

به نام خدا



وزارت علوم، تحقیقات، و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

خلاصه گزارش طرح پژوهشی سازمانی

عنوان طرح پژوهشی

ایجاد هستی‌شناسی از اصطلاح‌نامه علوم زیستی ایرانداک

تاریخ شروع طرح پژوهشی	۱۴۰۰/۱۱/۳	تاریخ پایان طرح پژوهشی	۱۴۰۱/۱۱/۳
نام و نام خانوادگی مجری طرح پژوهشی	ملوک‌السادات حسینی بهشتی		
نام و نام خانوادگی همکاران طرح پژوهشی	اسماعیل اکبری و امید قیاسوند		

درباره طرح پژوهشی

هستی‌شناسی یا آنتولوژی علم شناخت و دسته‌بندی مفاهیمی است که موجود هستند یا ممکن است در زمینه‌های مختلف وجود داشته باشند. به طور کلی علم آنتولوژی ریشه در فلسفه دارد و در راستای شناخت مفاهیم و در ادامه نگاه ارسطویی در طبقه‌بندی و دسته‌بندی مفاهیم موجود تلاش می‌کند.

هدف از ایجاد هستی‌شناسی علوم زیستی و به‌کارگیری آن در نظام‌های بازتابی اطلاعات، ساختارمند نمودن اطلاعات و در نهایت بازنمایی دانش به کمک ماشین با استفاده از فناوری‌ها و استانداردهایی است که به وسیله کنسرسیوم وب معنایی در حال تدوین و تکمیل است. به عبارت دیگر در وب معنایی ماشین‌ها با شناسایی داده‌های اسناد و مدارک موجود در فرایند تحلیل و استخراج اطلاعات مورد نظر کاربر مشارکت کرده و در این زمینه هوشمندانه‌تر عمل می‌کنند. در وب معنایی موتورهای جستجو بیش از پیش بر مفهوم داده‌ها و جداسازی آنها مسلط هستند و در این راستا از ابزار دقیق و قوی معنایی مانند آنتولوژی استفاده می‌کنند.

روش پژوهش

پژوهش حاضر از نظر ماهیت بنیادی و کاربردی است و با توجه به این که از دو رویکرد کیفی و کمی استفاده

شده است، روش ترکیبی است. برای ساختن ابزار اولیه پژوهش با استفاده از اصطلاح نامه علوم زیستی که قبلاً در ایرانداک تالیف و تدوین شده بود به منظور ابزار پژوهش استفاده شده است. جامعه آماری این پژوهش ۳۰۰۰ مدخل واژگانی مستخرج از اسناد موجود در حوزه علوم زیستی است که به صورت سلسله مراتبی و به شکل اصطلاح نامه در پایگاه اصطلاح نامه های ایرانداک موجود می باشد. از آنجا که پژوهش حاضر رویکرد تلفیقی دارد، سعی شده است با بررسی ادبیات نظری و منابع مرجع حوزه پژوهش (آنتولوژیهای موجود در سطح بین الملل) درک مناسبی برای شناخت مفاهیم فراهم گردد (روش قیاسی) و از طرف دیگر با بررسی داده ها به شناسایی مفاهیم اصلی و روابط معنایی میان آنها پردازیم (روش اسقرایی). در این پژوهش برای ساخت آنتولوژی مورد نظر از روش "مت آنتولوژی" ابداع گومز-پرز و همکاران (۲۰۰۴) استفاده شده است. و نهایتاً با نرم افزار "پروتزه" شبکه های مفهومی و در نهایت ساختار آنتولوژی ایجاد گردد.

یافته های طرح پژوهشی

در ادامه تبدیل اصطلاح نامه ها به آنتولوژی، طرح ایجاد هستی شناسی علوم زیستی با استفاده از اصطلاح نامه علوم زیستی ایرانداک که در سایت پژوهشگاه در دسترس عموم است، انجام پذیرفت. آنتولوژی حاصل پس از تبدیل با بیش از ۱۸۰ آنتولوژی علوم زیستی بین المللی مورد انطباق قرار گرفت و به توسعه مفاهیم و گسترش روابط معنایی آنها پرداختیم. در نهایت این انطباق سبب استاندارد سازی مفاهیم و برچسب های زبانی و نیز روابط معنایی میان مفاهیم منجر شده و حاصل آن هستی شناسی علوم زیستی دوزبانه بوده است.

نتیجه گیری و پیشنهادها

برای تبدیل اصطلاح نامه به هستی شناسی علوم زیستی از روشی که توسط سورگل و همکاران (۲۰۰۴) ذکر کرده، استفاده شده است. برای ساخت هستی شناسی ابتدا کدنویسی هایی که در هستی شناسی ها قابل تطبیق با نرم افزار پروتزه بودند، مورد مطالعه قرار گرفت. به همین منظور از ابتدا کدهای نوشته شده در هستی شناسی علوم زیستی گسترش داده شد. در این کدهای اولیه ویژگی های کلی فایل هستی شناسی مشخص گردید. در فاز دوم سعی بر استخراج روابط دودویی تعریف شده در هستی شناسی علوم زیستی شد که قبلاً در کشورهای دیگر توسعه یافته اند. برای استخراج روابط دودویی غنی سازی هستی شناسی به مجموعه ۱۸۳ هستی شناسی با بیش از هفت میلیون مفهوم می رسد. پس از بررسی روابط معنایی به فهرستی از روابط دودویی دست یافتیم که کلاس های هر دو طرف آنها را در اصطلاح نامه علوم زیستی ایرانداک یا آنتولوژی ساخته شده وجود دارند.