Designing and Implementing Basic Sciences Ontology Based on Concepts and Relationships of Relevant Thesauri

Molouk Sadat Hosseini Beheshti[1] | Fatemeh Ejei[2]

1. [Corresponding Author] PHD in Linguistic; Assistant Professor; Iranian Research Institute for Information Science and Technology; Tehran, Iran beheshti@irandoc.ac.ir
2. MS in Software Engineering; Sharif University of Technology; Tehran, Iran ejei.fa@gmail.com

Iranian Journal of Information Processing & Management

Iranian Research Institute for Science and Technology
ISSN 2251-8223
eISSN 2251-8231
Indexed in SCOPUS, ISC & LISA
Vol.30 | No.3 | pp: 357-372
Spring 2015

Abstract: Currently, the main portion of knowledge is stored in electronic texts and documents and for transferring that knowledge effectively, we must use proper methods to gather and retrieve relevant information. Ontologies provide means to produce structured documents and use intelligent search instead of keyword search. Ontology defines the common words and concepts used to describe and represent an area of knowledge. However, developing ontologies is a time consuming and labor work. So, many ontology developers try to facilitate and speed up this process by reusing other resources. In fact, thesaurus contains semantic information and hierarchical structure that make it an appropriate resource for ontology construction. Therefore, we determined to use the thesauri previously developed at Iranian Research Institute for Information Science and Technology (IRANDOC) to construct ontology in basic sciences domain. At first, we synchronized common concepts in thesauri before integrating them as a macro thesaurus and removed inconsistencies. To reduce the amount of time and human resources which were needed for synchronizing process, Thesaurus Synchronizer was developed to illustrate differences between matched cases of two thesauri. It provides powerful tools for demonstrating differences and suggestions for each of the existing matters. Thus, domain experts synchronized each two thesaurus semi-automatically. Then we merged thesauri and transformed the data format into ISO 25964 standard. The conceptual model have been designed based on the terms and their relationships in the integrated thesaurus and the concept maps that were designed by domain experts for each of basic sciences (Chemistry, Physics, Biology, Geology and Mathematics). We

used the methodology called METHONTOLOGY in this stage. The main activity in this methodology is conceptualization. It enables the construction of ontologies at the knowledge level. Ultimately, the ontology was generated by formalizing and implementing the conceptual model in OWL, a web ontology language which is recommended by the World Wide Web Consortium. The next step is to refine the relationships to more specific semantic relations. Our domain experts tried to refine some relationships manually but we decided to design an appropriate method for refining the ontology semi-automatically. Also, we need to add more details to our ontology and turn it into heavyweight ontology to get more advantages out of it in formal reasoning. The main aim of this project is to create a standard ontology to be used in information storage and retrieval system.

Keywords: Ontology; Thesaurus Transformation; Ontology Engineering; Knowledge Retrieval

طراحی و پیاده‌سازی هستی‌شناسی علوم پایه بر اساس مفاهیم و روابط موجود در اصطلاح‌نامه‌های مرتبط

ملوک‌السادات حسینی بهشتی[1] | فاطمه اژه‌ای[2]

۱. [پدیدآور رابط] دکتری زبان‌شناسی همگانی؛ استادیار؛ پژوهشگاه علوم و فناوری اطلاعات ایران beheshti@irandoc.ac.ir

۲. کارشناسی ارشد مهندسی کامپیوتر - نرم‌افزار؛ دانش‌آموختهٔ دانشگاه صنعتی شریف ejei.fa@gmail.com

مقالهٔ پژوهشی

دریافت: ۱۳۹۳/۰۲/۱۴
پذیرش: ۱۳۹۳/۰۸/۱۹

دورهٔ ۳۰ شماره ۳
ص.ص. ۶۷۷-۶۹۶

چکیده: هستی‌شناسی، واژگان و مفاهیم مشترک مورد استفاده برای توصیف و ارائهٔ یک حوزه از دانش را معین می‌کند و می‌تواند مبنایی برای سازماندهی و مدیریت دانش در یک حوزهٔ خاص باشد. با این وجود، ساخت هستی‌شناسی کاری دشوار و زمان‌بر است. به‌منظور تسهیل و سرعت‌بخشیدن به این امر، می‌توان منابع دیگر دانش نظیر اصطلاح‌نامه‌ها را مورد استفاده قرار داد. اصطلاح‌نامه به‌لحاظ دارابودن اطلاعات معنایی و ساختار سلسله‌مراتبیِ مفاهیم و نیز نشان‌دادن روابط میان آنها منبع مناسبی برای ساخت هستی‌شناسی به حساب می‌آید. ما با هدف شکل‌گیری هستی‌شناسی در حوزهٔ علوم پایه، اصطلاح‌نامه‌هایی را که پیش از این در پژوهشگاه علوم و فناوری اطلاعات ایران در حوزه‌های مختلف شیمی، فیزیک، زیست‌شناسی، زمین‌شناسی و ریاضی تدوین شده بودند، مبنای ساخت هستی‌شناسی قرار دادیم. برای این منظور، ابتدا مغایرت‌ها و هم‌پوشانی‌های مفاهیم مشترک حوزه‌های مختلف علوم پایه در اصطلاح‌نامه‌های مختلف را برطرف ساختیم تا بتوان از آنها در یک هستی‌شناسی واحد بهره برد. پس از آن کلیهٔ اصطلاح‌نامه‌ها با یکدیگر تلفیق و اصطلاح‌نامهٔ جامع تولید شد. سپس، طراحی مفهومی هستی‌شناسی با استفاده از روش «مت‌آنتولوژی» بر اساس مفاهیم و روابط موجود در اصطلاح‌نامهٔ جامع صورت گرفت. در نهایت، هستی‌شناسی علوم پایه بر اساس طراحیِ صورت‌گرفته پیاده‌سازی شد تا به‌عنوان ابزاری استاندارد جهت ذخیره و بازیابی اطلاعات به کار گرفته شود.

کلیدواژه‌ها: هستی‌شناسی؛ تبدیل اصطلاح‌نامه؛ مهندسی هستی‌شناسی؛ بازیابی دانش

پژوهشنامهٔ پردازش و مدیریت اطلاعات
فصلنامه | علمی پژوهشی
شاپا (چاپی) ۸۲۲۳-۲۲۵۱
شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱
نمایه در LISA، ISC و Scopus
http://jipm.irandoc.ac.ir
پژوهشگاه علوم و فناوری اطلاعات ایران

۱. مقدمه

پیدایش فناوری‌های اطلاعاتی و ارتباطی جدید و دسترسی به محیط وب که حاوی حجم زیادی از اطلاعات است، استفاده از ابزارهای بازنمایی دانش مانند هستی‌شناسی را ضروری ساخته است. وجود هستی‌شناسی‌ها با روابط معنایی دقیق برای بهبود و تقویت سیستم‌های بازیابی و فرایندهای خودکار استدلال ماشینی و همچنین وب معنایی بسیار مهم و لازم است، اما ایجاد آنها کاری پُرهزینه و زمان‌بر است. هستی شناسی در واقع ابزار بازنمون دانش در حوزهٔ سازماندهی دانش و هوش مصنوعی است و تأکید بر کاربرد آن در وب معنایی در دههٔ اخیر مورد توجه دانشمندان قرار گرفته است. هدف از کاربرد هستی‌شناسی، کاهش ابهام مفهومی در حوزه‌ای خاص است. کاهش ابهام، نتیجهٔ مستقیم وجود روابط معنایی روشن و دقیق میان مفاهیم یک علم می‌باشد.

اصطلاح‌نامه[1] یک منبع دانش سازماندهی‌شده توسط متخصصان موضوعی است که حاوی اصطلاحات یک حوزه و روابط معنایی بین این اصطلاحات است. به علاوه، روابط معنایی (اعم/ اخص) که در اصطلاح‌نامه موجود است، به ما این امکان را می‌دهد که ساختار سلسه مراتبی مورد نیاز برای هستی‌شناسی را ایجاد کنیم. از این رو، برآن شدیم تا از اصطلاح‌نامه به‌عنوان منبعی برای ساخت هستی‌شناسی استفاده کنیم. هستی‌شناسی‌ها در حوزهٔ هوش مصنوعی و مهندسی دانش رشد یافته و از پویایی بیشتری نسبت به اصطلاح‌نامه‌ها و فرااصطلاح‌نامه‌ها برخوردارند و می‌توانند در سازماندهی و مدیریت دانش مبنا قرار گیرند. بر این اساس، ارائهٔ الگویی از هستی‌شناسیِ علوم پایه می‌تواند در آینده مبنای ساخت هستی‌شناسی‌های دیگر قرار گرفته و برای بازیابی اطلاعات پایگاه‌های علمی- فنی و همچنین برنامه‌ریزی آموزش و پژوهش مورد استفاده قرار گیرد.

در این مقاله، ابتدا به معرفی هستی‌شناسی و بیان شباهت‌ها و تفاوت‌های آن با اصطلاح‌نامه پرداخته شده است. سپس، کارهای انجام‌شده در زمینهٔ تبدیل اصطلاح‌نامه به هستی‌شناسی مطرح و بررسی شده است. در ادامه، روش مورد استفاده برای تولید هستی‌شناسی و نحوهٔ طراحی و پیاده‌سازی آن بر مبنای این روش توضیح داده شده و در نهایت، پالایش و غنی‌سازی روابط معنایی در هستی‌شناسی مورد توجه قرار گرفته است.

1. thesaurus

۲. مبانی نظری پژوهش

هستی‌شناسی یک گرایش علمی در فلسفه است که به مطالعهٔ آنچه که موجود است و آنچه که باید فرض شود که وجود دارد، به‌منظور دستیابی به یک توصیف متقاعدکننده از واقعیت می‌پردازد. گروبر از پیشگامانِ معرفی فنی هستی‌شناسی در علم کامپیوتر است. هستی‌شناسی بر اساس نظر وی، «ذکر مختصات صریح و روشن از مفاهیم به اشتراک گذاشته‌شده» است (Gruber 1993). تعریف سُوا از هستی‌شناسی به این شرح است: «هستی‌شناسی، دسته‌بندی از موضوعات با عناصر موجود در یک حوزهٔ خاص را بررسی می‌کند و بر اساس آن بررسی، فهرستی از موضوعات را ارائه می‌دهد که به‌تفصیل، انواع موضوعات و روابط منطقی میان آنها را بیان می‌کند و از این‌رو، هستی‌شناسی نامیده می‌شود. از هستی‌شناسی در حوزه‌های مختلفی از جمله هوش مصنوعی، مهندسی نرم‌افزار و اطلاع‌رسانی استفاده شده است» (Sowa 2000).

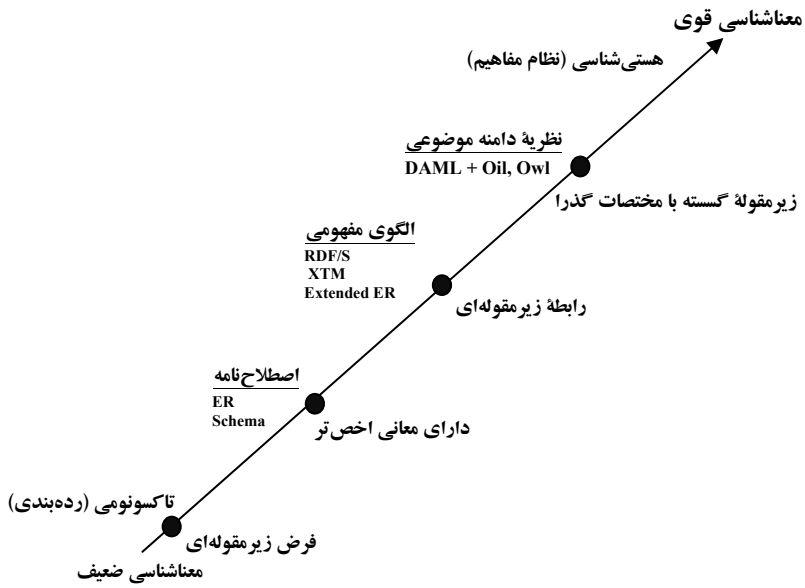
هستی‌شناسی در وب نیز به کار گرفته شده است و در بازیابی اطلاعات متن‌بنیاد مبتنی بر وب به کاربران کمک می‌کند. سورگل با توجه به این نکته، هستی‌شناسی را یک ابزار سازماندهی منابع وب و بازیابی دانش با کمک زبان طبیعی معرفی کرده است. محیط وب، دستیابی گسترده به اطلاعات را فراهم کرده. این محیط اطلاعاتی جدید و تغییر توانایی‌های کاربران و نیازهای اطلاعاتی متفاوت، نیاز به ابزارهای یاری‌دهنده در امر بازیابی و سازماندهی اطلاعات روبه‌رشد را ضرورت بخشیده و یکی از این ابزارها هستی‌شناسی‌ها هستند (Soergel 2004).

اگر بخواهیم میان رده‌بندی[1]، اصطلاح‌نامه و هستی‌شناسی تفکیک قائل شویم، می‌توانیم بگوییم که رده‌بندی‌ها فقط دارای تقسیم‌بندی سلسه‌مراتبی و ساختار پدر و فرزندی مفاهیم هستند و تمایز معنایی کمتری ایجاد می‌کنند؛ در حالی که، اصطلاح‌نامه‌ها علاوه بر آن، اصطلاحات و مترادف‌ها و روابط اعم و اخص و وابسته و ارجاعات را نیز نشان می‌دهند و از این‌رو، نسبت به رده‌بندی‌ها تمایز معنایی دقیق‌تری میان مفاهیم قائل می‌شوند. هستی‌شناسی‌ها در یک مرحلهٔ بالاتر با ارائهٔ الگوی مفاهیم، تقسیم‌بندی معناشناختی دقیق‌تری میان مفاهیم قائل هستند که بر یک نظریهٔ منطقی استوار بوده و از

1. taxonomy

دانش غنی‌تر و پیچیده‌تر با تمایزات معنایی بیشتری برخوردار می‌باشند (Daconta et al. 2003).

در نموداری که در شکل ۱ نشان داده شده، رده‌بندی‌ها با داشتن ساختار تقسیم‌بندی پدر و فرزندی مفاهیم از تمایزات معنایی نسبتاً ضعیف‌تری برخوردارند. در برخی موارد، روابط معنایی بر اساس طبقات و زیرطبقات و در مواردی دیگر بر اساس کل و جزء تعریف می‌شوند. به‌طورکلی، می‌توان در مورد روابط معنایی میان اجزاءِ معناییِ تشکیل‌دهندهٔ رده‌بندی و طبقه‌بندی آن گفت که نسبت به اصطلاح‌نامه‌ها و هستی‌شناسی‌ها از تمایزات معنایی ضعیف‌تری برخوردارند. اما رده‌بندی‌هایی که قوی‌تر و دقیق‌تر هستند، با استفاده از ابزار منطقی، تمایزات معنایی دقیق‌تری میان مفاهیم ایجاد کرده و از این روست که می‌توانند به‌عنوان اسکلت معنایی هستی‌شناسی‌ها مورد استفاده قرار گیرند.



شکل ۱. انواع ابزار رده‌بندی معناشناختی (Daconta et al 2003)

در این نمودار، گرهِ بعدی به اصطلاح‌نامه‌ها اختصاص یافته است. در اصطلاح‌نامه‌ها

بر اساس روابط معنایی میان اصطلاحات (واژه‌ها و عبارات)، ساختاری همانند رده‌بندی شکل گرفته است. بدین‌ترتیب، اصطلاح‌نامه‌ها نوعی رده‌بندی سلسه‌مراتبی و نیز روابط اعم و اخص و وابسته میان اصطلاحات را ارائه می‌دهند. بر اساس استاندارد اصطلاح‌نامهٔ تک‌زبانه[1] تعریف اصطلاح‌نامه عبارت است از: «واژگان کنترل‌شده که بر اساس نظامی شناخته‌شده دسته‌بندی شده است و ساختار آن بر اساس روابط معنایی هم‌ارز، سلسله‌مراتبی، اعم و اخص، و روابط وابستهٔ مفاهیم شکل گرفته است. اهداف اولیهٔ اصطلاح‌نامه، تسهیل بازیابی اسناد و مدارک و دستیابی به انسجام در نمایه‌سازی منابع مکتوب و غیرمکتوب و اسناد و مدارک علمی است. از این روست که اصطلاح‌نامه با داشتن واژگانی کنترل‌شده و استاندارد پشتیبان نمایه‌سازی و بازیابی اطلاعات قرار می‌گیرد تا «جستجوگر اطلاعات» را در گزینش اصطلاحات درست و منطبق با مفاهیم و دستیابی به داده‌های مورد نیازش هدایت کند».

کاربرد یک اصطلاح‌نامه تضمین می‌کند که:

- مفاهیم به‌شیوه‌ای منسجم توصیف شوند؛
- کاربرانِ باتجربه بتوانند به‌آسانی جستجوهای خود را طوری اصلاح نمایند که به نیازهای اطلاعاتی خود سریع‌تر دسترسی یابند؛ و
- کاربران نیازی به آشنایی کامل با اصطلاحات فنی و تخصصی و ریز مسائل آن نداشته باشند.

در شکل ۱، در مرحلهٔ بعد الگوی مفهومی که از بازنمایی معنایی بیشتری برخوردار است، قرار می‌گیرد. وقتی از رده‌بندی‌ها و اصطلاح‌نامه‌ها به‌سوی الگوی مفهومی و نظریهٔ منطقی مفاهیم حرکت می‌کنیم، در واقع به هستی‌شناسی‌ها می‌رسیم. مدل مفهومی الگویی است از مفاهیم نظام‌یافتهٔ یک حوزهٔ موضوعی علوم و دانش که در برخی موارد دامنه نیز نامیده می‌شود و نشانگر مدخل‌های اولیه (اشیاء و پدیده‌های یک دامنه) روابط ساختاری و معنایی میان این مدخل‌های واژگانی و ارزش‌گذاری مختصات معنایی و روابط آنها و در برخی موارد شامل قواعد حاکم بر مدخل‌ها، روابط و ممیزه‌های معنایی آنها نیز می‌شود.

هستی‌شناسی معرف واژگان همگانی و مفاهیم (معانی) آنهاست که در توصیف و

1. ANSI/ NISO

بازنمایی حوزه‌ای از دانش به‌کار می‌رود و از این رو، معانی اصطلاحات را معیارسازی می‌کند. هستی‌شناسی‌ها از سوی متخصصان، پایگاه‌های اطلاعاتی و دیگر سیستم‌های رایانه‌ای به کار گرفته می‌شوند تا اطلاعات موضوعی و تخصصی دامنه (حوزهٔ تخصصی دانش مانند پزشکی، شیمی، فیزیک و ...) را به اشتراک بگذارند. هستی‌شناسی‌ها شامل تعاریف مفاهیم پایهٔ یک دامنهٔ موضوعی و روابط معنایی میان آنهاست که برای رایانه کاربرد داشته باشد و دانش موجود در دامنه را، قابل استفاده و بهره‌برداری مجدد توسط متخصصان و پژوهشگران نماید.

۳ پیشینهٔ پژوهش

فعالیت‌های انجام‌شده در زمینهٔ هستی‌شناسی در زبان فارسی عموماً متمرکز بر وردنت فارسی و واژهستان‌شناسی‌ها بوده است (شمس‌فرد ۱۳۹۱؛ فدایی، قادر، و فیلی ۱۳۹۱). وردنت یا شبکه واژگانی، شامل مجموعه‌هایی از واژه‌های هم‌معنی است که با روابط مشخصی مانند شمول، جزءواژگی، اشتقاق و ... به هم مرتبط شده‌اند و از آن در بررسی مفهومی کلمات توسط رایانه در شاخه‌های مختلف پردازش زبان طبیعی استفاده می‌شود. هدف از این‌گونه از هستی‌شناسی‌ها توصیف ساختارهای معنایی زبان است و با هستی‌شناسی‌های دامنه، که شامل واژگان مربوط به مفاهیم یک حوزه و روابط آنها هستند و هدف آنها مدل‌کردن یک حوزهٔ خاص می‌باشد، متفاوت می‌باشند.

ساخت هستی‌شناسی در یک حوزه از دانش کاری دشوار و زمان‌بر است و یکی از روش‌هایی که به منظور تسهیل و سرعت‌بخشیدن به تولید هستی‌شناسی مورد توجه قرار گرفته، استفاده از منابع دانش سازماندهی‌شده‌ای است که پیش از این توسط متخصصان موضوعی تدوین شده است. یکی از این منابع که ساختار آن تشابه زیادی با ساختار اولیهٔ یک هستی‌شناسی دارد، اصطلاح‌نامه است. برای تبدیل اصطلاح‌نامه به هستی‌شناسی و رفع ابهام از روابط موجود، رویکردی از سوی سورگل و همکاران (۲۰۰۴) ارائه شده است. در این رویکرد، پس از تعریف ساختار هستی‌شناسی و مقداردهی این ساختار بر اساس اطلاعات موجود در اصطلاح‌نامه، فرایند تبدیل روابط صورت می‌گیرد. این فرایند برای دقیق‌ترکردن روابط موجود در هستی‌شناسی توسط متخصص موضوعی انجام می‌گیرد. متخصص موضوعی با شناسایی الگوهای مربوط به رخداد روابط معنایی بین مفاهیم،

قواعدی برای تصحیح معنایی روابط تعریف می‌کند تا بر روی موارد مشابه اعمال شود.

کاوتراکول و همکارانش رویکرد ارائه‌شده توسط سورگل و همکارانش را برای تبدیل اصطلاح‌نامه به هستی‌شناسی مورد استفاده قرار داده و قابلیت اجرای آن را با به‌کارگیری روش‌های یادگیری ماشین، بالا برده‌اند. آنها روشی برای استخراج قواعد به‌صورت خودکار و به‌منظور کشف روابط مبهم به‌خصوص روابط سلسله‌مراتبی و هم‌ارزی و پیشنهاد رابطهٔ مناسب ارائه داده‌اند (Kawtrakul et al. 2005). در مورد روابط وابسته، قواعد توسط متخصصان موضوعی و یا از طریق یادگیری ماشین شناسایی و پیشنهاد می‌شوند و پس از تأیید متخصص بر روی موارد مشابه اعمال می‌شوند. بر اساس همین روش، خسروی و وظیفه‌دوست طرحی برای ایجاد هستی‌شناسی فارسی در حوزهٔ کتابداری و علوم اطلاعات بر مبنای اصطلاح‌نامهٔ اصفا ارائه داده‌اند. در روش ارائه‌شده، روابط بین مفاهیم بر اساس مجموعه‌ای از متون مرتبط و با روش یادگیری به روابط دقیق‌تر هستی‌شناختی تبدیل می‌شوند (Khosravi and Vazifedoost 2008).

در رویکرد دیگری که از سوی هوانگ مطرح شده، از یک هستی‌شناسی پایه در حوزهٔ مرتبط با اصطلاح‌نامه استفاده شده و مفاهیم موجود در اصطلاح‌نامه به این هستی‌شناسی اضافه گردیده و آن را گسترش می‌دهند. در این رویکرد، ابتدا همهٔ مفاهیم بر اساس رده‌بندی حوزهٔ مربوط به اصطلاح نامه طبقه‌بندی می‌شوند. در مرحلهٔ دوم، روابط سلسله‌مراتبی موجود در هستی‌شناسی بررسی و پالایش می‌شوند تا انواع مختلف روابط سلسله‌مراتبی موجود (عام[1]، کل- جزء[2] و مصداقی[3]) از یکدیگر متمایز گردند. در مرحلهٔ آخر، روابط غیرسلسله مراتبی به روابط معنایی در هستی‌شناسی تبدیل می‌شوند. این تبدیل با استفاده از شناسایی الگوهای روابط معنایی که دامنه[4] و برد[5] آنها مشخص است و نیز بر اساس طبقه‌بندی مفاهیم که در مرحلهٔ اول صورت گرفته، انجام می‌شود (Huang et al. 2007).

روش دیگری در زمینهٔ تبدیل اصطلاح‌نامه به هستی‌شناسی از سوی کلِس مطرح شده

1. generic
2. whole-part
3. instance
4. domain
5. range

که شامل هشت گام مختلف است. این گام‌ها از مجموعهٔ اولیه‌ای از گام‌های منظم، که در یک بررسیِ موردیِ مهندسیِ مجدد به‌کار گرفته شده، استخراج شده‌اند. این روش بر مبنای تفاوت‌های ساختاری بین اصطلاح‌نامه و هستی‌شناسی تولید شده و رعایت ترتیب گام‌های آن تضمین می‌کند که در هستی‌شناسیِ به‌دست‌آمده کمترین خطا و ناسازگاری وجود داشته باشد. با این حال، طبق اذعان خود نویسندگان، در پیاده‌سازی این روشِ مهندسیِ مجدد، فقط تعدادی از گام‌ها را می‌توان تا اندازه‌ای به‌صورت خودکار انجام داد در صورتی که، به نظر می‌رسد امکان خودکارکردن گام‌های دیگر در حال حاضر، وجود ندارد و به‌کارگیری آن برای تعداد زیادی از اصطلاحات (یک یا چند اصطلاح‌نامه)، در عمل غیرممکن است (Kless et al. 2012).

٤. روش پژوهش

در پژوهش حاضر قصد بر این بود که از چند اصطلاح‌نامهٔ مختلف که شامل چندین هزار اصطلاح باشد، برای تولید هستی‌شناسی استفاده شود. بنابراین، باید روشی اتخاذ می‌شد که به‌صورت خودکار یا نیمه‌خودکار قابلیت اجرا داشته باشد. برای این منظور، روش مشخصی انتخاب شد و با توجه به چارچوب کلی ارائه‌شده از سوی Villazon-Terrazas et al. (2010) برای مهندسیِ مجددِ منابعِ دانشیِ دیگر و تبدیل آنها به هستی‌شناسی، مسیر کار مشخص گردید.

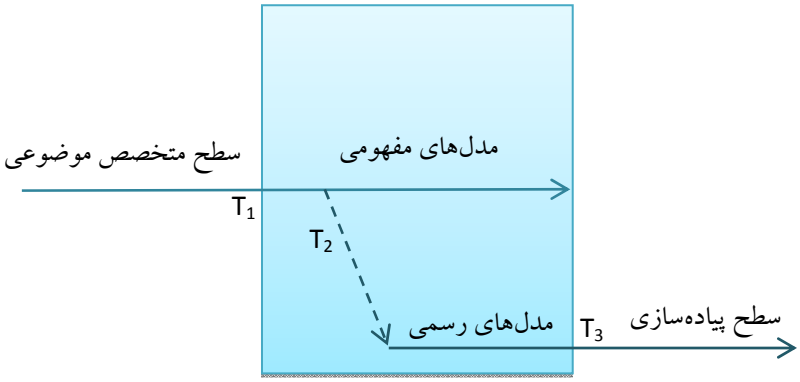
برای ساختِ هستی‌شناسی از روش مت‌آنتولوژی[1] استفاده شده است (Gómez-Pérez 2004). این روش، ساخت هستی‌شناسی‌ها را در سطح مفهوم‌بنیان مقدور می‌سازد. مبنای این روش، فعالیت‌های اصلی مشخص‌شده در فرایند تولید نرم‌افزار و متدولوژی‌های مهندسی دانش است و شامل مشخص‌کردن فرایند تولید هستی‌شناسی و روش‌های اجرای هر یک از فعالیت‌های مدیریت، توسعه و پشتیبانی می‌شود. در این روش، بیشترین میزان اکتساب دانش، مجتمع‌سازی و ارزیابی در مرحلهٔ مفهوم‌سازی صورت می‌گیرد و این مرحله از اهمیت ویژه‌ای برخوردار است.

مرحلهٔ مفهوم‌سازی در روش مت‌آنتولوژی نقش تعیین‌کننده‌ای در ادامهٔ ساخت هستی‌شناسی دارد. هدف این مرحله، سازماندهیِ دانشِ به‌دست‌آمده طی فعالیت کسب

1. metontology

دانش و نمایش آن به صورتی مستقل از الگوهای صوری و زبان‌های پیاده‌سازیِ مورد استفاده برای فرمول‌بندی و پیاده‌سازی هستی‌شناسی است. در نتیجه، فعالیت مفهوم‌سازی، رابطهٔ نیرومندی با فعالیت کسب دانش دارد. بعد از طراحی مدل مفهومی، متدلوژی ایجاب می‌کند که مدل مفهومی به مدل رسمی تبدیل شود. مدل رسمی با استفاده از یک زبان پیاده‌سازی هستی‌شناسی، پیاده‌سازی خواهد شد. در طول این فرایند، هستی‌شناسی به‌تدریج از سطح دانش به سمت سطح پیاده‌سازی حرکت می‌کند و درجهٔ رسمیت مدل دانش کم‌کم افزایش پیدا می‌کند، به‌طوری که برای ماشین قابل فهم خواهد شد.

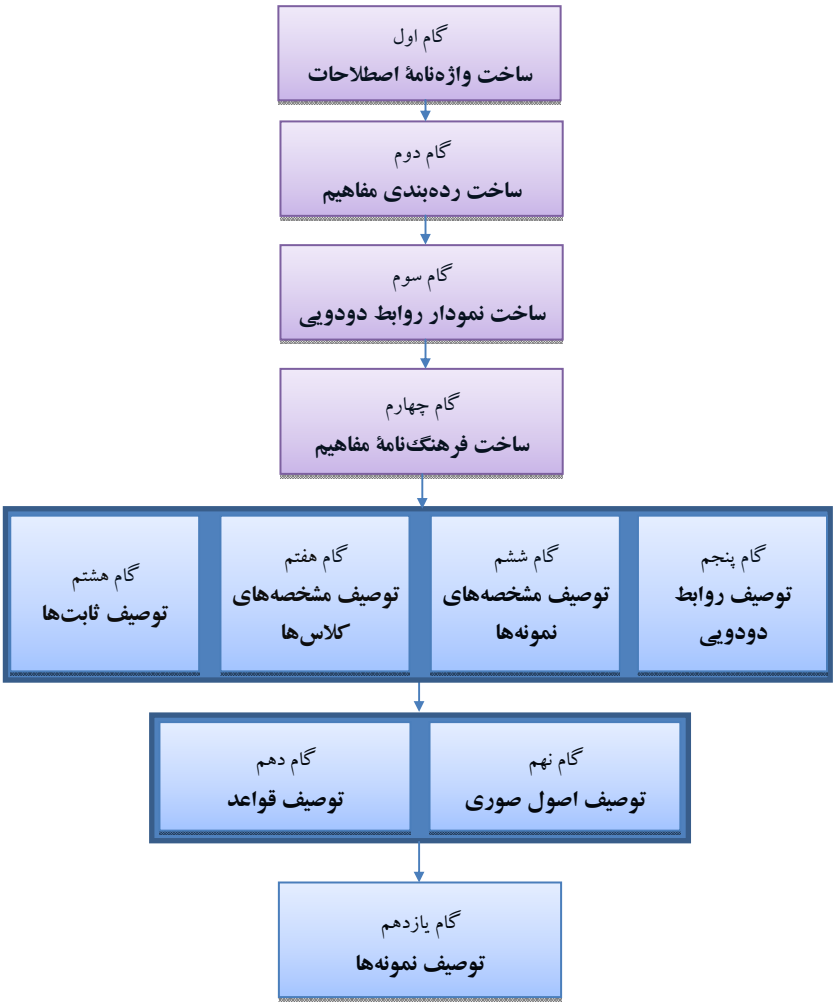
در شکل ۲، مدل فرایند اصلی در مهندسیِ هستی‌شناسی نشان داده شده است. تبدیل T1، نشانگر فرایند مدل‌سازی مفهومی است و می‌توان آن را به‌صورت تبدیل تصوری از دامنه به یک مدل مفهومی که آن تصور را شرح می‌دهد، در نظر گرفت. تبدیل T2 مدل مفهومی را به یک مدل رسمی تبدیل می‌کند. تبدیل T3 مدل رسمی را به مدلی که قابلیت اجرا در رایانه را داشته باشد، تبدیل می‌کند. در این شکل، T1 و T3 به‌صورت خط‌های پیوسته و T2 به‌صورت خط چین نشان داده شده است. این تفاوت به این علت است که قدری از دانش دامنه ممکن است در هنگام تبدیل از مدل مفهومی به مدل رسمی از دست برود.



شکل ۲. مدل فرایند اصلی در ایجاد هستی‌شناسی (Gómez-Pérez 2004)

فعالیت مدل‌سازی مفهومی در مت‌آنتولوژی شامل مجموعه‌ای از وظایف برای

سازماندهی دانش است که در شکل ۳ نشان داده شده است. در این شکل ترتیب پیشنهادی برای ایجاد مؤلفه‌های مختلف هستی‌شناسی در طول فعالیت مدل سازی مفهومی شرح داده شده است. در گام اول، واژه‌نامهٔ اصطلاحات باید ساخته شود. این واژه‌نامه، مجموعه‌ای از اصطلاحاتی را که هستی‌شناسی دربرمی‌گیرد، مشخص می‌کند. علاوه بر این، تعریف هر اصطلاح به زبان طبیعی و همچنین مترادف‌ها و متضادهای هر واژه در واژه‌نامه وجود دارد. در گام دوم، رده‌بندی مفهومی برای طبقه‌بندی مفاهیم ساخته می‌شود. خروجی این گام می‌تواند یک یا چند رده‌بندی باشد که در آنها مفاهیم، دسته‌بندی شده‌اند. گام سوم، مربوط به ساخت نمودارهای روابط دودویی بین مفاهیم برای مشخص‌کردن روابط بین مفاهیم در هستی‌شناسی است. گام چهارم، ساختن فرهنگ‌نامهٔ مفاهیم است که شامل نمونه‌های هر مفهوم، ویژگی‌های نمونه‌ها و کلاس‌ها و روابط بین آنهاست. وقتی فرهنگ‌نامهٔ مفاهیم ساخته شد، هستی‌شناس باید در گام‌های پنجم تا هشتم جزئیات هر یک از روابط دودویی، ویژگی‌های نمونه‌ها، و مشخصه‌های کلاس‌ها و نیز ثابت‌های اصلی دامنه را در فرهنگ‌نامهٔ مفاهیم به‌تفصیل شرح دهد. در گام‌های نهم و دهم اصول صوری و قواعدی که برای بررسی محدودیت و استنباط ارزش مشخصه‌ها به‌کار می‌رود، توضیح داده می‌شود. در نهایت، هستی‌شناس می‌تواند به‌صورت اختیاری اطلاعات بیشتری دربارهٔ نمونه‌ها اضافه کند. رعایت ترتیب گام‌های نشان داده‌شده در این مدل، سازگار و کامل‌بودن دانشِ ارائه‌شده را تضمین می‌کند.



شکل ۳. گام‌های فعالیت مدل‌سازی مفهومی بر طبق مت‌آنتولوژی (Gómez-Pérez 2004)

۵. طراحی و پیاده‌سازی هستی‌شناسی

برای ایجاد هستی‌شناسی علوم پایه، از اصطلاح‌نامه‌های تدوین شده در پژوهشگاه علوم و فناوری اطلاعات ایران که هر یک مرتبط با یکی از حوزه‌های علوم پایه (شیمی،

فیزیک، زیست‌شناسی، زمین‌شناسی و ریاضی) است، استفاده شد. این اصطلاح‌نامه‌ها شامل ده‌ها هزار واژه هستند که در زمان‌های مختلف توسط متخصصان موضوعی تولید شده‌اند. از آنجا که اصطلاحات هر حوزه توسط متخصصان موضوعی متفاوت تعیین شده بود، در ابتدا لازم بود مفاهیم مشترک حوزه‌های مختلف در اصطلاح‌نامه‌ها به‌صورتی یکسان بیان شود تا بتوان از آنها در یک هستی‌شناسی واحد برای کلیهٔ رشته‌های علوم پایه بهره برد. بدین‌منظور نرم‌افزاری برای همسان‌سازی اصطلاحات طراحی شد. این نرم‌افزار ابزاری برای نشان‌دادن موارد اختلاف بین مفاهیم مشترک بین دو اصطلاح‌نامه و پیشنهاد راهکارهای موجود به متخصص موضوعی برای برطرف‌کردن هر مورد است. در بسیاری از موارد، راهکار انتخاب‌شده توسط متخصص موضوعی می‌تواند مستقیماً توسط نرم‌افزار به اجرا درآید تا مفهوم مورد نظر بین دو اصطلاح‌نامه کاملاً یکسان شود، یا کاملاً از یکدیگر متمایز گردد. با کمک این نرم‌افزار، اصطلاحات موجود در همهٔ اصطلاح‌نامه‌های مورد نظر به‌صورت دو به دو توسط متخصصان موضوعی بررسی شد و موارد اختلاف بین مفاهیم مشترک رفع گردید (حسینی بهشتی ۱۳۹۱).

پس از همسان‌سازی اصطلاحات موجود در اصطلاح نامه‌ها بر اساس استاندارد ایزو ۲۵۹۶۴ (ISO25964-1 2011)، کلیهٔ اصطلاح‌نامه‌های حوزه‌های مختلف علوم پایه با یکدیگر تلفیق و اصطلاح‌نامهٔ جامع تولید شد و با انتقال اصطلاح‌نامه از استاندارد ایزو ۵۹۶۴ به استاندارد ایزو ۲۵۹۶۴، اصطلاح‌نامه از حالت واژه‌بنیان به حالت مفهوم‌بنیان تبدیل شد. در استاندارد ایزو ۲۵۹۶۴، هر مفهوم در اصطلاح نامه به‌وسیلهٔ یک اصطلاح مرجح به‌ازای هر زبان و به‌وسیلهٔ هر تعداد اصطلاحات نامرجح نشان داده می‌شود و یادداشت‌های دامنه، حاشیه‌نویسی‌ها و روابط سلسله‌مراتبی و وابسته، به‌جای اصطلاح مرجح، به کل یک مفهوم متصل می‌شوند.

در مرحلهٔ بعد، طراحی هستی‌شناسی بر مبنای اصطلاح‌نامهٔ جامع، با توجه به روش مت‌آنتولوژی، که در بخش قبل توضیح داده شد، طی مراحل زیر صورت گرفت:

۱. ساخت واژه‌نامهٔ مجموعهٔ اصطلاحات

واژه‌نامهٔ اصطلاحات بر مبنای اصطلاحات موجود در اصطلاح‌نامهٔ جامع تشکیل داده شد. این اصطلاحات توسط متخصصان موضوعی هر حوزه به یکی از دو روش قیاسی و یا استقرائی از منابع مختلف استخراج و گزینش شده بود. در روش قیاسی، ابتدا چارچوب

کلی زمینهٔ موضوعی ترسیم می گردد، یعنی مشخص می شود که محدودهٔ آن چیست و چه موضوع‌های کلی را دربرمی گیرد. سپس، هر یک از این موضوع‌ها به موضوع‌های فرعی‌تر تقسیم می گردد و این کار تا مرحلهٔ تعیین اخص‌ترین مفاهیم ادامه پیدا می کند. در این روش، مفاهیم اصطلاحات معمولاً از منابع متعددی نظیر طرح‌های رده‌بندی، اصطلاح‌نامه‌های موجود، دایره‌المعارف‌ها، فرهنگ‌ها و لغت‌نامه‌های تخصصی، نمایه‌نامه‌های ادواری، فهرست مندرجات و متون کتاب‌های درسی پایه و غیره استخراج می‌شوند. در روش استقرایی، ابتدا مجموعهٔ نمونه‌ای از انواع مدارکی که در عمل زیر پوشش نظام ذخیره و بازیابیِ مورد نظر قرار خواهند گرفت، انتخاب می شود. سپس، از نمایه‌سازی این مدارک، مجموعه‌ای از مفاهیم و اصطلاحات فراهم می‌آید.

۲. ساخت رده‌بندی مفاهیم

رده‌بندی مفاهیم بر اساس نظام مفهومی موجود در اصطلاح‌نامه‌ها شکل گرفته است. هر اصطلاحِ مرجح به‌عنوان یک مفهوم یا کلاس در نظر گرفته شده و روابط (اخص/ اعم) به روابط (زیررده/ بالارده) تبدیل شده است؛ بدین‌معنی که اگر یک مفهوم بر اساس اصطلاحی در نظر گرفته شده باشد که نسبت به یک اصطلاح دیگر اخص باشد، این مفهوم زیرردهٔ مفهومی خواهد بود که بر اساس اصطلاح دوم در هستی‌شناسی در نظرگرفته شده است. نمونه‌ای از رده‌بندی مفاهیم در هستی‌شناسی در شکل ۴ نشان داده شده است.

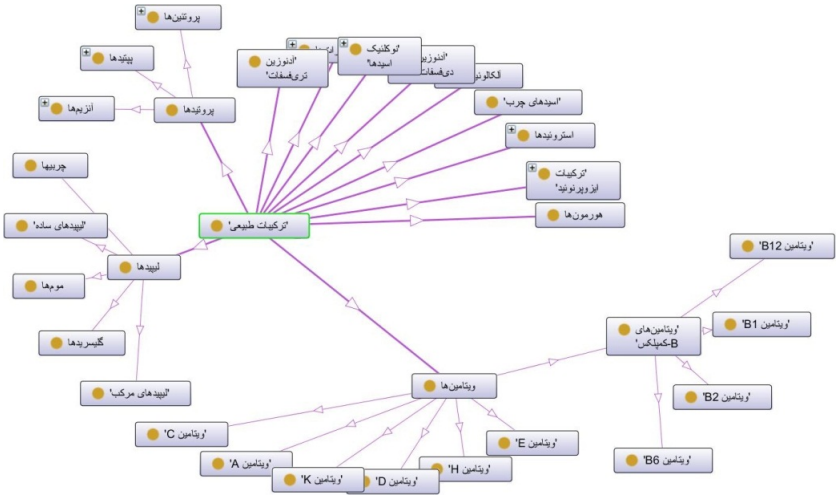
۳. تعریف روابط دودویی میان مفاهیم

روابط دودویی بین مفاهیم بر اساس روابط مفهومی موجود در اصطلاح‌نامه، برای هستی‌شناسی در نظر گرفته شد. علاوه بر روابط سلسه‌مراتبی که در مرحلهٔ قبل بین مفاهیم در نظر گرفته شده بود، بین هر دو مفهومِ وابسته در اصطلاح‌نامه نیز یک رابطهٔ دودویی با برچسب وابسته در هستی‌شناسی در نظر گرفته شد. روابط دیگر نیز به‌صورتی مشابه در هستی‌شناسی لحاظ شده است.

۴. ساخت فرهنگ‌نامهٔ مفاهیم

اصطلاحات نامرجح به‌صورت منفرد تعریف می‌شوند. هر اصطلاح نامرجح نمونه‌ای از کلاس مربوط به اصطلاح مرجح خود در نظر گرفته شده است. روابط دودویی بین نمونه‌ها بر اساس نوع روابط بین اصطلاحات مربوط به آنها در اصطلاح‌نامه مشخص می‌شود. ترجمهٔ اصطلاحات و عبارت‌های اختصاری نیز به‌عنوان ویژگی‌های مفاهیم

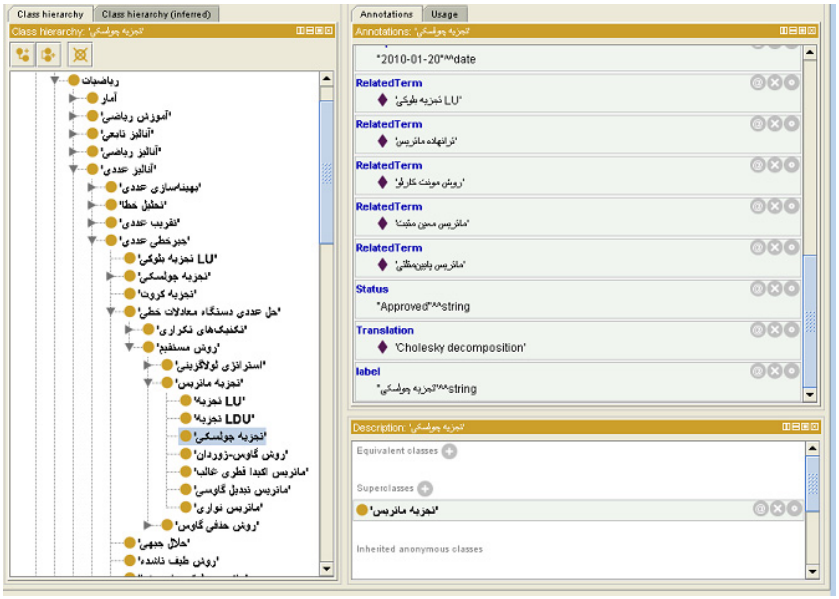
مربوطه در نظر گرفته می‌شوند.



شکل ٤. بخشی از رده‌بندی مفاهیم مشترک بین حوزهٔ زیست‌شناسی و شیمی

پس از طراحی مفهومی هستی‌شناسی و بررسی مفاهیم مشترک و روابط موجود بین آنها با مفاهیم دیگر توسط متخصصان موضوعی، مرحلهٔ پیاده‌سازی صورت گرفت. در این مرحله، طرح مفهومی در قالب زبان «اُدبلیوِال.»[1]، که زبان استاندارد توصیه‌شده توسط کنسرسیوم جهانی وب[2] برای هستی‌شناسی است، پیاده‌سازی شد و هستی‌شناسی علوم پایه تولید گردید. بخشی از هستی‌شناسی تولیدشده، با کمک نرم‌افزار پروتژ[3] در شکل ۵ نشان داده شده است.

1. OWL
2. World Wide Web Consortium (W3C)
3. Protégé



شکل ۵. بخشی از هستی‌شناسی تولیدشده

۵. پالایش و غنی‌سازی روابط معنایی در هستی‌شناسی

پس از تولید هستی‌شناسی، به‌منظور دقیق‌تر کردن روابط معنایی تعریف‌شده در هستی‌شناسی، در مرحلهٔ بعد باید روابط معنایی توسط متخصص موضوعی پالایش شود. این روابط، شامل روابط سلسه مراتبی و روابط وابسته هستند. برای این منظور، ما قصد داریم از روشی مشابه روش ارائه‌شده از سوی Soergel et al. (2004) استفاده کنیم. در این رویکرد، متخصص موضوعی در هستی‌شناسی پیش می‌رود و به‌دنبال الگوهای روابط معنایی بین مفاهیم می‌گردد و بر اساس این الگوها قواعدی را تنظیم می‌کند که قابل اِعمال بر روی موارد مشابه بعدی هستند. این الگوها نوع خاصی از محدودیت می‌باشند. مبنای این محدودیت‌ها روابط موجود در اصطلاح‌نامه از یک طرف و نوع موجودیت‌های مفاهیم مورد بحث از طرف دیگر است. جدول ۱ چند نمونه از محدودیت‌ها را بر مبنای روابط اصطلاح‌نامه نشان می‌دهد.

جدول ۱. نمونه‌هایی از محدودیت‌ها بر مبنای روابط اصطلاح‌نامه (Soergel et al. 2004)

روابط اصطلاح‌نامه	روابط ممکن در هستی‌شناسی
NT/BT	*<hasMember>/<isMemberOf>* *<includesSpecific>/<isa>* *<spatiallyInclude>/<spatiallyIncludeIn>* ...
RT	*<similarTo>* *<growsIn>/<EnvironmentForGrowing>* *<treatmentFor>/<treatedWith>* *<hasMember>/<memberOf>* ...

به‌عنوان اولین گام از روابط معنایی که در جدول ۲ نشان داده شده، برای پالایش محدود روابط معنایی در هستی‌شناسی بهره بردیم. روابط سلسه‌مراتبی در اصطلاح‌نامه معمولاً به یکی از روابط معنایی در دو ردیف اول و روابط وابسته اغلب به یکی از روابط نشان‌داده‌شده در ردیف آخر تبدیل می‌شوند.

روابط نشان‌داده‌شده در جدول ۲ معمولاً نامتقارن هستند. به‌عنوان مثال، اولین رابطه در جدول به‌مفهوم این است که Y همهٔ مشخصات X را داراست و نوع خاصی از X محسوب می‌شود. این، یک رابطهٔ عام استاندارد بین مفاهیم است و برای نشان‌دادن روابط سلسه‌مراتبی کاربرد دارد. در مثال زیر، این رابطه بین دو مفهوم «سلول» و «سلول خون» نشان داده شده است. رابطهٔ بین این دو مفهوم در اصطلاح‌نامه نیز از نوع سلسه‌مراتبی (اعم/اخص) است.

سلول خون *<includeSpecific>* سلول

سلول *<isa>* سلول خون

در اصطلاح‌نامه، سلول خون علاوه بر سلول، دارای یک واژهٔ اعم دیگر یعنی «خون» نیز هست. رابطه بین مفاهیم مربوط به این دو اصطلاح در مرحلهٔ طراحی مفهومی به رابطهٔ *<isa>* تبدیل شده است. در مرحلهٔ غنی‌سازی، رابطهٔ سلسه‌مراتبی این دو مفهوم از نوع کل- جزء تعیین می‌شود و به‌صورت زیر تغییر پیدا می‌کند. این رابطه نشان می‌دهد که سلول خون یکی از مواد تشکیل‌دهندهٔ خون است.

سلول خون *<containsSubstance>* خون

خون *<substanceContainedIn>* سلول خون

نمونهٔ دیگری از روابط سلسه‌مراتبی عام که به رابطهٔ کل- جزء تبدیل شده، در مثال زیر نشان داده شده است. این رابطه نشان می‌دهد که کروموزوم بخشی از هستهٔ سلول است، با این حال، هویت خود را حتی زمانی که بخشی از یک شیء دیگر است، حفظ می‌کند.

کروموزوم <*hasComponent*> هستهٔ سلول

هستهٔ سلول <*componentOf*> کروموزوم

جدول ۲. نمونه‌هایی از روابط بین مفاهیم (Soergel et al 2004)

X و Y مفاهیم هستی‌شناسی هستند
روابط سلسه‌مراتبی (*Isa*)
X <*includesSpecific*> Y / Y <*isa*> X
X <*inheritsTo*> Y / Y <*inheritsFrom*> X
روابط کل- جزء
X <*containsSubstance*> Y / Y <*substanceContainedIn*> X
X <*hasIngredient*> Y / Y <*ingredientOf*> X
X <*madeFrom*> Y / Y <*usedToMake*> X
X <*yieldsPortion*> Y / Y <*portionOf*> X
X <*spatiallyIncludes*> Y / Y <*spatiallyIncludedIn*> X
X <*hasComponent*> Y / Y <*componentOf*> X
X <*includesSubprocess*> Y / Y <*subprocessOf*> X
X <*hasMember*> Y / Y <*isMemberOf*> X
مثال‌هایی از روابط دیگر
X <*causes*> Y / Y <*causedBy*> X
X <*instrumentFor*> Y / Y <*performedByInstrument*> X
X <*processFor*> Y / Y <*usesProcess*> X
X <*beneficialFor*> Y / Y <*benefitsFrom*> X
X <*treatmentFor*> Y / Y <*treatedWith*> X
X <*harmfulFor*> Y / Y <*harmedBy*> X
X <*hasPest*> Y / Y <*afflicts*> X
X <*growsIn*> Y / Y <*growthEnvironmentFor*> X
X <*hasProperty*> Y / Y <*propertyOf*> X
X <*hasSymptom*> Y / Y <*indicates*> X
X <*similarTo*> Y / Y <*similarTo* > X
X < *oppositeTo* > Y / Y < *oppositeTo* > X
X <*hasPhase*> Y / Y <*phaseOf*> X
X <*growsIn*> Y / Y <*EnviromentForGrowing*> X
X <*ingests*> Y / Y < *ingestedBy*> X

در برخی موارد حتی ممکن است روابطی که در اصطلاح‌نامه به‌صورت سلسه‌مراتبی در نظر گرفته شده، در هستی‌شناسی از نوع روابط معنایی غیرسلسه‌مراتبی تعیین شود. به‌عنوان مثال، رابطه بین «گردش خون» و «فشار خون» در اصطلاح‌نامه از نوع اعم/ اخص تعریف شده است، اما این رابطه در هستی‌شناسی به رابطهٔ معنایی دقیق‌تر زیر تبدیل می‌شود.

فشار خون <*hasProperty*> گردش خون

گردش خون <*propertyOf*> فشار خون

مشابه مثال بالا در حوزهٔ فیزیک، رابطه بین «تجسم رد ذره» با «اتاقک حباب» است. اصطلاح مربوط به مفهوم اول یک واژهٔ اعم برای اصطلاح مربوط به مفهوم دوم در اصطلاح‌نامه محسوب می‌شود. در صورتی که در هستی‌شناسی رابطهٔ این دو مفهوم باید کاملاً به‌طور دقیق مشخص شود. از آنجا که «اتاقک حباب» ابزاری برای «تجسم رد ذره» محسوب می‌شود، رابطهٔ این دو مفهوم در هستی‌شناسی به‌صورت زیر اصلاح می‌شود.

تجسم رد ذره < *equipmentFor* > اتاقک حباب

با توجه به مثال‌های بیان‌شده، تبدیل روابط اصطلاح‌نامه به روابط دقیق معنایی در هستی‌شناسی کاری دشوار و زمان‌بر و دارای پیچیدگی‌های خاص خود است. بنابراین، لازم است این کار با استفاده از روش‌های مناسب و نیمه‌خودکار تسهیل گردد تا در حجم زیاد در زمان مناسب قابل انجام گردد.

۶. نتیجه‌گیری

به‌منظور مهندسی مجدد اصطلاح‌نامه و تبدیل آن به هستان‌شناسی علوم پایه پس از همسان‌سازی اصطلاح‌نامه‌ها و رفع موارد اختلاف بین مفاهیم در اصطلاح‌نامه‌های مختلف، اصطلاح‌نامه‌ها از استاندارد ایزو ۵۹۶۴ به استاندارد ایزو ۲۵۹۶۴ منتقل و با یکدیگر تجمیع شدند. با انتقال اصطلاح‌نامه به این استاندارد، اصطلاح‌نامه در واقع از حالت واژه‌بنیان به حالت مفهوم‌بنیان تبدیل شد. این استاندارد از سوی (Kless et al. 2012) نیز قالب پیش‌فرض اصطلاح‌نامهٔ مورد استفاده برای مهندسی مجدد و تبدیل به هستی‌شناسی در نظر گرفته شده است. در مرحله دوم، تبدیل نحوی صورت گرفت. در این مرحله، کلیهٔ

اصطلاح‌نامه‌های علوم پایه در قالب زبان «اُدبلیوِال» ارائه و به هستی‌شناسی تبدیل شدند. از آنجا که ایزو ۲۵۹۶۴ مفهوم‌بنیان است، همهٔ مفاهیم اصطلاح‌نامه به‌عنوان کلاس‌های هستان‌شناسی و اصطلاحات اصطلاح‌نامه به‌عنوان برچسب کلاس‌ها در نظر گرفته شدند. تعاریف، یادداشت‌های دامنه، و دیگر یادداشت‌ها و اطلاعات به توضیحات تبدیل شدند. روابط سلسه‌مراتبی در اصطلاح‌نامه، به‌عنوان روابط سلسه‌مراتبی عام در هستی‌شناسی در نظر گرفته شد و از طریق اصل زیرکلاسی نمایش داده شد. روابط وابسته که نشان می‌دهد رابطه‌ای معنایی بین دو مفهوم وجود دارد، بخشی از توصیف صوری یک مفهوم به‌عنوان یک کلاس در هستی‌شناسی در نظر گرفته شد.

در ادامه، روابط معنایی هستی‌شناسی باید پالایش و غنی‌سازی شوند. تعداد محدودی از روابط توسط متخصص موضوعی به روابط معنایی دقیق‌تر تبدیل شده‌اند. پالایش و غنی‌سازی همهٔ روابط موجود در هستی‌شناسی هدف بعدی خواهد بود. برای این منظور، باید روش‌های مختلف بررسی گردد و روشی مورد استفاده قرار گیرد که قابلیت اجرای خودکار یا نیمه‌خودکار برای پالایش تعداد زیاد روابط موجود در هستی‌شناسی را داشته باشد. هستی‌شناسی با روابط معنایی دقیق‌تر می‌تواند بازنمایی بهتری از دانش یک حوزه را ارائه دهد و توان استدلال پیشرفته‌تری بر اساس این بازنمایی دانش عرضه کند و در نهایت، جستجو و بازیابی اطلاعات را بهبود بخشد.

۷. یادداشت

مقاله حاضر برگرفته از طرح پژوهشی «طراحی هستی‌شناسی علوم پایه و پایگاه اطلاعات مفهوم‌بنیان» است که در پژوهشگاه علوم و فناوری اطلاعات ایران انجام شده است.

فهرست منابع

حسینی بهشتی، ملوک‌السادات. ۱۳۹۱. ایجاد هستی‌شناسی واژگانی از اصطلاح‌نامه‌های فارسی. در *مجموعه مقالات هشتمین همایش زبان‌شناسی ایران*. ۱۴۸-۱۶۴. تهران. انتشارات دانشگاه علامه طباطبایی.

شمس‌فرد. مهرنوش. ۱۳۹۱. ساخت نیمه‌خودکار وردنت فارسی: از واژه تا واژهستان‌شناسی. در *مجموعه مقالات دومین هم‌اندیشی زبان‌شناسی رایانشی*. ۱۰۱-۱۲۲. تهران. انتشارات اهورا. انجمن زبان‌شناسی ایران.

فدایی، مرضیه، حمیدرضا قادر، و هشام فیلی. ۱۳۹۱. تولید خودکار وردنت فارسی با استفاده از روش شناختی و کاربرد آن در بازیابی اطلاعات. در *مجموعه مقالات دومین هم‌اندیشی زبان‌شناسی رایانشی*. ۱۳۱-۱۵۲. تهران. انتشارات اهورا. انجمن زبان‌شناسی ایران.

Daconta, Michael C., Leo J. Obrst, and Kevin T. Smith. 2003. The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management. Wiley Publishing, Inc.

Gómez-Pérez, Asunción, Mariano Fernández-López, and Oscar Corcho. 2004. Ontological Engineering; 2nd ed. London: Springer-Verlag.

Gruber, T. R. 1993. A translation approach to portable ontology specifications. *Knowledge acquisition* 5 (2): 199-220.

Huang, Jin-Xia, Ji-Ae Shin, and Key-Sun Choi. 2007. *Enriching core ontology with domain thesaurus through concept and relation classification*. Proc. OntoLex, ISWC.

ISO 25964-1. 2011. *Information and documentation -- Thesauri and interoperability with other vocabularies -- Part 1: Thesauri for information retrieval.* International Organization for Standardization, International Standard ISO 25964-1, Aug. 2011.

Kawtrakul, Asanee, Aurawan Imsombut, Aree Thunkijjanukit, Dagobert Soergel, Anita Liang, Margherita Sini, Gudrun Johannsen and Johannes Keizer. 2005. Automatic term relationship cleaning and refinement for AGROVOC. *In Proceedings of EFITA/WCCA و Vila Real, Portugal*, 247-260.

Khosravi, F. and A. Vazifedoost. 2008. Creating a Persian Ontology through Thesaurus Reengineering for Organizing the Digital Library of the National Library of Iran. *Faslname Ketab* 70: 19-36.

Kless, D., L. Jansen, J. Lindenthal, and J. Wiebensohn. 2012. A method for re-engineering a thesaurus into an ontology. In: *Proceedings of the 7th International Conference, Graz, Austria*, *Volume* 239, 133-146.

Soergel, Dagobert و Boris Lauser و Anita Liang, Frehiwot Fisseha, Johannes Keizer, and Stephen Katz. 2004 Reengineering Thesauri for New Applications: the AGROVOC Example. *Journal of Digital Information* 4 (4).

Sowa, J. F. 2000. *Knowledge representation: logical, philosophical, and computational foundations*, Pacific Grove, CA: Brooks Cole.

_____. 2010. The role of logic and ontology in language and reasoning. In *Theory and Applications of Ontology: Philosophical Perspectives*, Netherlands: Springer 231-263.

Villazón-Terrazas, B., M. C. Suárez-Figueroa, and A. Gómez-Pérez. 2010. A pattern-based method for re-engineering non-ontological resources into ontologies. *International Journal on Semantic Web and Information Systems* 6 (4): 27-63.