

به نام خدا



وزارت علوم، تحقیقات، و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران

(ایرانداک)

### خلاصه گزارش طرح پژوهشی سازمانی

#### عنوان طرح پژوهشی

توسعه و روزآمدسازی اصطلاحنامه‌های ایرانداک

۱۳۹۸/۲/۲۲

تاریخ پایان طرح پژوهشی

۱۳۹۵/۸/۱

تاریخ شروع طرح پژوهشی

ملوک السادات حسینی بهشتی

نام و نام خانوادگی مجری طرح پژوهشی

اسماعیل اکبری، مه‌ری صدیقی، مریم نوروزی اقبالی، سعیده وفایی،

تقی رجبی، زهرا دهرسرای، مهرداد نوروزی اقبالی، امید قیاسوند

نام و نام خانوادگی همکاران طرح پژوهشی

#### درباره طرح پژوهشی

توسعه علم و فناوری، تولید واژه‌های جدید، افزایش انتشارات علمی، واژه‌هایی که توسط کاربران وارد نظام سازماندهی اطلاعات می‌شود، و تغییراتی که در روابط بین مفاهیم و موضوعات ایجاد می‌شود از مجموعه دلایل روزآمدسازی اصطلاحنامه‌ها هستند. در این پژوهش افزایش کمی، ویرایش و توسعه روابط معنایی میان واژه‌های اصطلاحنامه‌های علمی و فنی تدوین شده در پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک) صورت گرفته است. فرایند روزآمدسازی و توسعه شامل مراحل مرور ادبیات و تهیه منابع، طراحی و بررسی نقشه‌های مفهومی، ترجمه واژگان، تعیین روابط معنایی، یکسان‌سازی، واپایش روابط و خطایابی، تولید خروجی‌های اصطلاحنامه‌ای، ویرایش و انتشار است. طی این فرآیند تعداد ۹۰ هزار واژه به مجموعه واژه‌های قبلی افزوده شدند و روابط معنایی میان اصطلاحات توسعه و بهبود یافته است. همچنین در این پروژه وب سرویس اصطلاحنامه‌های روزآمد شده، طراحی و اجرا شده است و هم اکنون از طریق وبگاه ایرانداک قابل دسترسی است. این وبگاه می‌تواند راهنمای مناسبی برای کاربران و نمایه‌سازان در سازماندهی و بازیابی اطلاعات باشد.

#### روش پژوهش

در روزآمد سازی و توسعه اصطلاحنامه‌ها مراحل زیر باید طی شود و یا مورد بازنگری قرار گیرند: شناخت حوزه‌های تحت پوشش، یافتن اصطلاحنامه‌های مناسب لاتین و روزآمد شده، تهیه منابع تخصصی چاپی و الکترونیک، استخراج و جمع‌آوری واژگان و اطلاعات واژگانیفارسی، طراحی نقشه مفهومی و ایجاد شبکه واژگانی اولیه، تفکیک حوزه‌های مختلف علمی و توزیع واژه‌های رده بالا

بین آنها، ترجمه واژه‌های اصطلاح‌نامه موضوعی پایه به زبان فارسی، ترجمه واژه‌های اصطلاح‌نامه‌های موضوعی پایه به زبان انگلیسی، تعیین و توسعه واژه‌های اعم، اخص و مرتبط، تعیین ارجاعات، تعیین اختصارات، تهیه و بررسی یادداشت‌های دامنه، یکدست کردن و هماهنگی صوری واژگان از نظر رسم‌الخط، یکدست نمودن واژگان از نظر جمع و مفرد، یکدست نمودن و هماهنگی واژگان از نظر قواعد صرفی، یکدست نمودن و هماهنگی معنایی واژگان و پیش‌همارایی، واپایش روابط معنایی (سلسله مراتبی و مرتبط)، واپایش ارجاعات، کوته‌نوشته‌ها و روابط دوظفه واژگان، ویرایش خطاهای منطقی، نمایه‌های الفبایی، سلسله مراتبی و نمایه‌گردان، و بازنگری و ویرایش نهایی توسط اساتید متخصص هر حوزه به تفکیک حوزه دانش. با توجه به لزوم انطباق کار روزآمدسازی اصطلاح‌نامه با نیازهای کاربران و نمایه‌سازان، از طرف بخش سازماندهی و تحلیل اطلاعات در ایرانداک، مجموعه‌ای در حدود ۴۰۰۰۰ کلیدواژه پرکاربرد حوزه‌های مختلف علوم پایه و مهندسی بدست آمده از اصطلاحات سامانه ویرایش پایگاه اطلاعات علمی ایران (گنج) در اختیار همکاران طرح قرار گرفته که پس از تفکیک موضوعی و تعیین روابط سلسله مراتبی و انجام فرایندهایی که در بالا ذکر شد، با همکاری متخصصان موضوعی در حوزه‌های مرتبط به اصطلاح‌نامه‌ها افزوده شده است.

### یافته‌های طرح پژوهشی

اصطلاح‌نامه‌های روزآمد شده ایرانداک تنها اصطلاح‌نامه‌های دو زبانه فارسی-انگلیسی در حوزه علوم پایه در زبان فارسی هستند. از آنجا که در تدوین، توسعه و روزآمدسازی آنها از مجموعه واژه‌های پایگاه گنج ایرانداک استفاده شده، بنابراین از مزیت پشتوانه انتشاراتی مناسب برخوردار هستند. ارائه تحت وب این اصطلاح‌نامه‌ها از طریق ویگاه ایرانداک به صورت فرامتن و امکان مشاهده ارتباطات معنایی میان واژه‌ها و اصطلاحات توسط کاربران و نمایه‌سازان، افزایش تعداد واژگان و ارتباطات معنایی دوسویه و استفاده از معادلهای فرهنگستان زبان و ادب فارسی و مراجع معتبر علمی، از ویژگیهای مثبت این پژوهش است. طراحی نقشه‌های مفهومی برای تدوین و روزآمدسازی اصطلاح‌نامه‌های شیمی و علوم زیستی و استفاده از اسامی گیاهان بومی ایران در این اصطلاح‌نامه، همچنین استفاده از اصطلاحات مرتبط با تقسیمات کشوری و جغرافیای بومی ایران در حوزه علوم زمین از امتیازات این پژوهش است. استفاده از نرم‌افزار تزاروی بیلدر که از نرم‌افزارهای برتر در حوزه تدوین و مدیریت اصطلاح‌نامه است و امکان بکارگیری استاندارد ۲۵۹۶۴ (موسسه استاندارد بریتانیا ۱۳۹۱) و ارائه شکلهای مختلف واژگان مانند سلسله مراتبی، الفبایی و گردشی را فراهم می‌کند، از دیگر ویژگیهای مثبت این پژوهش است.

### نتیجه‌گیری و پیشنهادها

استفاده عملیاتی در سازماندهی و بازیابی اطلاعات در مرکز سازماندهی اطلاعات ایرانداک، این اصطلاح‌نامه‌ها را در شرایط واقعی استفاده و بنابراین در معرض ارزشیابی قرار می‌دهد. تعیین میزان جامعیت و مانعیت در جستجوهای انجام شده، میزان اخص بودن واژگان، میزان دقت در واژه‌گزینی، میزان مطابقت با استانداردهای بین‌المللی، نداشتن ارجاعات کور و رها، تناسب واژه‌های ترکیبی نسبت به تکواژه‌ها، تناسب سطح پیش‌همارایی با توجه به گستره اصطلاح‌نامه، سنجش ارتباط و سنجش دسترس‌پذیری از شاخص‌های ارزشیابی هستند که می‌تواند دستمایه پژوهش‌های بعدی در ارزیابی‌های کمی و کیفی این اصطلاح‌نامه‌ها برای محققان باشد. توسعه وب سرویس اصطلاح‌نامه‌ها و ارائه خدمات برخط به کاربران و نمایه‌سازان امکان روزآمدسازی و توسعه در محیط وب، تدوین فرااصطلاح‌نامه و هستان‌شناسی و کاربرد آن در نمایه‌سازی، یادگیری ماشینی و وب‌معنایی از چشم‌اندازهای آینده این پژوهش است.