ندرت صورت می‌گیرد. و دیگر این که از آنجا که بازیابی به لاتین هم صورت می‌گیرد شناسایی مسائل مدارکی که به زبان لاتین هستند، نیز ضروری است.

مسئله دیگری که لازم است روشن شود که آیا کلیدواژه‌های بکار رفته از سوی کاربران به چه میزان با کلیدواژه‌های ذخیره شده هماهنگی دارد. به عبارتی چند درصد از کلیدواژه‌های نمایه‌سازی شده توسط کاربران مورد استفاده قرار می‌گیرد.

در نهایت می‌توان مسائل بازیابی که ریشه در مسائل زبانی متن فارسی دارند را به صورت زیر برشمرد.

الف- تفاوت در رکورد بازیابی ناشی از کاربرد صورتهای متفاوت نوشتاری:

- ۱- کاربرد صورت‌ها و علامتهای مختلف نوشتاری: تشدید، مد، همزه، تنوین، آوا، نقطه، خط تیره
- ۲- کاربرد صورتهای متفاوت کسره اضافه
- ۳- کاربرد صورتهای مختلف جمع
- ۴- کاربرد صورتهای مختلف اختصارات

ب- تفاوت در رکورد بازیابی ناشی از مسائل معنایی-صرفی-نحوی زبان فارسی

۱- کلمات چندمعنا با املای یکسان (اما صورت آوایی یکسان یا متفاوت)

۲- صورت‌های متفاوت واژگانی با معنای یکسان یا مرتبط

۳- صورت‌های کوتاه / کامل اسامی یا سازمان‌ها

فصل چهار

ارائه راهکار برای حل مسائل ساماندهی و بازیابی اطلاعات در گنج

۴. مقدمه

در دو فصل گذشته برخی از مسائل زبانی ساماندهی و بازیابی اطلاعات در متون فارسی که با بررسی میدانی پایگاه بدست آمد، را برشمردیم.

در بخش ساماندهی و نمایه سازی برخی مسائل زبانی به صورت زیر است:

الف- استفاده از معادل‌های متعدد برای یک واژه لاتین: به نظر می‌رسد ساماندهی این موارد باید به گونه‌ای صورت گیرد که معادل‌های متعدد در کنار هم آورده شود. بدین صورت که معادل‌های متعددی که برای یک واژه بکار می‌رود، به نحوی در اختیار کاربر قرار گیرد و به او پیشنهاد شود. در این صورت بازیابی آنها نیز بهتر صورت می‌گیرد.

ب- تعیین کلیدواژه: (۱) تعیین معادل فارسی کلیدواژه‌های لاتین (۲) تعیین صورت‌های اختصاری کلیدواژه‌ها

ت- نوشتن صورت‌های مخفف لاتین به اشکال مختلف: مثلاً با علائم نگارشی نظیر نقطه، خط فاصله و رقم (اچ. وی- سی ۲ یا اچ وی سی نوع ۲ و غیره). این تفاوت‌های نگارشی رکوردهای بازیابی متفاوتی را بدست می‌دهد.

ث- تعیین اختصار فارسی برای یک واژه لاتین: به عنوان مثال اصطلاح *voluntary counselling and testing* دارای صورت اختصاری (VCT) است اما صورت مختصر فارسی آن (ماد) برای معادل «مشاوره و آموزش اختیاری» تعیین شده است. بازیابی با این صورت‌های مختلف منجر به تولید رکوردهای مختلف می‌شود. لازم است در بازیابی همه این صورت‌ها با هم دیده شود.

ج- تعدد صورت‌های نوشتاری اسامی خاص افراد (پدیدآوران): اول نام پدیدآوران ممکن است به چند صورت ملاحظه شود مثلاً شمس لیالستانی یا فیض شمس لیالستانی. دوم اینکه یک نام ممکن است به صورت‌های مختلفی نوشته شود. (یدالله یا یداله یدا...) به دلیل صورت‌های مختلف نوشتاری، وجود فاصله یا نیم فاصله و غیره و سوم این که نام افراد (مانند نام سازمانها) ممکن است در طی زمان تغییر کند و یا صورت‌های دیگری مثل اسم مستعار بکار رود که در این صورت لازم است همه صورت‌ها در بازیابی مشاهده شود.

ح- انواع خط‌های املایی و نگارشی در کلیدواژه‌ها، چکیده و پدیدآوران شخص در هر دو سطح فارسی و لاتین

خ- درج مطالب نامربوط در چکیده: به عنوان مثال درج کلمه چکیده یا گذاشتن کلیدواژه‌ها در چکیده

د- مسائل مربوط به درج علائم ریاضی و فرمولهای شیمی در چکیده و عنوان

مسائل بازیابی که با بررسی میدانی در پایگاه گنج بدست آمدند، به صورت زیر است:

الف- صورت‌های متفاوت نوشتاری

۱- ناشی از کاربرد علامتهای مختلف نوشتاری: تشدید، مد، همزه، تنوین، آوا، نقطه، خط تیره

۲- صورت‌های متفاوت کسره اضافه

۳- صورت‌های مختلف جمع

۴- صورت‌های مختلف اختصارات

ب- مسائل معنایی-صرفی-نحوی

۵- کلمات چندمعنا با املاهای یکسان (اما صورت آوایی یکسان یا متفاوت)

۶- صورت‌های متفاوت واژگانی با معنای یکسان یا مرتبط

۷- صورت‌های کوتاه / کامل اسامی یا سازمان‌ها

۱-۴. مسائل سازماندهی و نمایه سازی و راهکارهای پیشنهادی

با نگاهی به مسائلی که در بالا مطرح شده روشن می شود که این مسائل انواع متفاوتی دارند. حل این مسائل مستلزم فعالیت ها و اقدام های متفاوتی است. به نظر نگارنده برای حل این مسائل دانش، فعالیت و اقداماتی لازم است که در زیر به آنها اشاره شده است و مسائل بر اساس نیازی که به دانش، فعالیت و اقدامات متفاوت دارند، در دسته های مختلف قرار گرفته اند:

۴-۱-۱. دانش تخصصی

مورد زیر به عنوان مسئله ای در ساماندهی و نمایه سازی مطرح شد:

- تعیین کلیدواژه ها: ۱) تعیین معادل فارسی کلیدواژه‌های لاتین ۲) تعیین صورت‌های اختصاری کلیدواژه‌ها

تعیین کلیدواژه‌های جدید و یافتن معادل فارسی کلیدواژه‌های لاتین، صورتهای مخفف و اختصاری مستلزم دانش تخصصی رشته‌های مختلف علمی و نیز دانش تخصصی در زمینه زبان فارسی است، بنابراین حل این مسئله در گرو تعامل با متخصصان رشته‌های مختلف علمی و نیز زبان فارسی و زبان‌شناسی است.

۴-۱-۲. جستجو، تصمیم‌گیری و بهبود قابلیت نرم افزار نمایه سازی

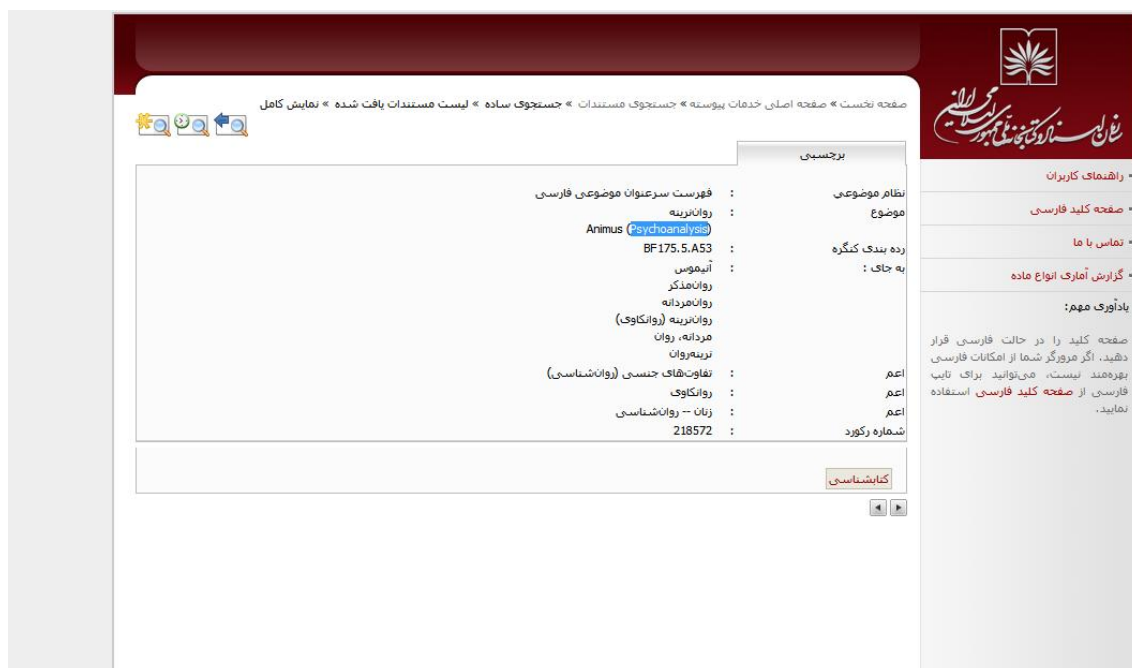
دسته ای از مسائلی که بالا به آنها اشاره شد، به صورت زیر هستند:

- صورت‌های متفاوت اسامی خاص و پدیدآورندگان شخص
- وجود معادل‌های متعدد برای یک واژه لاتین
- صورت اختصاری بدست آمده از معادل فارسی یک واژه لاتین (بنابراین متفاوت با صورت اختصاری لاتین آن واژه)
- صورت‌های متفاوت نوشتاری برای صورتهای مخفف لاتین

در مسائل بالا، نمایه ساز با تنوع و تفاوت در تعداد کلیدواژه‌ها و یا اسامی خاص روبرو است. بنابراین نمایه ساز برای تصمیم‌گیری درباره اینکه کدام صورت درست است و می‌تواند ذخیره شود، با دشواری روبرو است و دچار سردرگمی و تردید است. علاوه بر آن این تنوع در تعداد بر میزان دقت بازیابی نیز تأثیر منفی دارد. به این صورت که کاربر وقتی مدخلی را برای بازیابی وارد می‌کند به همه رکوردهای آن مدرک دسترسی پیدا نمی‌کند زیرا به عنوان مثال دسته ای از مدارک تنها با صورت «شمس لیالستانی» ذخیره شده‌اند و بنابراین کاربر به مدخل‌هایی که با صورت کاملتری مثل «یدالله فیض شمس لیالستانی» ذخیره شده‌اند، دسترسی ندارد و در نتیجه دقت بازیابی کاهش پیدا می‌کند.

برای حل مسائل بالا برخورداری از دانش تخصصی در زمینه‌های مختلف علمی و دانش تخصصی زبان فارسی ضروری است، علاوه بر آن لازم است منابع مختلف، ادبیات موجود در زمینه کلیدواژه‌ها و اختصارات به درستی مورد بررسی قرار گیرد. تصمیم‌گیری برای یکدست کردن یا بکار بردن صورت‌های مختلف گام دیگری است که مستلزم دانش کارشناسی در زمینه کلیدواژه و زبان فارسی است. از سویی به نظر می‌رسد باید به گونه ای نرم افزار نمایه سازی بهبود یابد تا امکان ذخیره و ارائه همه صورت‌های متفاوت برای کاربر باشد. به عنوان مثال اگر کاربر در جستجوی VCT است بهتر است مخفف فارسی آن یعنی «ماد» نیز به کاربر پیشنهاد گردد.

به عنوان مثال جستجوی رکوردها در سایت کتابخانه ملی امکان مشاهده سایر کلیدواژه‌ها و ارجاعات را نیز در اختیار کاربر قرار می‌دهد (شکل زیر). مدلی که می‌تواند مورد استفاده قرار گیرد.



شکل ۴-۱: مثالی برای همه ارجاعات

۴-۱-۳. ویرایش و اصلاح متن

مسائل زیر در حال حاضر به طور دستی از سوی نمایه ساز برطرف می‌شود. برای حل این موارد و افزایش سرعت ویرایش و اصلاح متن نیاز به نرم افزارهای ویرایش متن فارسی است.

- انواع خطاهای املایی و نگارشی
- حذف مطالب نامربوط در چکیده و یا عنوان
- مسائل مربوط به درج علائم ریاضی و فرمولهای شیمی

در ادامه این فصل نگارنده کیفیت خطایاب‌های موجود برای زبان فارسی را بررسی می‌نماید تا روشن شود آیا امکان استفاده از این خطایابها برای کاهش میزان خطاهای املایی و تایپی مدارک موجود در پایگاه گنج میسر است؟

۴-۱-۴. خطایابی و رفع اشتباهات تایپی و نگارشی

اکثر خطاهای تایپی و نگارشی را می‌توان به کمک خطایاب رفع نمود. خطایابهای موجود برای زبان فارسی شامل موارد زیر است: ویراستیار، وفا، ویرا. ویراسباز، مجموعه زبانی میکروسافت. با بررسی نگارنده به نظر می‌رسد از بین خطایاب‌های موجود، خطایاب وفا از کارایی بهتری برخوردار باشد. برای بررسی این موضوع نگارنده خطایاب مذکور را نصب نمود و عملکرد آن را برای خطایابی در چکیده‌ها مورد بررسی قرار داد. با نصب این خطایاب، آیکون آن در نوار بالای برنامه «ورد» ایجاد می‌شود و با انتخاب متن و کلیک بر روی آیکون آن، خطاهای متن به دو صورت اصلاح می‌شوند. بدین نحوه که دسته‌ای از خطاها به کاربر پیشنهاد و جایگزینی آنها درخواست می‌شود و برخی خطاها نظیر ایجاد نیم فاصله به جای فاصله به طور خودکار صورت می‌گیرد.

۴-۱-۵. خطایاب وفا

شکل زیر نمونه‌ای از خطاهای تایپی مثل اشتباه در تایپ کلمه، درج فاصله به جای نیم فاصله، اتصال دو کلمه به دلیل حذف فاصله بین کلمات، درج نادرست علائم سجاوندی و غیره را نشان می‌دهد.

بررسی رابطه دینداری و نگرش نسبت به دموکراسی؛ مطالعه موردی: دانشجویان دانشگاه تربیت معلم (0) (0) (0) (0) (0) (0)

بایان‌نامه . دولتی - دانشگاه پیام نور (وزارت علوم، تحقیقات، و فناوری) - دانشگاه پیام نور استان تهران - مرکز پیام نور تهران - دانشکده علوم اجتماعی . 1388 . کارشناسی ارشد

موضوع: علوم اجتماعی > جامعه‌شناسی

استاد راهنما: محمدجواد زاهدی مازندرانی | استاد مشاور: پروانه دانش | پدیدآور: مهراڤ صولتی

دسترسی به فایل تمام‌متن پیشینه‌هایی (رکوردهایی) که نشانه «پی.دی.اف.» ندارند، شدنی نیست.

پژوهش حاضر به منظور بررسی رابطه دینداری دانشجویان دانشگاه تربیت معلم سبزوار با نگرش نسبت به دموکراسی می‌باشد و سعی در ارائه تحلیل جامعه‌شناختی از رابطه دینداری و نگرش نسبت به دموکراسی دارد. روش این تحقیق پیمایشی است. به لحاظ معیار زمان این پیمایش از نوع مقطعی بوده و به لحاظ معیار زمانی از نوع پیمایشی است که داده‌های آن با استفاده از پرسشنامه گردآوری شده است. سنجش میزان دینداری افراد براساس مدل چند بعدی گلاک و استارک صورت گرفته است. همچنین نگرش نسبت به دموکراسی در سه بخش حقوق و آزادی‌های مدنی، حکومت دموکراتیک و پاسخگو و جامعه دموکراتیک سنجیده شده است. نتایج پژوهش نشان می‌دهد بعد مناسکی، سایر ابعاد دینداری از جمله اعتقادی، عاطفی و پیامدهای رابطه معنی‌داری در سطح اطمینان 95 درصد با نگرش نسبت به دموکراسی ندارند. یافته‌های پژوهش همچنین نشان می‌دهد که بین متغیرهای زمینه‌ای سن، جنس محل سکونت و مدت حضور در دانشگاه و نگرش نسبت به دموکراسی در سطح اطمینان 95 درصد رابطه وجود دارد.

نمایه‌ها:

دینداری | دموکراسی | تعصب دینی | مردم‌سالاری | دانشگاه تربیت معلم | سبزوار | دانشجویان | ساختار اجتماعی | اختلال رفتاری |

شکل ۴-۲: چکیده دارای خطاهای املائی

شکل زیر نشان می‌دهد که چگونه برخی از این خطاهای در نرم افزار وفا تصحیح شد: نرم افزار کلماتی که با رنگ آبی مشخص شده‌اند، را به کاربر برای تغییر پیشنهاد می‌کند و کلماتی که با رنگ خاکستری نشان داده شده‌اند، را به طور خودکار اصلاح کرده است. و کلماتی که دور آنها دایره رسم شده است، بدون تغییر و اصلاح باقی مانده‌اند.

پژوهش حاضر به منظور بررسی رابطه دینداری دانشجویان دانشگاه تربیت معلم سبزوار با نگرش نسبت به دموکراسی می‌باشد. وسعی در ارائه تحلیل جامعه شناختی از رابطه دینداری و نگرش نسبت به دموکراسی دارد. روش این تحقیق پیمایشی است. به لحاظ معیار زمان این بی مایش از نوع مقطعی بوده و به لحاظ معیار ژرفایی از نوع بهنانگر است که داده‌های آن با استفاده از پرسشنامه گردآوری شده است. سنجش میزان دینداری افراد براساس مدل چند بعدی گلاک و استارک صورت گرفته است. همچنین نگرش نسبت به دموکراسی در سه بخش حقوق و آزادی‌های مدنی، حکومت دموکراتیک و پاسخگو و جامعه دموکراتیک سنجیده شده است. نتایج پژوهش نشان می‌دهد بجز بعد مناسکی، سایر ابعاد دینداری از جمله اعتقادی، عاطفی و پیامدهای رابطه معنی داری در سطح اطمینان ۹۵ درصد با نگرش نسبت به دموکراسی ندارند. یافته‌های پژوهش همچنین نشان می‌دهد که بین متغیرهای زمینه ای سن، جنس محل سکونت و مدت حضور در دانشگاه و نگرش نسبت به دموکراسی در سطح اطمینان ۹۵ درصد رابطه وجود دارد.

با اصلاح وفا ۱۴ مورد اشتباه تایپی تشخیص داده شد که برخی خودکار اصلاح شده‌اند (رنگ خاکستری) و برخی برای اصلاح به کاربر پیشنهاد می‌شوند (رنگ آبی) و ۹ مورد خطا باقیمانده است (کلماتی که دور آنها دایره کشیده شده است. کلمه «پیامدهای» که دور آن دایره کشیده شده است، دارای اشتباه دستوری است که نرم افزار قادر به تشخیص آن نیست. و تنها به وسیله نویسنده به عنوان موردی از خطا گزارش شده است). خطاهای باقیمانده از چند نوع هستند:

- ۱- عدم تصحیح خطا تایپی کلمه مثل تایپ «چند» به جای «چند»، یا «بی مایش به جای پیمایش»
- ۲- عدم تشخیص صورت نوشتاری نادرست مثال: «بجز» به جای «به جز»
- ۳- عدم درج نیم فاصله، مثل «جامعه شناختی و زمینه ای»
- ۴- عدم درج فاصله مثل «وسعی» به جای «و سعی»

در شکل زیر نمونه دیگری از یک مدرک دارای خطاهای تایپی و نگارشی متعدد در پایگاه گنج قابل مشاهده است.

تحلیلی بر متریک فعالیت توسعه‌دهنده درشناسایی آسیب‌پذیری نرم‌افزار با استفاده از کاوش مخازن نرم‌افزاری

(0) (0) (0) (0) (0) (0)

مقاله‌های همایش‌های ایران . دوازدهمین کنفرانس ملی سیستم‌های هوشمند . انجمن سیستم‌های هوشمند ایران . 1392

نویسنده: جمیله شفیعی . اشکان سامی

شناسایی آسیب‌پذیری نرم‌افزار، بعد از انتشار، نه تنها موجب تحمیل هزینه‌های اضافی جهت رفع آن می‌شود، بلکه باعث از بین رفتن اعتبار شرکت نرم‌افزاری نیز می‌شود. بنابراین هر سازمانی باید شناسایی آسیب‌پذیری را در اولویت‌های خود قرار داده و از رخداد آن، جلوگیری نماید. تکنیک‌هایی که برای جلوگیری و شناسایی آسیب‌پذیری نرم‌افزار قبل از انتشار وجود دارند، تکنیک‌های هستند که به پایش چشمی کد می‌پردازند، بنابراین دشوار و زمان‌بر هستند. علاوه بر این، پایش چشمی همگی قسمت‌های کد نیز کاری پرهزینه است و با توجه به منابع محدود تیم توسعه‌دهنده، این کار امکان‌پذیر نمی‌باشد. بنابراین در این مقاله، به دنبال خودکارسازی شناسایی آسیب‌پذیری نرم‌افزار، هستیم. با توجه به این که توسعه‌ی نرم‌افزار، به طور معمول، توسط تیم توسعه‌ای که با هم روی یک پروژه‌ی مشترک کار می‌کنند، انجام می‌شود، روابط انسانی نقش‌بررنگی در توسعه‌ی هر پروژه دارند. علاوه بر این، بازبینیها و نظارت‌هایی که انجام می‌شود نیز توسط عوامل انسانی انجام می‌شوند. همچنین عدم پیوستگی تیم و برداشت نادرست در ارتباطات، می‌تواند موجب ایجاد نرم‌افزاری گردد که آسیب‌پذیر است. با توجه به اهمیت موارد مطرح شده، استفاده از متریک‌های استخراجی حاصل از بررسی روابط انسانی و توسعه‌دهندگان، متریک فعالیت توسعه‌دهنده، تأثیر بسزایی در شناسایی نقاط آسیب‌پذیر کد خواهد داشت. در این مقاله، با طرح سه فرضیه، به بررسی و تحلیل تأثیر متریک فعالیت توسعه‌دهنده، بر شناسایی آسیب‌پذیری یک نرم‌افزار متن باز (یازده نسخه مرورگر وب موزیلا فایرفاکس)، پرداخته می‌شود. طی بررسی‌های انجام شده، تأثیر زیاد متریک فعالیت توسعه‌دهنده بر شناسایی فایل‌های آسیب‌پذیر، نسبت به فایل‌های فاقد آسیب‌پذیری، از نظر آماری چشمگیر است. بنابراین از این متریک می‌توان جهت شناسایی خودکار آسیب‌پذیری و هدایت سازمان و تیم توسعه‌دهنده جهت شناسایی مکان کد آسیب‌پذیر، صرف منابع و بهبود کد، استفاده کرد

شکل ۴-۲: چکیده دارای خطاهای املائی

چکیده متن بالا به وسیله نرم افزار وفا به صورت زیر اصلاح شد:

- ۱- کلماتی که به رنگ زرد نشان داده شده‌اند، برای اصلاح به وسیله نرم افزار به کاربر پیشنهاد می‌شوند.
- ۲- کلماتی که با رنگ خاکستری مشخص شده‌اند، به طور خودکار به وسیله نرم افزار اصلاح شده‌اند.
- ۳- کلماتی که دور آنها دایره رسم شده است، به وسیله نرم افزار اصلاح نشده‌اند.

شناسایی آسیب‌پذیری نرم‌افزار، بعد از انتشار، نه تنها موجب تحمیل هزینه‌های اضافی جهت رفع آن می‌شود، بلکه باعث از بین رفتن اعتبار شرکت نرم‌افزاری نیز می‌شود. بنابراین هر سازمانی باید شناسایی آسیب‌پذیری را در اولویت‌های خود قرار داده و از رخداد آن، جلوگیری نماید. تکنیک‌هایی که برای جلوگیری و شناسایی آسیب‌پذیری نرم‌افزار قبل از انتشار وجود دارند، تکنیک‌های هستند که به پایش چشمی کد می‌پردازند، بنابراین دشوار و زمان‌بر هستند. علاوه بر این، پایش چشمی همه قسمت‌های کد نیز کاری پرهزینه است و با توجه به منابع محدود تیم توسعه‌دهنده، این کار امکان‌پذیر نمی‌باشد. بنابراین در این مقاله، به دنبال خودکارسازی شناسایی آسیب‌پذیری نرم‌افزار، هستیم. با توجه به اینکه توسعه نرم‌افزار، به طور معمول، توسط تیم توسعه ای که با هم روی یک پروژه مشترک کار می‌کنند، انجام می‌شود، روابط انسانی نقش پررنگی در توسعه هر پروژه دارند. علاوه بر این، بازبینیها و نظارت‌هایی که انجام می‌شود نیز توسط عوامل انسانی انجام می‌شوند. همچنین عدم پیوستگی تیم و برداشت نادرست در ارتباطات، می‌تواند موجب ایجاد نرم‌افزاری گردد که آسیب‌پذیر است. با توجه به اهمیت موارد مطرح شده، استفاده از متریک‌های استخراجی حاصل از بررسی روابط انسانی و توسعه دهندگان، متریک فعالیت توسعه‌دهنده، تأثیرپذیری در شناسایی نقاط آسیب‌پذیر کد خواهد داشت. در این مقاله، با طرح سه فرضیه، به بررسی و تحلیل تأثیر متریک فعالیت توسعه‌دهنده، بر شناسایی آسیب‌پذیری یک نرم‌افزار متن‌باز (یازده نسخه مرورگر وب موزیلا فایرفاکس)، پرداخته می‌شود. طی بررسی‌های انجام شده، تأثیر زیاد متریک فعالیت توسعه‌دهنده بر شناسایی فایل‌های آسیب‌پذیر، نسبت به فایل‌های فاقد آسیب‌پذیری، از نظر آماری چشمگیر است. بنابراین از این متریک می‌توان جهت شناسایی خودکار آسیب‌پذیری و هدایت سازمان و تیم توسعه‌دهنده جهت شناسایی مکان کد آسیب‌پذیر، صرف منابع و بهبود کد، استفاده کرد.

همان‌طور که نتیجه اصلاح خطاها نشان می‌دهد تعداد موارد اصلاح شده یا پیشنهاد شده به کاربر برای بازنگری، ۳۷ مورد اصلاح است و در مقابل ۲۲ مورد خطا باقیمانده است. نکاتی که قابل مشاهده است، این است که: ایجاد یا حذف نیم فاصله که گاهی به عنوان خطا در نظر گرفته شده مثلاً در کلمات «نرم‌افزار یا نرم افزار، می‌شود، اولویت‌های» ممکن است در جاهای دیگر به عنوان خطا پیشنهاد نشده است، به عنوان مثال در «آسیب‌پذیری، نمی‌باشد، تکنیک‌هایی و غیره». در کلمه «نرم‌افزار» گاهی خطا گرفته شده است و گاهی نشده است.

۲-۴. نتیجه‌گیری

در نرم افزار وفا تاحدی خطاها اصلاح می‌شوند ولی نرم افزار از دقت کافی برای اصلاح همه خطاهای موجود در متن برخوردار نیست. از سویی این نرم افزار در تعامل با کاربر کار می‌کند و تنها بخشی از خطاها به طور خودکار تصحیح می‌شوند. بنابراین نقش کمکی در تصحیح خطاها دارد. بررسی نشان می‌دهد که این برنامه تاحدودی خطاها را اصلاح می‌کند ولی از دقت لازم برخوردار نیست. نکته دیگر این که هر گونه برنامه خطایابی باید در مرحله نمایه سازی روی چکیده، عنوان یا سایر قسمت‌های یک مدرک فعال شود و لازم نباشد که نمایه ساز متن را به داخل برنامه «ورد» انتقال دهد زیرا در این صورت زمان و حوصله بیشتری لازم است که با توجه به حجم بالای مدرک‌ها در پایگاه گنج امکان پذیر نیست.

۳-۴. مسائل بازیابی و راهکارهای پیشنهادی

در فصل گذشته به برخی از مسائل زبانی بازیابی اطلاعات که با بررسی میدانی پایگاه گنج بدست آمد، اشاره کردیم. این مسائل به دو دسته تقسیم شدند.

الف- صورت‌های متفاوت نوشتاری

۱- ناشی از کاربرد علامتهای مختلف نوشتاری: تشدید، مد، همزه، تنوین، آو، نقطه، خط تیره

۲- صورت‌های متفاوت کسره اضافه

۳- صورت‌های مختلف جمع

۴- صورت‌های مختلف اختصارات

ب- مسائل معنایی-صرفی-نحوی

۱- کلمات چندمعنا با املاي یکسان (اما صورت آوایی یکسان یا متفاوت)

۲- صورت‌های متفاوت واژگانی با معنای یکسان یا مرتبط

۳- صورت‌های کوتاه / کامل اسامی یا سازمان‌ها

در ادامه به برخی منابع که سعی در ارائه راه حل‌ها و راهکارهایی برای حل مشکلات زبانشناختی بالا و بهبود بازیابی اطلاعات بر آن اساس پرداخته‌اند، ارائه می‌شود.

۴-۳-۱. مدل‌های زبانشناختی بازیابی اطلاعات

آرامپاتزیس^۶ و همکاران (۱۹۹۸) به این نتیجه می‌رسند که سیستم‌های بازیابی اطلاعات که مبتنی بر کلیدواژه هستند از مسائلی رنج می‌برند که ریشه در آنچه آنها تنوعات زبانشناختی^۷ می‌نامند، است. آنها این تنوعات زبانی را به ۴ دسته تقسیم می‌کنند (شامل همان مواردی است که در بالا با مطالعه میدانی بدست آمد):

۱- تنوعات واژگانی^۸: سیستم‌های بازیابی مبتنی بر کلیدواژه در مواردی که واژه‌های متفاوت بیانگر

معنا یا مفهومی یکسان در سؤال کاربر و مدارک هستند، عمل نمی‌کنند. به عنوان مثال اگر

کلیدواژه سؤال کاربر «film» باشد، مدارکی که در آنها کلمه مترادف movie وجود داشته

باشد، بازیابی نمی‌شوند.

^۶ . Arampatzis A. T.

^۷ . linguistic variation

^۸ . lexical variation

۲- تنوعات معنایی^۹: سیستم‌های بازیابی مبتنی بر کلیدواژه در مواردی که یک کلمه چند معنا دارد، بین این معانی تفاوت و تمایز ایجاد نمی‌کنند. به عنوان مثال اگر یک آوازخوان به دنبال کلمه «بندها» bands باشد ممکن است با بازیابی «بندهای فرانکسی رادیو» radio frequency bands مواجه شود.

۳- تنوعات نحوی^{۱۰}: این سیستم‌ها در رابطه با مسئله تنوعات نحوی عملکرد خوبی ندارند. به عنوان مثال اگر مدرکی حاوی جمله «نزدیک رودخانه، آلودگی هوا یک مسئله بزرگ است» باشد، این مدرک ربطی به جستجوی «آلودگی رودخانه» ندارد.

۴- تنوعات صرفی و ساختاری^{۱۱}: کلیدواژه‌ها ممکن است به دلیل تنوعات ساختاری و صرفی به صورتهای مختلف جمع wolf- wolves یا شکل‌های مختلف صرفی مثل man-man's ظاهر شوند. بازیابی برحسب تنوعات صرفی یکسان نیست.

همه این مسائل که در بالا به آنها اشاره شد، به دقت و بازخوانی یک سیستم بازیابی اطلاعات آسیب می‌زنند. دقت یعنی نسبت مدارک مرتبط با بازخوانی شده به کل مدارک بازیابی شده و بازخوانی یعنی نسبت مدارک مرتبط با بازخوانی شده به همه مدارک مرتبط در یک سیستم. پرداختن به تنوعات زبانی ۲ و ۳ دقت بازیابی را افزایش می‌دهد و پرداختن به تنوعات زبانشناختی ۱ و ۴ بازخوانی را بهبود می‌بخشد. اما به طور کلی معمولاً بهبود دقت سبب کاهش بازخوانی می‌شود و برعکس (آرامپاتزیس و همکاران، ۱۹۹۸).

روشی که آرامپاتزیس^{۱۲} و همکاران (۱۹۹۸) برای حل مسائل زبانشناختی بکار می‌برند، مبتنی بر تکنیک‌های پردازش زبان طبیعی (NLP) است و «بازیابی اطلاعات مبتنی بر عبارت»^{۱۳} نامیده می‌شود. این روش بنابراین تنوعات زبانشناختی را دربر می‌گیرد. کاربرد پردازش زبان طبیعی در بازیابی، رویکردی پذیرفته شده برای بهبود کارایی بازیابی است. بنابراین آنها بازیابی عبارت را به عنوان جایگزین بازیابی کلیدواژه معرفی می‌کنند و رویکرد آنها در بازیابی عبارت مبتنی بر گروه اسمی^{۱۴} است. بدین صورت که بازیابی بر مقیاسی از نزدیکی^{۱۵} گروه اسمی در سؤال با گروه اسمی در مدرک مبتنی است. اگر دو یا چند اسم و

^۹ . semantical variation

^{۱۰} . syntactical variation

^{۱۱} . morphological variation

^{۱۲} . Arampatzis A. T.

^{۱۳} . phrase-based IR

^{۱۴} . noun phrase (NP)

^{۱۵} . nearness

وابسته هایشان در یک NP دیده شوند، فرض بر این است که بین این اسمها نوعی رابطه با هم دارند حتی بدون دانستن معنای آنها. به عنوان مثال در جمله زیر

... tracks were recorded at the BBC studios for later radio programs ...

اسمهای radio و programs و اسم خاص BBC که در یک NP جمله قرار دارند، به لحاظ معنایی به هم مربوطند. بنابراین جستجوی سوالی مثل radio programs on BBC منجر به بازیابی مدارکی که حاوی عباراتی مثل بالا هستند می شود و نه سایر اشکال هم رخدادی^{۱۶} این اسمها مثل مدارک زیر:

Document ۱: the transmission of his first **radio programs** resembled the early years of the creation of **BBC** empire which...

Document ۲: ten musicians from the **BBC** Symphony Orchestra were interviewed in several **radio programs** of L.A. stations...

عبارات بالا پذیرفته نمی شوند به دلایل نحوی^{۱۷} زیرا سه کلمه سوال در یک NP قرار ندارند. آنها برای تجزیه نحوی یک تقطیع گر^{۱۸} NP ساختند. نتایج آنها نشان می دهد که کاربرد هم رخدادی NP باعث افزایش سطح دقت بازیابی به میزان بسیار چشمگیری می شود، بالای ۹۰ درصد در همه انواع جستجو ها. اما در مقابل آن بازخوانی بسیار پایین نشان داد. به نظر می رسد بازخوانی بهتر با هزینه اندکی کاهش دقت بدست می آید. آنها در این تحقیق خود بازنمایی ای از عبارات مناسب برای نمایه سازی و معماری ای برای چنین سیستم بازیابی پیشنهاد می کنند. و نرمالسازی نحوی با هدف بهبود کارایی بازیابی ارائه می شود.

وودز و همکاران (۲۰۰۰) نیز برای بهبود بازیابی اطلاعات از روش های پردازش زبان طبیعی استفاده کردند. نتایج آنها نشان می دهد که دانش زبانی می تواند به طور معناداری عملکرد بازیابی اطلاعات را بهبود ببخشد. دانش زبانی مورد نظر آنها شامل کاربرد روابط ساختواری بین کلمات، روابط طبقه ای^{۱۹} بین مفاهیم و روابط کلی معنایی^{۲۰} بین کلمات و مفاهیم است. نتایج آنها نشان می دهد که ترکیب این سه منبع دانش می تواند به طور معناداری توانایی یافتن پاسخ های مناسب به سوالها را بهبود ببخشد. نمایه سازی مفهومی^{۲۱} و سیستم بازیابی که آنها بکار بردند کلمات و عبارات را از متن استخراج کرده و آنها را

^{۱۶} . cooccurrence

^{۱۷} . syntactic

^{۱۸} . parser

^{۱۹} . taxonomic relationships

^{۲۰} . semantic entailment

^{۲۱} . conceptual indexing

به صورت یک شبکه معنایی که تکمیل‌کننده روابط ساختوازی، معنایی و نحوی است، سازمان‌دهی می‌کند. سیستم بازیابی آنها از یک واژگان استفاده می‌کند که حاوی اطلاعات صرفی، نحوی و معنایی درباره کلمات، معنای کلمات و عبارات است. علاوه بر آن دارای یک سیستم گسترده از توابع و قواعد ساختوازی دانش بنیاد برای تجزیه و تحلیل کلماتی که در واژگان موجود نیستند، است و سیستم علاوه بر قواعدی برای کار با صورتهای تصریفی و اشتقاقی کلمات شناخته شده، شامل قواعدی برای ترکیبات واژگانی و قواعدی برای حدس منطقی در مورد کلمات کاملاً ناشناخته است.

بنابراین به نظر می‌رسد روش‌های پردازش زبان طبیعی می‌تواند به بهبود بازیابی اطلاعات کمک کند. در زبان فارسی نیز برخی پژوهشها در این زمینه صورت گرفته است.

پژوهش ایزدی بروجنی (۱۳۹۱) نشان می‌دهد که کاربرد مدلهای زبانی سبب بهبود معیارهای دقت و فراخوانی در بازیابی اطلاعات می‌شود.

ستوده و هنرجویان (۱۳۹۱) نیز به بیش از ۴۰ دشواری نگارشی در رابطه با جستجو و بازیابی اطلاعات در زبان فارسی اشاره کرده‌اند. به گفته آنها «این گونه‌گونی نگارشی به نایکدستی و تطور بسیار در نگارش فارسی می‌انجامد که می‌تواند اثربخشی بازیابی را بویژه از منظر کاهش دقت یا ریزش کاذب و نیز کاهش جامعیت بازیابی، متأثر سازد. در نتیجه، ضروری است در طراحی الگوریتمهای سامانه‌های جستجو و بازیابی فارسی، به‌هنجارسازی تنوعات و چنددستیهای نگارشی و دستوری توجه شود». از جمله راهکارهای ارائه شده توسط آنها «تدوین استاندارد نگارش فارسی، استفاده از سیاهه‌های از پیش تعیین شده، تجهیز پایگاه اطلاعاتی به اصطلاحنامه و فرهنگهای املائی، و تدوین دستنامه یا راهنمای جستجو است. این راهکارها با وجود جامع نبودن، کم و بیش اثربخش به نظر می‌رسند».

هماوندی و همکاران (۱۳۹۵) نیز به مسائل بازیابی ناشی از ویژگی‌های زبان فارسی اشاره می‌کنند. «مواردی مانند پیوسته نویسی و جدانویسی واژگان مشتق، مرکب و مشتق-مرکب، گوناگونی جمعها، واژگان دخیل و معادل آنها در بخش نوشتاری و چندمعنایی، همنامی و در بخش معنایی از این دست ویژگیها هستند. فقدان پوشش مناسب ویژگیهای یاد شده در مراحل ذخیره سازی و پردازش و عدم آگاه نمودن کاربر از آن، جهت اصلاح فرایند کاوش در مرحله بازیابی اطلاعات در پایگاههای اطلاعاتی مورد پژوهش، اثرات نامطلوبی بر فرایند کاوش و بازیابی دارد».

۴-۴. نتیجه‌گیری

مسائل و راهکارهای ارائه شده در جدول زیر به صورت خلاصه ارائه شده است. بدیهی است مسائل ساماندهی در بازیابی نیز ملاحظه می‌شود.

جدول ۴-۱: مسائل و راهکارهای ارائه شده در زمینه ساماندهی اطلاعات

مسائل ساماندهی	راهکار
ترجمه به فارسی یا انگلیسی اطلاعات کتابشناختی	۱- بهره‌گیری از زبانشناسان و متخصصان زبان فارسی ۲- افزایش قابلیت نرم‌افزار نمایه‌سازی ۳- تشکیل گروه تصمیم‌ساز متشکل از زبانشناس، متخصص حوزه‌های علمی و نمایه‌ساز
استفاده از معادل‌های متعدد برای یک واژه لاتین	
تعیین معادل فارسی کلیدواژه‌های لاتین	
تعیین اختصار فارسی برای یک واژه لاتین	
نوشتن صورت‌های مخفف لاتین به اشکال مختلف	
تعدد صورت‌های نوشتاری اسامی خاص افراد (پدیدآوران)	نرم‌افزارهای خطایاب زبان فارسی موجود در بازار کمک‌چندانی به حل مسائل ناشی از خطاهای مختلف نگارشی و املائی و نیز درج فرمول‌ها نمی‌کنند. هر چند می‌توانند مقداری از مسائل و خطاها را به نمایه‌ساز نشان دهند.
انواع خطاهای املائی و نگارشی در هر دو سطح فارسی و لاتین	
مسائل مربوط به درج علائم ریاضی و فرمول‌های شیمی در چکیده و عنوان	
مطالب نامربوط در چکیده	

مسائل بازیابی و راهکارهای مطرح شده نیز به طور مختصر در جدول زیر آمده است.

جدول ۴-۲: مسائل و راهکارهای ارائه شده در زمینه بازیابی اطلاعات

مسائل بازیابی	راهکار
صورت‌های متفاوت نوشتاری ۱- ناشی از کاربرد علامتهای مختلف نوشتاری: تشدید، مساد، همزه، تنوین، آوا، نقطه، خط تیره ۲- صورت‌های متفاوت کسره اضافه ۳- صورت‌های مختلف جمع ۴- صورت‌های مختلف اختصارات	بررسی ادبیات موجود نشان می‌دهد که با کاربرد روش‌ها و تکنیک‌های پردازش زبان طبیعی در مدل‌های بازیابی می‌تواند دقت بازیابی را افزایش داد. بنابراین با مشخص کردن روش بازیابی اطلاعات در سامانه گنج و سپس افزودن تکنیک‌های NLP به مدل‌های موجود می‌توان دقت بازیابی در این سامانه را بهبود بخشید.
مسائل معنایی-صرفی-نحوی ۱- کلمات چندمعنا با املاهای یکسان (اما صورت آوایی یکسان یا متفاوت) ۲- صورت‌های متفاوت واژگانی با معنای یکسان یا مرتبط ۳- صورت‌های کوتاه / کامل اسامی یا سازمان‌ها	

فهرست منابع

احترامی، تینا ۱۳۹۵. مستندسازی فرایندهای مرکز اطلاعات و مدارک علمی ایران. پژوهشگاه علوم و فناوری اطلاعات ایران.

ایزدی بروجنی، سارا. ۱۳۹۱. بکارگیری تکنیک‌های پردازش زبان طبیعی برای تطبیق سوال در سیستم‌های پرسش و پاسخ فارسی، پایان‌نامه کارشناسی ارشد، دانشگاه یزد.

ستوده، هاجر و هنرجویان، زهره. ۱۳۹۱. مروری بر دشواریهای زبان فارسی در محیط دیجیتال و تاثیرات آنها بر اثر بخشی پردازش خودکار متن و بازیابی اطلاعات. فصلنامه علمی-پژوهشی کتابداری و اطلاع‌رسانی، دوره ۱۵، شماره ۴ (پیاپی ۶۰).

پارسی اصفهانی، امیر. ۱۳۷۹. ساماندهی کلیدواژه‌های پایگاههای مرکز اطلاعات و مدارک علمی ایران، گزارش طرح پژوهشی، تهران مرکز اطلاعات و مدارک علمی ایران.

رجبی، تقی. و همکاران ۱۳۸۴. بهینه‌سازی اطلاعات علمی شیمی و مهندسی شیمی (سمینارها) در بانک جامع اطلاعات پژوهشگاه اطلاعات و مدارک علمی ایران و طراحی نرم‌افزار بانک ترکیبات شیمیایی، گزارش طرح پژوهش، تهران: پژوهشگاه اطلاعات و مدارک علمی ایران.

رجبی، تقی. و همکاران ۱۳۷۹. بررسی اطلاعات علمی شیمی و مهندسی شیمی ایران در بانک جامع اطلاعات مرکز به منظور بهینه‌سازی آنها (پایان‌نامه‌ها)، گزارش طرح پژوهش، تهران: مرکز اطلاعات و مدارک علمی ایران.

سمائی، سید مهدی. و رجبی، تقی. ۱۳۷۸. مشکلات یکسان‌نویسی اسامی ترکیبهای شیمیایی آلی در زبان فارسی، نخستین هم‌اندیشی مسائل واژه‌گزینی و اصطلاح‌شناسی، تهران: فرهنگستان زبان و ادب فارسی.

هماوندی، هدی، یعقوبی، نوروز و حسینی بهشتی، ملوک السادات ۱۳۹۵. بررسی مشکلات جستجو و بازیابی اطلاعات در پایگاههای اطلاعاتی از جنبه ویژگیهای نگارشی زبان فارسی. فصلنامه علمی-پژوهشی پردازش و مدیریت اطلاعات. پژوهشگاه علوم و فناوری اطلاعات ایران.

Arampatzis, A. T. et al. ۱۹۹۸. Phase-based information retrieval, Journal of Information Processing & Management, Vol.۳۴, No. ۶.

Lancaster, F. Wilfrid, ۱۹۷۹. Information Retrieval Systems: Characteristics, Testing and Evaluation. Second Edition. John Wiley and Sons.

William A. Woods, et al. ۲۰۰۰. Linguistic knowledge can improve information retrieval. Proceedings of the applied NLP conference. Seattle.

Abstract

In this research, firstly the linguistic problems of information organizing and retrieving at Iran Scientific Information System (GANJ) were determined via exploring the database and then some possible solutions were given based on relevant literature. Indexing problems include: the translation of bibliographic information, having different Farsi equivalents for a foreign word, finding Farsi equivalent for foreign keywords, finding Farsi abbreviation for a foreign word, writing a foreign abbreviation in various Farsi forms, spelling and typing errors both in Persian and English versions, problems in typing and inserting mathematical and chemical formula and symbols, and deleting irrelevant information from documents abstracts. For improving information organizing it is recommended to develop more functional softwares for indexing. Besides it is needed to make benefit of linguistic knowledge and Persian language and consult with other experts in different fields of science.

Information retrieval problems are of two sorts: A) different word forms due to using different writing symbols, different genitive and plural forms and different abbreviated forms. B) different forms due to semantic morphological and syntactic issues: polysemy words with same form, different word forms with the same or related meaning and reduced forms for proper names of people or organizations. Information retrieval literature shows that applying techniques of NLP has significant impact on increasing IR systems' precision and recall.

Keywords: Iran Scientific Information System (GANJ), Text retrieving, Text organizing, Text processing, linguistic structure of the text.



Iranian Research Institute for Information Science and Technology

Research Project

**linguistic problems of information
organizing and retrieving at Iran
scientific information system (GANJ)**

By:

Zahra Mahmoodzade

August ۲۰۱۸