

الله أكبر صلوات



وزارت علوم، تحقیقات و فناوری

پژوهشگاه علوم و فناوری اطلاعات ایران (ایراندک)

طرح پژوهشی

ایجاد هستی‌شناسی از اصطلاح‌نامه علوم زیستی ایراندک

مجری

ملوک‌السادات حسینی بهشتی

همکاران

اسماعیل اکبری

امید قیاسوند

بهمن ۱۴۰۱

حقوق: پژوهشگاه علوم و فناوری اطلاعات ایران

طرح پژوهشی ایجاد هستی‌شناسی از اصطلاح‌نامه علوم زیستی ایرانداک پیرو قرارداد شماره ۳۳/۹۴۴۰ مورخ ۱۴۰۰/۱۱/۴ میان پژوهشگاه علوم و فناوری اطلاعات ایران (کارفرما) و خانم ملوک‌السادات حسینی به‌شتی اجرا شده است. گزارش حاضر یک جلد از مستندات این طرح است.

این گزارش و تمامی حقوق مادی آن بر اساس قانون حمایت حقوق مؤلفان و مصنفان و هنرمندان، مصوب ۱۳۴۸ و اصلاحیه‌های بعدی آن و همچنین آیین‌نامه‌های اجرایی این قانون متعلق به پژوهشگاه علوم و فناوری اطلاعات ایران و هرگونه استفاده از تمامی یا پاره‌ای از آن، شامل: نقل قول، تکثیر، انتشار، کاربرد نتایج، تکمیل و مانند آنها به‌صورت 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 بر عهده مجری طرح پژوهشی است.

صفحه تأييد پژوهشگاه

ایجاد هستی‌شناسی از اصطلاح‌نامه علوم زیستی ایرانداک

مدیر طرح پژوهشی: ملوک‌السادات حسینی بهشتی، پژوهشگاه علوم و فناوری اطلاعات ایران (ایرانداک)

نشانی: تهران، خیابان انقلاب، چهارراه فلسطین، شماره ۱۰۹۰، طبقه سوم، اتاق ۳۱۸

صندوق پستی: ۱۳۱۵۷۷۳۳۱۴ تلفن: ۰۲۱-۶۶۴۹۴۹۸۰ (داخلی ۳۲۰)، دورنگار: ۰۲۱-۶۶۴۶۲۲۵۴

وبگاه: <http://www.irandoc.ac.ir>

رایانامه: beheshti@irandoc.ac.ir

همکاران اصلی:

اسماعیل اکبری، استاد فیزیولوژی پزشکی، دانشگاه علوم پزشکی مازندران

امید قیاسوند، محقق کتابخانه ملی اقتصاد آلمان

چکیده

هستی‌شناسی یا آنتولوژی علم شناخت و دسته‌بندی مفاهیمی است که موجود هستند یا ممکن است در زمینه‌های مختلف وجود داشته باشند. به‌طور 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.

در آنتولوژی سعی می‌شود مفاهیم یک رشته تخصصی دسته‌بندی و کلاس‌های مفاهیم، ارتباطات معنایی بین مفاهیم، ویژگی‌های مفاهیم، توابع مفاهیم، رویدادهایی که برای یک مفهوم می‌تواند رخ دهد، قوانین جاری الگوهای معنایی و محدودیت‌های مفاهیم و نیز شرایط قیاس بین مفهوم‌ها مشخص شود. به این ترتیب با نگاه به مفاهیم مورد نیاز امکان ایجاد هستی‌شناسی مورد نیاز در نظام‌های بازیابی اطلاعات و جستجوی دقیق‌تر و سریع‌تر امکان‌پذیر می‌باشد. برای ساختن آنتولوژی از یک فرهنگ واژگان توسعه‌یافته مانند تزاروس استفاده می‌کنیم. این فرهنگ واژگان نه تنها حاوی معنی اصطلاحات است، بلکه اصطلاحاتی را که به نوعی با مفهوم مورد نظر ارتباط منطقی دارند مشخص کرده است و به این ترتیب همه مفاهیم مورد نظر در بحث خاص خود اعم از اشیاء یا اشخاص و ویژگی‌ها را از فرهنگ واژگان یا اصطلاح‌نامه استخراج و آنتولوژی شکل می‌گیرد. در نهایت با ساخت آنتولوژی و تدوین الگوهای منطقی به زبان مشترک و قابل فهم برای انسان و ماشین دست خواهیم یافت.

هدف از ایجاد هستی‌شناسی زیست‌شناسی و به‌کارگیری آن در نظام بازیابی اطلاعات، ساختارمند نمودن اطلاعات و در نهایت بازنمایی دانش به کمک ماشین با استفاده از فناوری‌ها و استانداردهایی است که به وسیله کنسرسیوم وب جهان‌گستر 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 پیش از پیش بر مفهوم داده‌ها و جداسازی آن‌ها مسلط هستند و در این راستا از ابزار دقیق و قوی معنایی مانند آنتولوژی استفاده می‌کنند.

در ادامه تبدیل اصطلاح‌نامه‌ها به آنتولوژی، طرح ایجاد هستی‌شناسی علوم زیستی با استفاده از اصطلاح‌نامه علوم زیستی ایرانداک و با انطباق با General Ontology for Biology که در سایت Bioportal Ontologies در دسترس است، به توسعه مفاهیم و گسترش روابط معنایی آن می‌پردازیم. بدین منظور نرم‌افزار پروتزه، که دسترس‌پذیر و قابل استفاده برای زبان فارسی است به کار گرفته و در نهایت هستی‌شناسی علوم زیستی با ۵۰۰۰ مدخل حاصل خواهد شد.

کلیدواژه‌ها: اصطلاح‌نامه، هستی‌شناسی، علوم زیستی، نظام بازیابی اطلاعات، وب معنایی

فهرست مطالب

عنوان	شماره صفحه
چکیده.....الف	
فصل نخست: مقدمه.....	۱
۱-۱. مقدمه و بیان مسئله.....	۲
۲-۱. روش پژوهش.....	۳
۳-۱. پیشینه پژوهش.....	۵
فصل دوم: مبانی نظری.....	۸
۱-۲. مقدمه.....	۹
۲-۲. تعریف هستی‌شناسی.....	۱۰
۳-۲. نظریه‌های مبتنی بر هستی‌شناسی.....	۱۳
۴-۲. هستی‌شناسی در مقایسه با ابزارهای دیگر سازماندهی و بازنمون دانش.....	۱۵
۵-۲. طبقه‌بندی هستی‌شناسی ها.....	۲۰
۶-۲. ساختار هستی‌شناسی.....	۲۳
۷-۲. معیارهای طراحی هستی‌شناسی.....	۲۳
۸-۲. زبان‌های هستی‌شناسی.....	۲۵
۹-۲. زبان‌های بازنمایی هستان‌شناسی.....	۲۷
۱۰-۲. مقایسه زبان‌های هستی‌شناسی.....	۳۳
۱۱-۲. ابزارهای ساخت هستی‌شناسی.....	۳۵
۱۲-۲. رویکردهای طراحی هستی‌شناسی.....	۳۷
۱۳-۲. روش‌شناسی‌های مهندسی هستی‌شناسی.....	۳۹
۱۴-۲. مقایسه روش‌شناسی‌های مهندسی هستی‌شناسی.....	۴۷
۱۵-۲. روش‌های طراحی هستی‌شناسی.....	۴۸

۱۶-۲. روش های ارزیابی هستی شناسی	۵۰
۱۷-۲. محدودیت های هستی شناسی موجود در حوزه علوم زیستی	۵۲
۱۸-۲. مقایسه اصطلاح نامه و هستی شناسی	۵۶
۱۹-۲. هستی شناسی در بازیابی اطلاعات	۵۷
فصل سوم: روش تبدیل اصطلاح نامه علوم زیستی به هستی شناسی	۶۳
۱-۳. مقدمه	۶۴
۲-۳. نحوه تبدیل اصطلاح نامه به هستی شناسی	۶۴
۳-۳. یافتن روابط دودویی در هستی شناسی	۶۷
۴-۳. نتایج و تحلیل نهایی	۷۰
۵-۳. پیوست ها	۷۰
فصل چهارم: تحلیل روابط معنایی مفاهیم هستی شناسی	۷۵
۱-۴. تحلیل روابط معنایی مفاهیم	۷۶
فصل پنجم: جمع بندی و نتیجه گیری	۸۹
۱-۵. جمع بندی	۹۰
۲-۵. نتیجه گیری	۹۷
فهرست منابع و مآخذ	۹۹

فهرست جدول‌ها

عنوان	شماره صفحه
جدول ۱-۲. مقایسه ابزارهای مختلف سازماندهی و بازنمون دانش	۱۶
جدول ۲-۲. رویکردهای مختلف در طبقه‌بندی انواع هستی‌شناسی	۲۲
جدول ۳-۲. مقایسه زبان‌های بازنمایی هستان‌شناسی‌ها	۳۴
جدول ۴-۲. هستی‌شناسی‌های موجود در حوزه علوم زیستی	۵۲

فهرست شکل‌ها

عنوان	شماره صفحه
شکل ۱-۲. مثلث معنا.....	۱۸
شکل ۲-۲. اصطلاح‌نامه در مقابل هستی‌شناسی.....	۱۹
شکل ۳-۲. طیف هستی‌شناسی.....	۲۰
شکل ۴-۲. زبان‌های هستی‌شناسی.....	۲۷
شکل ۵-۲. زیرشاخه‌های اُ. دیلو. ال.....	۳۳
شکل ۶-۲. نمایش رویکردهای سه‌گانه طراحی هستی‌شناسی.....	۳۸
شکل ۷-۲. مراحل ساخت هستی‌شناسی.....	۴۵
شکل ۱-۳. برشی از نقشه مفهومی سلول و زیرکلاس‌ها.....	۷۲
شکل ۲-۳. برشی از نقشه مفهومی هورمون‌ها.....	۷۳
شکل ۳-۳. برشی از نقشه مفهومی اسیدهای آلی.....	۷۴
شکل ۴-۳. برشی از نقشه مفهومی فیزیولوژی گیاهی.....	۷۴

فصل نخست

مقدمه

۱-۱. مقدمه و بیان مسئله

در فرایند تدوین اصطلاح‌نامه، ایجاد یک ساختار سلسله‌مراتبی از مفاهیم، اولین ضرورت و ایجاد و تعیین نوع روابط بین مفاهیم دومین ضرورت است که در رویکرد ساختار مبنای، فراروی تدوین کنندگان اصطلاح‌نامه قرار دارد. اما پیچیدگی‌های شرایط نمایه‌سازی و نیز دیدگاه‌های متفاوت کاربران نظام‌های بازیابی اطلاعات، ضرورت سومی را با رویکرد کاربرمدار به عرصه سازماندهی و بازیابی اطلاعات وارد ساخته است. نمایه‌سازان در بسیاری از موارد هنگام استفاده از اصطلاح‌نامه در سازماندهی اطلاعات، با توصیفگرهای متعدد از چند مقوله موضوعی متفاوت مواجه می‌شوند که برای نمایاندن یک موضوع اطلاعاتی، مناسب به نظر می‌رسد. ابهام در استفاده از توصیفگرهای مناسب‌تر هنگام سازماندهی از یک سو و دیدگاه‌های متفاوت کاربران نهایی نظام‌های بازیابی اطلاعات از سوی دیگر، گسترش روابط معنایی در اصطلاح‌نامه را ضروری می‌سازد. شناخت حدود و ثغور مفاهیم و مصداق‌های آن‌ها با ابهام کمتری روبه‌رو می‌گردد. در چنین شرایطی اختصاص توصیفگرها با قطعیت بیشتر هنگام سازماندهی و امکان تعیین درخواست جستجوی اطلاعات با استفاده از مختصات معنایی، سطح اثربخشی نظام بازیابی اطلاعات را از دید کاربران افزایش خواهد داد. همچنین وارد ساختن مختصات معنایی به ساختار مفاهیم در اصطلاح‌نامه آن را به سطح مدل مفهومی سلسله‌مراتب هستی‌شناسی‌های تخصصی نزدیک می‌کند. سازماندهی دانش با هدف دسته‌بندی مفاهیم مشابه به جهان‌های ممکن نیز می‌نگرد. هر پدیده‌ای به منزله یک مفهوم - در هستی، دارای ارتباطی معنایی با سایر مفاهیم است، که مجموعه این پدیده‌ها به همراه ارتباط معنایی شان با یکدیگر، شبکه مفهومی از هستی را، در جهانی ممکن به تصویر می‌کشد. جدای از آنکه دسته‌بندی و مدل‌سازی مفهومی را حاصل کشفیات علمی یا ساخت انسانی بدانیم، تعیین مفاهیم و روابط معنایی در بافت زبانی تعیین شده صورت می‌پذیرد. هستی‌شناسی واژگانی در واقع ابزار بازنمون دانش در حوزه سازماندهی و هوش مصنوعی است و تأکید بر کاربرد آن در وب معنایی در دهه اخیر مورد توجه دانشمندان این رشته است. هدف از کاربرد هستی‌شناسی واژگانی، کاهش ابهام مفهومی در حوزه‌ای خاص است. کاهش

ابهام، نتیجه مستقیم وجود روابط معنایی روشن و دقیق میان مفاهیم یک علم است. روابط معنایی مجاز در هستی‌شناسی‌ها بر اساس نرم‌افزارهای گوناگون و به‌ویژه با زبان هستی‌شناسی OWL 2 و جدیدترین استاندارد زبانی کنسرسیوم جهانی ۲۵۹۶۴ در بازنمون دانش امکان‌پذیر می‌باشد. هستی‌شناسی‌های واژگانی در حوزه پردازش زبان طبیعی NLP 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.

اصطلاح‌نامه یک منبع دانش سازماندهی شده توسط متخصصان موضوعی است. روابط سلسله‌مراتبی (اعم/اخص) که در اصطلاح‌نامه موجود است می‌تواند برای ایجاد ساختار سلسله‌مراتبی مورد نیاز برای هستی‌شناسی مورد استفاده قرار گیرد. روابط وابسته در اصطلاح‌نامه، وجود روابط معنایی بین مفاهیم را نشان می‌دهد. در نتیجه تبدیل اصطلاح‌نامه از حالت واژه-بنیان به مفهوم-بنیان با انتقال آن به استاندارد ۲۵۹۶۴ خواهد بود. در مرحله تبدیل نحوی، انتقال مفاهیم اصطلاح‌نامه به مفاهیم هستی‌شناسی صورت می‌گیرد. اصطلاحات مرجع و نامرجح، برچسب مفاهیم می‌گیرد و روابط اعم/و اخص به روابط سلسله‌مراتبی تبدیل می‌شود. روابط وابسته بخشی از توصیف صوری مفاهیم خواهد بود و نهایتاً پیاده‌سازی خودکار هستی‌شناسی بر اساس تبدیل نحوی در قالب زبان OWL صورت می‌گیرد.

هدف اصلی: ایجاد هستی‌شناسی با استفاده از اصطلاح‌نامه علوم زیستی ایرانداک.

هدف‌های فرعی: ۱. غنی‌سازی روابط معنایی آنتولوژی علوم زیستی؛

۲. کاربرد هستی‌شناسی به منظور ابزاری قدرتمند در بازیابی اطلاعات.

۲-۱. روش پژوهش

پژوهش حاضر از نظر ماهیت کاربردی است و با توجه به این که از دو رویکرد کیفی و کمی استفاده شده است، روش ترکیبی (آمیخته) است. برای ساختن ابزار اولیه پژوهش با استفاده از منابع کتابخانه‌ای و اسنادی، مقوله‌های موضوعات اصلی و فرعی شناسایی می‌شود. در این روش گردآوری داده‌ها در بخش کیفی و کمی دارای اهمیت است. در بخش کیفی از روش تحلیل موضوع به منظور ساخت ابزار پژوهش استفاده شده است. البته برای تعیین تعداد و فراوانی کمی مفاهیم اصلی و فرعی از روش کمی نیز استفاده شده است.

روش کتابخانه‌ای و اسنادی برای گردآوری اطلاعات و شناسایی مقوله‌ها و تعیین نمایه‌های موضوعی به کار گرفته می‌شود. در مرحله بعد تحلیل محتوا، کار با داده‌ها و سازماندهی آنها، تجزیه

داده‌ها به واحدهای قابل کنترل، ترکیب آنها و جستجوی الگوها خواهد بود. این روش قابلیت این را دارد که مفاهیم به‌دست آمده را بسط ساختاری دهیم و طبقه‌بندی نموده و در قالب نقشه مفهومی و مدل ارائه دهیم.

جامعه آماری این پژوهش تعداد ۳۰۰۰ مدخل واژگانی مستخرج از اسناد موجود حوزه علوم زیست است که به صورت سلسله‌مراتبی و به شکل اصطلاح‌نامه در پایگاه اصطلاح‌نامه‌های ایرانداک موجود می‌باشد. بنابراین از سویی برای استخراج مفاهیم و مقوله‌بندی آنها و برقراری روابط اصطلاح‌نامه‌ای از داده‌های اسنادی و از سوی دیگر برای تعیین میزان فراوانی مفاهیم اصلی و فرعی، از این روش استفاده می‌شود. و در نهایت با مقوله‌بندی پربسامدترین مقوله‌های نمایه‌های موضوعی به تدوین شبکه معنایی مفاهیم رشته علوم زیستی پرداخته می‌شود. به‌طور کلی فرایند مقوله‌بندی در تحلیل محتوا وابسته به اهداف پژوهش می‌تواند قیاسی یا استقرایی صورت پذیرد. از آنجا که پژوهش حاضر رویکرد تلفیقی دارد، سعی می‌کنیم با بررسی ادبیات نظری و منابع مرجع حوزه پژوهشی (آنتولوژی‌های موجود در سطح بین‌المللی) درک مناسبی برای شناخت مفاهیم فراهم گردد (روش قیاسی) و از طرف دیگر با بررسی داده‌ها به شناسایی مفاهیم اصلی و روابط معنایی میان آنها پردازیم (روش استقرایی). در این پژوهش برای ساخت آنتولوژی مورد نظر از روش «مت آنتولوژی» ابداع گومز-پرز و همکاران (۲۰۰۴) استفاده می‌شود. و نهایتاً با نرم‌افزار «پروتژه» Protege شبکه‌های مفهومی و در نهایت ساختار آنتولوژی را ایجاد می‌نماییم.

گام‌های فعالیت مدل‌سازی مفهومی طبق روش «مت آنتولوژی».

۱. ساخت واژه‌نامه اصطلاحات؛

۲. ساخت رده‌بندی مفاهیم؛

۳. ساخت نمودار روابط دودویی؛

۴. ساخت فرهنگ‌نامه مفاهیم؛

۵. توصیف روابط دودویی؛

۶. توصیف مشخصه‌های نمونه‌ها؛

۷. توصیف مشخصه‌های کلاس؛

۸. توصیف ثابت‌ها؛

۹. توصیف اصول صوری؛

۱۰. توصیف قواعد؛

۱۱. توصیف نمونه‌ها.

کاربرد

آنتولوژی علوم زیستی به عنوان ابزار جستجوی و بازیابی اطلاعات مورد نیاز مراکز اطلاع رسانی و کتابخانه‌های تخصصی محسوب می‌شود. همچنین پیاده سازی و به کارگیری آنتولوژی علوم زیستی حاصل از تزاروس ایرانداک می‌تواند ابزاری برای بازیابی دقیق تر اطلاعات پایگاه گنج باشد و اصل مهم مانعیت و جامعیت را در سازماندهی دانش این پایگاه تحقق بخشد.

گام‌های مورد نظر این پژوهش عبارتند از:

۱. مرحله مطالعاتی و مرور اسناد در حوزه ایجاد هستی‌شناسی علوم زیستی؛
۲. ایجاد هستی‌شناسی علوم زیستی، تبدیل اصطلاح‌نامه و گسترش روابط معنایی آن.

۱-۳. پیشینه پژوهش

«پایگاه‌های اطلاعات واژگانی همانند شبکه واژگان (شامل واژگان، اصطلاح‌نامه‌ها و فرهنگ‌های لغت) است که به قاعده‌مند نمودن معنی‌شناسی واژگانی (یعنی معنی واژه‌های یک یا چند زبان طبیعی) می‌پردازد. این معانی بر اساس شبکه مفاهیم و مبتنی بر فهم مشترک اکثریت جامعه علمی، و دارای تعاریف خاص، و دربرگیرنده معانی مختلف متداول و کاربردهای خاص و ویژه و ترکیبات گوناگون واژه‌های مبتنی بر «پیش‌نمونه‌های» و نیز مفاهیم پایه علوم است.

وردنت^۱ که توسط میلر^۲ و همکاران وی در سال‌های ۱۹۹۰ تا ۱۹۹۵ به عنوان نمونه یک پایگاه و سیع اطلاعات واژگانی در زبان انگلیسی در دانشگاه پرینستون و بر اساس نظریه‌های روان‌شناسی طراحی و پیاده‌سازی شد روان‌شناسی زبان حوزه میان رشته‌ای تحقیقاتی است که متمرکز بر علوم شناختی و توانش زبان‌شناختی است (فلبوم و میلر ۱۹۹۰)^۳. وردنت یک نوع سازماندهی اطلاعات واژگانی و روابط منطقی میان مفاهیم است که از مشهورترین و پرکاربردترین فرهنگ‌های معنابنیان رایانه‌ای به شمار می‌رود. جورج میلر، روان‌شناس و محقق در زمینه واژگان که خود بنیان‌گذار وردنت است، آغاز شکل‌گیری وردنت را سال ۱۹۸۵ میلادی می‌داند.

وردنت همانند یک فرهنگ معنابنیان واژه‌ها را در قالب شبکه‌ای از روابط مفهومی و به گونه‌ای که واژگان ذهنی گویشوران یک زبان سازماندهی شده است، ارائه می‌دهد. روابط هم‌معنایی، شمول معنایی، رابطه جزء و کل و یا جزء واژگی، و تضاد یا تقابل معنایی در این شبکه از اصلی‌ترین روابط هستند. وردنت مبتنی بر معنی‌شناسی مؤلفه‌ای است که واژه‌ها را به مؤلفه‌های کوچک‌تر و زیرمقوله‌ای تجزیه می‌کند و روابط میان آن‌ها را ایجاد می‌نماید. واژه‌ها به مقولات دستوری اسم، فعل، صفت و

¹ Word Net

² Miller

³ Fellbaum & Miller 1990

قید تقسیم و با معیار معنای مشترک به مجموعه‌هایی^۱ دسته‌بندی شده‌اند و هر واژه بر اساس معنای واحد در یکی از مجموعه‌ها جای می‌گیرد و چنانچه معانی متفاوتی داشته باشد در مجموعه‌های واژگانی جداگانه قرار می‌گیرد در نتیجه روابط متفاوتی با دیگر واژه‌های آن مجموعه ایجاد می‌نماید. این مجموعه‌ها از واژه‌های « مترادف »^۲ یا به عبارتی هم‌معنا تشکیل می‌شوند که در بافت زبانی، قابلیت جانشینی دارند. هم‌معنایی به دلیل ایجاد مجموعه‌ها، رابطه معنایی اصلی در وردنت می‌باشد. رابطه معنایی دیگری که در این شبکه اهمیت ویژه‌ای دارد، رابطه شمول معنایی است که سبب ایجاد ساختاری سلسله‌مراتبی یا درختی در مجموعه‌ها و به‌ویژه در میان اسامی است. علاوه بر این، دو رابطه معنایی جزء و کل یا جزء واژگی^۳ در ایجاد روابط میان اسامی و رابطه تضاد یا تقابل معنایی در ایجاد شبکه واژگانی به کار گرفته شده است. هم‌اکنون زبان‌های زیادی علاوه بر انگلیسی مانند فرانسه، آلمانی، ایتالیایی، اسپانیایی و چینی دارای وردنت هستند.

در زمینه کاربرد هستی‌شناسی در ترجمه ماشینی و «وردنت فارسی» در دانشکده برق و کامپیوتر دانشگاه شهید بهشتی، طرح‌های پژوهشی در دست تکمیل و اجراست، من جمله «وردنت فارسی» یا "FARSNET"^۴ توسط دکتر مهرانوش شمس‌فرد در دانشکده مهندسی و علوم کامپیوتر دانشگاه شهید بهشتی و با حمایت مرکز تحقیقات مخابرات در دست اجرا می‌باشد و در واقع «شبکه واژگانی» بر اساس واژه‌نامه‌های متعدد من جمله فرهنگ آریان‌پور و با استفاده از واژه‌های عمومی زبان فارسی انجام شده و در حال تکمیل است.

نمونه دیگر از این دست طرح‌های پژوهشی در دانشگاه شریف توسط دکتر ابوالحسنی^۵ و همکاران در دست اجرا می‌باشد که هدف نهایی آن ترجمه ماشینی و مهم‌تر از آن ایجاد شبکه معنایی است که مراحمی از پروژه پشت سر گذاشته شده و در حال تکمیل است.

طرح دیگری که دکتر هشام فیلی در دست مطالعه و اجرا داشته و آن نیز بیشتر بر نحو زبان فارسی متمرکز بوده و هدف نهایی آن نیز «ترجمه ماشینی»^۶ دو زبان انگلیسی به فارسی و بالعکس می‌باشد. علاوه بر این در مورد ترجمه ماشینی زبان فارسی به انگلیسی، پروژه کلانی در آزمایشگاه تحقیقاتی کامپیوتر، دانشگاه مکزیکوسیتی (نیومکزیک) توسط فردی به نام «کارین مگر دو میان»^۷ و همکاران در دست اجرا و تکمیل است که از سال ۲۰۰۰ گزارش آن در دست است که بر روی text و

^۱ Synsets

^۲ Synonyms

^۳ meronymy

^۴ <http://farsnet.nlp.sbu.ac.ir>

^۵ موسس و سرپرست آزمایشگاه وب هوشمند (Web Intelligence Lab)

^۶ NLP-IIS@ut at SemEval-2021

^۷ Karine Megerdooian

متن ماشین‌خوان استوار است و هدف آن بیشتر ترجمه اخبار از فارسی به انگلیسی است و پستوانه آن واژه‌نامه عمومی است شامل ۵۰ هزار واژه که به تدریج از سال ۱۹۹۷ تا ۲۰۰۰ از متون خبری فارسی استخراج و از نظر ساخت‌واژی و نحوی^۱ در حجم کلان پیکره زبانی دو زبانه برچسب‌گذاری^۲ شده و تحت عنوان پروژه شیراز^۳ در حال تکمیل می‌باشد.

در زمینه کاربرد هستی‌شناسی‌ها در بازیابی اطلاعات علمی-فنی مقالات فروانی من جمله داورپناه (۱۳۸۴)، اسدی و جمالی مهمویی (۲۰۰۴)، شریف (۱۳۸۴)، نشاط (۱۳۸۵)، زندی‌روان (۱۳۸۶) و غیره در زمینه کاربرد آنتولوژی‌ها در فرایند سازماندهی اطلاعات و مدیریت نگاشته شده است.

در زمینه «پردازش زبان طبیعی و بازیابی اطلاعات» نیز مقالاتی از سوی دکتر جعفر مهرداد (مهرداد ۱۳۸۴) تدوین یافته و طرح پژوهشی «ریشه‌ساز فارسی» در کتابخانه منطقه‌ای شیراز در دست اجرا می‌باشد.

تفاوت میان طرح‌های موجود و در دست اجرا با آنچه هم‌اکنون از سوی پیشنهاددهنده طرح حاضر ارائه شده آن است که اکثر این واژه‌نامه‌ها و پیکره‌های زبانی^۴ شامل مجموعه‌ای از واژه‌های عمومی زبان فارسی می‌باشد. حال آنکه طرح حاضر، اصطلاح‌نامه‌های تخصصی تدوین یافته در پژوهشگاه به‌ویژه در حوزه «علوم زیستی» را به‌عنوان «پیکره زبانی» این تحقیق قرار داده که همه آن‌ها مجموعه‌ای از اصطلاحات تخصصی این رشته‌ها می‌باشد و مهم‌تر آن که دارای روابط سلسله‌مراتبی منطقی و در نتیجه قابل تجزیه و تحلیل معنایی و تعیین مختصات معنایی^۵ علاوه بر تجزیه و تحلیل ساخت‌واژی (صرفی) می‌باشد. این ویژگی سبب می‌گردد که بتوان با کاهش هم‌پوشانی این اصطلاحات در حوزه‌های تخصصی به سرعت به «شبکه واژگان علمی» دست یافت و «هستی‌شناسی‌های واژگانی» مرتبط با این حوزه‌های تخصص که نسبت به اصطلاح‌نامه‌ها از انعطاف بیشتری برخوردار بوده و قابل گسترش و روزآمدسازی بیشتری است طراحی گردد در نهایت با ارائه «نقشه مفهومی»^۶ هر یک از رشته‌ها، بازیابی اطلاعات علمی-فنی به صورت خودکار و ماشینی درآید. علاوه بر این می‌توان از این نقشه مفهومی در بازنگری برنامه‌ریزی درسی رشته‌های مرتبط و نیز ترسیم نقشه جامع علمی کشور استفاده کرد و نیز ایجاد «پایگاه دانش واژگانی» مانند جمع‌آوری مجموعه‌ای از واژه‌نامه‌ها، اصطلاح‌نامه‌ها و شبکه واژگانی می‌تواند به عنوان ابزار کلیدی در شکل‌گیری «هستی‌شناسی‌های واژگانی» بر اساس محتوای متون علمی، به‌ویژه مرتبط با پایگاه اطلاعات علمی و فنی باشد.

¹ Parsing

² tagged

³ Shiraz Machine Translation Project (Proceedings of the NAACL HLT 2010)

⁴ Corpora

⁵ semantic features

⁶ Concept map

فصل دوم

مبانی نظری

۲-۱. مقدمه

هستی‌شناسی اغلب مفاهیمی را دربرمی‌گیرد که برای شناخت انسان از دنیای پیرامون وی، بنیادی و اساسی باشند و دانش «پیش‌نمونه»^۱ها را بازنمایی کنند. این که ساختار هستی‌شناسی بر اساس اعم و اخص یا سطوح بالا و پایین طراحی شوند، بیشتر به اهداف و حوزه مورد نظر بستگی دارد تا حل مشکلات مدیریت واژگان، زیرا تفات معنادار و واقعی میان این دو نوع هستی‌شناسی وجود دارد. هستی‌شناسی بر اساس سلسله‌مراتب و مختصات معنایی، سعی دارد حوزه معنایی خاصی را که مبتنی بر مفاهیم دقیقاً تعریف شده است را به دست آورد. به عنوان مثال در علوم پایه (ریاضی، فیزیک، شیمی، زیست‌شناسی و زمین‌شناسی) دانش مربوطه به راحتی فرمول‌بندی می‌شود، زیرا در همه این حوزه‌ها نظام مفاهیم حاکم است که می‌توان آن را به واسطه اصطلاحات علمی، با تعریف دقیق بیان نمود. در چنین مواردی سعی می‌شود اشیاء در عالم مقال و نیز به‌طور محض و انتزاعی و با الگوهای آرمانی و ساده‌شده‌ای از پدیده‌های واقعی در جهان ارائه شوند (گومز - پرز و همکاران ۲۰۰۶).

هستی‌شناسی‌ها به‌طور گسترده در مهندسی دانش، هوش مصنوعی و علوم رایانه، مدیریت دانش، پردازش زبان طبیعی^۲، تجارت الکترونیک، بازیابی اطلاعات، طراحی پایگاه‌های اطلاعاتی، آموزش و برنامه‌ریزی درسی، و نیز در حوزه جدید «شبکه معنایی»^۳ کاربرد دارند. ابداع شبکه معنایی مرحله‌ای دیگر در تکامل هستی‌شناسی‌ها را به وجود آورده است. بنا بر نظریه برنرزی و همکاران ۲۰۰۱^۴ شبکه معنایی، توسعه شبکه جهانی اطلاعات کنونی است که در آن، اطلاعات بر مبنای عناصر معنایی تعریف شده و با همکاری انسان و رایانه، ذخیره و اشاعه اطلاعات انجام می‌شود. در این همکاری، هستی‌شناسی‌ها به عنوان ابزار کلیدی توسعه شبکه معنایی در بازنمایی یک دامنه موضوعی و به عنوان الگوهای فرایندی استدلالی تبادل اطلاعات نقش مؤثری ایفا می‌کنند. دیویس و همکاران ۲۰۰۷^۵ نگاه

¹ Prototypes

² NLP (Natural Language Processing)

³ Semantic Web

⁴ Berners-Lee 2001

⁵ Davis et al. 2007

جدیدی به آینده «شبکه جهانی اطلاعات»^۱ دارند که در آن، تمامی دانش موجود به صورت قاعده‌مند رمزگذاری می‌شود تا عوامل هوشمند بتوانند به طریقی خود سامان به همه انسان‌ها خدمات اطلاعاتی ارائه دهند. این کار مستلزم تعریف رسمی و قاعده‌مند هستی‌شناسی‌ها مبتنی بر «وب» است تا بیان‌کننده دانش قابل فهم انسان و نیز عوامل هوشمند باشد و ارزش افزوده بر دانش و توسعه آن به صورت «فرا داده‌ها»^۲ خواهد بود، که در حال حاضر تعریف شده و قابل کاربرد در مجموعه‌های محدودی از تقاضاهاست. پیش‌بینی می‌شود که در دهه‌های آتی، ساختار دانش مورد درخواست بیشتر به صورت رسمی و به شکل نظام‌های هستی‌شناسی مبتنی بر «وب» و ارائه آن‌ها به زبان‌های گوناگون، زمینه فعالی برای تحقیقات و مطالعات بیشتر باشد که حاصل آن تعداد قابل توجهی طرح‌های تحقیقاتی و تصویب معیارها و استانداردهاست که در قالب «آر.دی.اف»^۳ و «اکس‌ام‌ال/آر.دی.اف»^۴ و جدیدترین آن (OWL) به عنوان زبان‌های ارائه دانش به کار می‌روند. شبکه معنایی ارتباطات مفاهیم گوناگون را مقدور می‌سازد و برای عملکرد دقیق‌تر رایانه، نیاز به ساختارمند نمودن اطلاعات و نیز اعمال قواعد معنایی دارد.

۲-۲. تعریف هستی‌شناسی

واژه هستی‌شناسی ریشه در فلسفه دارد. تعاریف متعددی از هستی‌شناسی بر اساس دیدگاه‌های مختلف ارائه شده است. گروبر (۱۹۹۳) به عنوان یکی از پیشگامان مطالعات هستی‌شناسی در تعریفی که مقبول بسیاری از آثار این حوزه است، آن را در بافت هوش مصنوعی به عنوان «ذکر خصوصیات شفاف از مفاهیم به اشتراک گذاشته شده» تعریف می‌کند. وی معتقد است اصطلاح «هستی‌شناسی» از فلسفه به امانت گرفته شده در جایی که یک هستی‌شناسی بیان نظام‌مندی از وجود است و برای سیستم‌های دانش‌بنیان آنچه موجود است دقیقاً همان است که می‌تواند بازنمون شود. او همچنین به نقش هستی‌شناسی‌ها به عنوان واژگان مشترک توصیف فرا داده‌ها، اشتراک‌گذاری و استفاده مجدد دانش در میان سامانه‌ها هوشمند اشاره می‌کند (گروبر ۱۹۹۳).

همچنین گوارینو و گیارتا^۵ (۱۹۹۵) هفت تعریف از هستی‌شناسی را برای شفاف‌سازی معنی آن تحلیل نمودند. آن‌ها هستی‌شناسی را به عنوان «نظریه‌ای منطقی که مفهوم‌سازی جزئی و آشکار را ارائه می‌کند» تعریف کردند. در جای دیگر گوارینو هستی‌شناسی را مطالعه آنچه موجود است و آنچه باید فرض شود که وجود دارد؛ به منظور نائل آمدن به یک توصیف متقاعدکننده از واقعیت تعریف

¹ World Wide Web

² Metadata

³ RDF (Resource Description Framework)

⁴ XML/ RDF

⁵ Guarino and Giaretta

می‌کند.

گوارینو، هستی‌شناسی را سلسله‌مراتبی از مفاهیم مرتبط توسط رده‌ها توصیف می‌کند؛ از نظر وی در موارد پیچیده‌تر اصول مناسب به‌منظور توضیح دیگر روابط بین مفاهیم و همچنین برای تحمیل تفسیرهای مورد نظر اضافه می‌شوند. از سوی دیگر در کاربرد متداول در حوزه هوش مصنوعی، هستی‌شناسی اشاره به ایجاد یک مهندسی دارد که متشکل از واژگان خاص به کار رفته برای توصیف یک واقعیت خاص به‌علاوه گروهی از فرضیات است که متوجه معانی مورد نظر واژگان فرهنگ لغت (فهرست واژگان) هستند (گوارینو ۱۹۹۹، ۴).

از دیدگاه سُوا^۱ نیز به‌عنوان مبتکر نقشه‌های مفهومی، موضوع هستی‌شناسی، مطالعه طبقات اشیاء موجود یا ممکن‌الوجود در بعضی حوزه‌ها است که مفاهیم و منطق بین آن‌ها را با جزئیات بازنمون می‌کند. آنسکفلد نیز در تعریفی مشابه سُوا هستی‌شناسی را مجموعه‌ای از مفاهیم (موجودیت‌ها، نگرش‌ها و فرایندها)، تعاریف آن‌ها و روابط داخلی بین آن‌ها می‌داند (سوا ۲۰۰۰).

همچنین به عقیده برنرزی و همکاران^۲ (۲۰۰۱) برای پژوهشگران رایانه، هوش مصنوعی و وب، هستی‌شناسی پیشینه یا فایلی است که به صورت رسمی روابط میان اصطلاحات را تعریف می‌کند و قوانین استنباطی آن قدرت بیشتری در مقایسه با تاکسونومی‌ها فراهم می‌کند. مدچ^۳ (۲۰۰۲) هم هستی‌شناسی را ساختاری رسمی^۴ از کلاس‌ها، روابط و اصول می‌داند که اجازه ایجاد روابط چندگانه از جمله همنامی^۵ و مترادف^۶ را بین موجودیت‌های واژگانی می‌دهد.

همان‌طور که از تعاریف قابل درک است واژه هستی‌شناسی در بافت‌های مختلف به صورت‌های متفاوتی تعریف شده است. رسی و همکاران^۷ (۲۰۱۱) در مطالعه خود به این نکته اشاره دارند که پژوهش‌های مربوط به هوش مصنوعی در ابتدا واژه «هستی‌شناسی» را از فلسفه قرض گرفتند، سپس واژگان آن در بسیاری دامنه‌های علمی گسترش یافت و هم‌اکنون هستی‌شناسی‌ها در شمار زیادی از حوزه‌ها و در سطحی وسیع استفاده می‌شوند. اما به صورت کلی اغلب تعاریف مورد اشاره، دارای همه یا برخی از مفاهیم زیر هستند که تشریح آن‌ها به حصول درک کاملی از تعریف هستی‌شناسی کمک می‌کند:

رسمی بودن:

به این نکته اشاره دارد که هستی‌شناسی باید ماشین‌خوان بوده و با سیستم‌های هوش مصنوعی

^۱ Sowa

^۲ Berners-Lee, Hendler and Lissila

^۳ Meadche

^۴ Formal

^۵ Homonymy

^۶ Synonymy

^۷ Roussey, Pinet, Kang, and Corcho

پردازش شده باشد. بر همین اساس لازم است هستی‌شناسی با یک زبان رسمی تعریف شود این بازنمون سبب می‌شود که هستی‌شناسی قابلیت به‌کارگیری و تفسیر در سیستم‌های اطلاعاتی را داشته باشد.

مختصات یا ذکر خصوصیات^۱:

به معنی تشخیص و تعیین ویژگی‌های نحوی زبان برای حصول معیارهای خاصی از جمله دقت، ابهام‌زدایی، پیوستگی و یکپارچگی و جامعیت توصیفات و مفاهیم است چرا که هستی‌شناسی در واقع ابزار ارتباطی بر مبنای دانش تسهیم شده و مورد اجماع کاربران است. این مرحله شامل تعیین مختصات معنایی است زیرا هستی‌شناسی ویژگی‌های مفاهیم را به زبان طبیعی ارائه می‌کند و فارغ از سطح پیچیدگی و شفافیت در تعیین ویژگی‌های معنایی وابسته به زبانی است که بتواند فرمول‌بندی هستی‌شناسی را ارائه کند (نقل در ساوسا ۲۰۱۳).

تسهیم شده^۲:

کمک به ایجاد درک مشترک از یک حوزه موضوعی و مفاهیم مرتبط با آن از اهداف و کارکردهای اصلی هستی‌شناسی‌ها است. عبارت تسهیم شده که در تعاریف بسیاری از آن یاد شده به این نکته اشاره دارد که هستی‌شناسی دانش مورد توافق و اجماع را بازنمون می‌کند، چرا که این ابزار حاصل فعالیت گروهی از افراد است تا این که بر مبنای نظرات شخصی افراد شکل گیرد.

مفهوم‌سازی^۳:

به نقل از گومز پرز و همکاران (۲۰۰۴) مفهوم‌سازی را جهان‌گفتمان در همه حالات ممکن امور برای حوزه خاصی از اشیاء تعریف می‌کند که هدف هستی‌شناسانه نظام‌های اطلاعاتی است. در این مرحله دانش دامنه موضوعی به صورت مدلی مفهومی که شامل واژگان آن حوزه است شکل می‌گیرد، سپس مفاهیم توصیف می‌شوند. زبان و سیله منحصر به فرد برای ارائه مفاهیم است و از آنجا که عامه مردم به طور معمول از زبان طبیعی استفاده می‌کنند برای مفیدتر بودن مفهوم‌سازی لازم است که این مرحله میان انسان‌ها به اشتراک گذاشته شود و در مورد آن توافق حاصل شود. به عبارت دیگر مفهوم‌سازی که در زبان طبیعی بازنمایی می‌شود نه یک فرایند واحد بلکه مجموعه‌ای از فرایندها و اطلاعات حاصله تعریف شده از سوی جمعی از گویشوران زبان است.

¹ Specification

² Shared

³ Conceptualization

۲-۳. نظریه‌های مبتنی بر هستی‌شناسی

از منظر علم اطلاعات، مبانی نظری هستی‌شناسی از سویی ریشه در فلسفه و از سوی دیگر ریشه در نظریه‌های طبقه‌بندی دارد. نظریه هستی‌شناسی به فلاسفه یونان باستان مانند پارمنیدس^۱ (۵ و ۴ ق.م) و ارسطو (۴ ق.م) بازمی‌گردد. اصطلاح هستی‌شناسی ریشه در فلسفه دارد که با مطالعه «وجود» و «هستی» مرتبط است. در فلسفه از هستی‌شناسی به عنوان نظریه‌ای در مورد ماهیت وجود یاد می‌شود. ارسطو به عنوان پیشگام در بحث هستی‌شناسی و نویسنده کتاب متافیزیک^۲ (اثری که از آن با نام هستی‌شناسی نیز یاد می‌شود) در مطالعه خود در مورد وجود، حالات مختلفی از وجود و هستی را برای پایه‌گذاری سیستمی از مقوله‌بندی‌ها (ماده، کیفیت، کمیت، رابطه، عمل، حالت، مکان و زمان) مشخص کرد تا بتوان هر آنچه را که به پدیده یا شیئی دیگری در جهان اشاره دارد، طبقه‌بندی نمود. برای مثال وقتی می‌گوییم: «این رایانه روی میز قرار دارد» نسبت به وقتی که می‌گوییم: «این رایانه خاکستری است» حالت متفاوتی از بودن را فرض می‌کنیم. بیان نخستین در مقوله مکان طبقه‌بندی می‌شود در حالی که بیان دوم در مقوله کیفیت جای می‌گیرد. این نوع طبقه‌بندی ارائه شده توسط ارسطو تا قرن هجدهم به‌طور گسترده‌ای پذیرفته شده بود.

در دوران جدید امانوئل کانت^۳ (۱۷۲۴-۱۸۰۴ م) معتقد بود که وجود پدیده‌ها و موجودات نه تنها با خود آن فی نفسه بلکه با هم‌افزایی و همکاری ادراک افراد و فهم آن‌ها نیز مشخص می‌گردد. از نظر وی سؤال کلیدی این است که «ذهن ما از چه ساختارهایی برای تسخیر و درک واقعیت اسفاده می‌کند؟» پاسخ این سؤال، کانت را به سمت مقوله‌بندی سوق داد. چارچوب ارائه شده وی در چهار کلاس سازماندهی شده که هر یک از آن‌ها الگویی چندوجهی را ارائه می‌نماید: کمیت (یگانگی (وحدت)، تعدد، کلیت)، کیفیت (واقعیت، خنثی بودن، محدودیت)، رابطه (توارث، رابطه علت و معلولی یا علیت، اجتماع) و چگونگی (احتمال، وجود، ضرورت). بنابراین یک «هستی‌شناسی»^۴ طبقه‌بندی از مقولات است در حالی که هستی‌شناسی^۵ شاخه‌ای از فلسفه است. در علوم رایانه و علم اطلاعات، هستی‌شناسی اصطلاحی فنی است که دلالت بر مصنوعی دارد که به منظوری خاص، یعنی ایجاد توانایی به منظور مدلسازی دانش در یک دامنه خاص واقعی یا تخیلی طراحی می‌شود (کرچو و همکاران ۲۰۰۷، ص ۴۴-۷۰).

¹ Parmenides

² MetaPhysics

³ Emmanuel Kant

⁴ An ontology

⁵ Ontology

همچنین «مرئولوژی»^۱ به‌عنوان یک نظریه رسمی از روابط جزء-کل^۲ اساساً از طریق آثار هوسرل^۳ (۱۸۵۹-۱۹۳۸) و استانیسلاو لسنووسکی^۴ (۱۹۳۹-۱۸۸۶) مطرح شد. هوسرل نظریه مرئولوژی را به‌عنوان اساسی برای هستی‌شناسی رسمی معرفی نمود. این نظریه بخشی از پروژه‌ای بزرگ‌تر با هدف توسعه چارچوبی کلی و عمومی برای هستی‌شناسی غیررسمی بود که در آن هوسرل علاوه بر کل-جزء شامل نوع و گونه، سوژه و کیفیت، رابطه و مجموعه، یگانگی و از این قبیل، چندین ایده دیگر را مطرح نمود. صرف‌نظر از این که آیا دامنه تعریف ما دربرگیرنده موضوعاتی هم‌سو با واقعیت، موجودیت‌های انتزاعی به‌هم پیوسته و از این قبیل است یا نه؛ برخی ویژگی‌های کلی باید نمایش داده شوند و از دستورات و قوانینی کلی پیروی کنند. وظیفه هستی‌شناسی -آنچه به‌طور رسمی درک و پذیرفته شده- این است که ویژگی‌ها و قوانین را کشف کند. به‌عنوان نمونه فارغ از این که موجودیت ما چیست، وظیفه هستی‌شناسی پشتیبانی از این است که هر موجودیت از قوانین خاص مرتبط با هویتش از قبیل انعطاف‌پذیری، تناسب (تقارن) یا انتقال‌پذیری تبعیت کند:

- (1) $x = x$
- (2) if $x = y$, then $y = x$
- (3) if $x = y$ and $y = z$, then $x = z$.

درستی و صحت این قوانین به این که موجودیت‌ها به چه متغیرهای منفردی اشاره دارند، بستگی ندارد. X ، y و z دقیقاً همان‌گونه که حقیقت قوانین زیر -که مرتبط با معادل آن‌ها هستند- به آنچه که به متغیرهای جمله‌ای " p "، " q " و " r " اختصاص داده شده بستگی ندارند:

- (1') $p _ p$
- (2') if $p _ q$, then $q _ p$
- (3') if $p _ q$ and $q _ r$, then $p _ r$.

به عقیده هوسرل و ورزی، همان‌گونه که ساوسا (۲۰۱۳) اشاره می‌کند، مرئولوژی نظریه‌ای رسمی مرتبط با ساختارهای کل و جزء است در حالی که توپولوژی نظریه تمامیت و کل است که روابط مرتبط با موجودیت‌ها و چگونگی قرارگیری این اجزاء درون سیستم را تعریف می‌کند. سیستم مترولوژیکال^۵ پایه به‌صورت $M = (E, \leq)$ است که در آن E موجودیت‌های دامنه است و $\leq$ روابط دودویی. روابط دودویی بین E و $\leq$ دلالت بر این دارد که M می‌تواند به‌عنوان پایه و اساس روابط مرئولوژی در نظر گرفته شود.

به‌طور خلاصه همان‌طور که ساوسا (۲۰۱۳) در پژوهش خود به نقل از پُلی (۲۰۰۱) اشاره می‌کند،

¹ Mereology

² part-whole relation

³ Husserl

⁴ Stanisław Lesniewski

⁵ Metrological

تعدادی از نظرات فلاسفه و متفکرانی مانند ارسطو، هارتمن و هوسرل که هستی‌شناسی را به عنوان چارچوبی برای مقوله‌بندی و ابزاری برای کاوش جهان هستی می‌دانند از سوی بسیاری محققین پذیرفته شده است. در مجموع، بر خلاف طیف گسترده پژوهش‌های هستی‌شناسی از ماهیت وجود در جهان توسط فلاسفه گرفته تا ساخت مدل‌هایی از اهداف تعیین شده دانشمندان علوم رایانه، همه این نظریه‌ها در کنار هم راهنمایی برای ایجاد هستی‌شناسی‌های حوزه‌های گوناگون و برر سی تعاریف و کاربردهای پیوندهای معنایی پایه مانند is-a، part of، instance of فراهم می‌کنند (دانشگاه اوساکا، آزمایشگاه میزوگوچی^۱ ۲۰۰۶).

با در نظر گرفتن هستی‌شناسی‌ها از دیدگاه علم اطلاعات و کتابداری و طرح‌های طبقه‌بندی مورد توجه در آن می‌توان دریافت که نظریه طبقه‌بندی پویای به نقل از ساوسا (۲۰۱۳) بیش از مجموعه‌ای از چهریزه‌های پیوسته به هم از طریق مقولات پایه است. بخش‌های دیگر نظریه طبقه‌بندی او با هستی‌شناسی چهریزه‌ها به عنوان ریزواحد‌های معنایی سروکار دارد. رانگانانان نظریه تحلیل چهریزه‌ها را از آنجایی توسعه داد که از ناکارآمدی سیستم‌های سنتی شمارشی مرتبط با طبقه‌بندی کتابشناختی خصوصاً در مورد نحوه بیان موضوعات مرکب ناراضی بود. سیستم‌های طبقه‌بندی مانند دهدهی دیوئی و کنگره تلاش می‌کنند تا موضوعات مورد نظر در آثار منتشر شده را به صورت شمارشی و با کدها نمایش دهند. با این وجود سیستم‌های یاد شده به سهولت قابلیت ترکیب اصطلاحات را از قسمت‌های مختلف طرح‌های طبقه‌بندی به منظور بیان موضوعات مرکب نمی‌دهند. نظریه‌ها و طرح‌های طبقه‌بندی در قالب‌هایی مانند تاکسونومی‌ها^۲ و اصطلاحنامه‌ها (شامل کلاس‌ها^۳ و زیرکلاس‌ها^۴) ظاهر شده‌اند تا مقولات را غالباً بر اساس نظریه ارسطو که از آن یاد شد، طبقه‌بندی کنند. هستی‌شناسی‌ها را می‌توان نوع تکامل یافته و غنی شده معنایی طرح‌های طبقه‌بندی دانست.

در مجموع می‌توان دریافت که هستی‌شناسی نیز به عنوان موجودیتی که سعی در تعریف مفاهیم یک دامنه و ارتباط آن‌ها با یکدیگر دارد نیز مانند دیگر ابزارهای سازماندهی علوم و دانش، وام دار فلاسفه و طرح‌های طبقه‌بندی ارائه شده توسط آنان است که در پیوند با علوم رایانه، علم اطلاعات، زبان‌شناسی و ریاضیات به صورت کنونی ظاهر شده‌اند.

۲-۴. هستی‌شناسی در مقایسه با ابزارهای دیگر سازماندهی و بازنمون دانش

انسان برای درک نظام‌مند جهان هستی به شیوه‌های گوناگون به مقوله‌بندی و رده‌بندی پدیده‌های هستی می‌پردازد. مقوله‌های ایجاد شده توسط برخی فلاسفه از جمله ارسطو را می‌توان یکی از این

^۱ Osaka University, Mizoguchi Lab

^۲ Taxonomy

^۳ superclass

^۴ subclass

تلاش‌ها دانست. درک کنونی ما از رده‌بندی‌های سلسله‌مراتبی در واقع از دیدگاه‌های وی اقتباس شده است. او معتقد بود طبیعت از یک کل واحد و یکپارچه تشکیل شده که می‌تواند به چندین رده طبیعی و همچنین رده‌های فرعی‌تر تقسیم شود. این فرایند بر اساس مجموعه‌ای از قواعد منظم پیوند^۱ و تمایز^۲ انجام می‌شود. زمانی که مفاهیم شکل گیرند و روابط میان آن‌ها درک شود، رده‌بندی می‌تواند به‌عنوان بازنمونی غنی از شناخته‌ها به کار رود. در راستای تلاش انسان‌ها برای طبقه‌بندی پدیده‌ها به تدریج نظام‌ها و طرح‌ها و نظام‌های رده‌بندی مختلفی شکل گرفت که از جمله آن‌ها می‌توان به طرح‌های سلسله‌مراتبی^۳، درختواره^۴ و تجزیه و تحلیل چهریزه‌ای^۵ اشاره داشت که بر اساس ایده‌های پیشگامانی از جمله ملویل دیوئی و رانگاناتان ارائه شدند. این تلاش‌ها در واقع برای ایجاد نظام‌هایی انجام شد که امکان «سازماندهی و بازنمایی دانش» را از طریق ارائه مفاهیم و روابط میان آن‌ها فراهم می‌کنند.

ساوسا (۲۰۱۳) انواع نظام‌های سازماندهی و بازنمون دانش را که بر اساس تعریف او می‌توانند به‌عنوان «ابزارهای معناشناسی» در نظر گرفته شوند که تفسیری سازمان‌یافته از ساختارهای دانش ارائه می‌دهند، به سه گروه کلی تقسیم می‌کند:

۳. فهرست‌های واژه‌ها: ابزارهایی که شامل فهرست‌هایی از واژه‌ها اغلب همراه با تعریف آن‌ها هستند. مانند واژه‌نامه‌ها^۶ و فرهنگ‌ها^۷؛

۴. رده‌بندی‌ها و طبقه‌بندی‌ها: ابزارهایی که مجموعه‌های موضوعی از مفاهیم و روابط میان آن‌ها ایجاد می‌کنند. مانند سرعنوان‌های موضوعی، طرح‌های رده‌بندی و طبقه‌بندی‌ها^۸؛

۵. فهرست‌های رابطه‌ای^۹: ابزارهایی که به ارائه روابط و پیوندهای میان واژه‌ها و مفاهیم می‌پردازند. این طبقه شامل اصطلاح‌نامه‌ها، شبکه‌های معنایی و هستی‌شناسی‌ها است.

در همین راستا ساوسا (۲۰۱۳) در پژوهش خود به مقایسه این ابزارها از جنبه‌های مختلف پرداخته که خلاصه‌ای از آن در جدول ۱-۲ قابل مشاهده است.

جدول ۱-۲. مقایسه ابزارهای مختلف سازماندهی و بازنمون دانش (ساوسا ۲۰۱۳)

¹ Association

² Distinction

³ Hierarchy

⁴ Trees

⁵ Faceted Analysis

⁶ Dictionary

⁷ Glossary

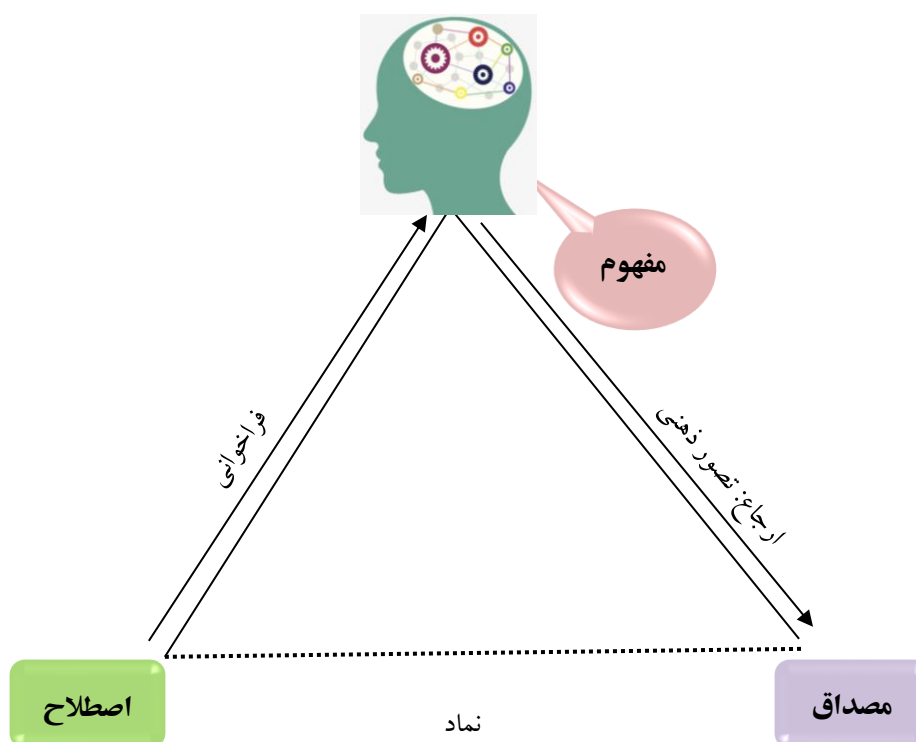
⁸ Taxonomies

⁹ Relationship Lists

ابزار	تاکسونومی	تزاروس	مدل مفهومی	وب معنایی	هستی‌شناسی
مورد مقایسه	مترادف	واژگان کنترل شده	-	-	مدل
خروجی	طبقه‌بندی مفاهیم، اصطلاحات و اشیاء	فهرستی از کلمات و مترادف‌های سازماندهی شده بر مبنای معنی اصطلاحات	مدل ذهنی در حوزه خاصی از دانش	توصیف اطلاعات در وب گسترده جهانی	بازنمون معنای پیچیده مفاهیم و روابط
ساختار	سلسله‌مراتبی، درختی	روابط هم‌ارزی، همبستگی و سلسله‌مراتبی بر اساس استانداردهای مربوط	سلسله‌مراتب پیچیده‌ای از دانش	درختی	روابط بین مقولات
پیوند بین مفاهیم	والد و فرزندی	روابط اعم، اخص و مرتبط	موجودیت، روابط، ارزش‌ها و قواعد	سلسله‌مراتبی	کلاس‌ها، نمونه‌ها، روابط، ویژگی‌ها و محدودیت‌ها
اساس شکل‌گیری	واژه‌نامه و اصطلاح‌نامه	واژه‌نامه	تاکسونومی	تاکسونومی	تاکسونومی
هدف	طبقه‌بندی اشیاء	راهبری مفهومی، کاوش و بازیابی اطلاعات	بازنمون موجودیت‌های اصلی یک دامنه	خود کارسازی و تلفیق و استفاده مجدد از اطلاعات بین برنامه‌های کاربردی	گردآوری و بازنمون معانی یک دامنه
عملکرد در بازیابی اطلاعات	ضعیف	ضعیف	ضعیف	ضعیف	قوی

بنابر آنچه ذکر شد، یکی از روش‌های درک بهتر ویژگی‌های هستی‌شناسی تشریح تمایز آن با دیگر ابزارهای سازماندهی دانش خصوصاً اصطلاح‌نامه است که به‌نوبه خود دربردارنده تمایز اصطلاح

در مقابل مفهوم می‌شود. نگاهی به مثلث معنای^۱ آگدن و ریچاردز^۲ (۱۹۲۳) که در شکل ۱-۲ آمده به درک بهتر فرایند دریافت معنا در ذهن انسان کمک می‌کند. این مثلث بیانگر این است که چطور افراد از نشانه‌ها به عنوان بازنمون اشیاء استفاده می‌کنند تا بتوانند با مفاهیم اشیاء حتی در غیاب آن‌ها ارتباط برقرار کنند. بر همین اساس در دنیای واقعی بین افکار (مفاهیم)، واژه‌ها (اصطلاحات) و موجودیت‌ها (اشیاء) تعاملاتی وجود دارد که در مثلث معنا توصیف شده‌اند. اولین عنصر این مثلث اصطلاح است که در واقع نماد یا واژه است و قواعد نحوی حاکم بر آن می‌تواند به صورت جمله یا عبارت نیز ارائه شود. نمادها به تنهایی بی‌معنا هستند مگر آن‌که در یک مثلث معنایی با مفاهیم و مصداق‌های آن‌ها در جهان هستی پیوند حاصل کنند.



شکل ۱-۲. مثلث معنا (آگدن و ریچاردز ۱۹۲۳)

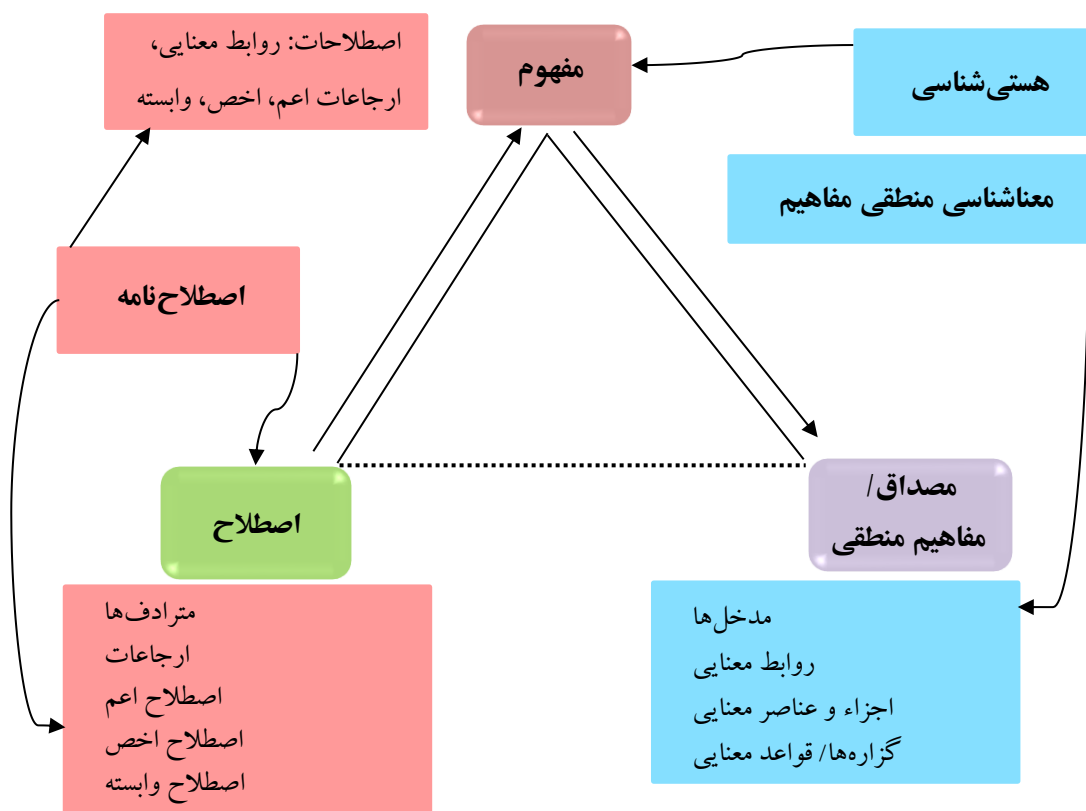
اصطلاح‌نامه به مفهوم نیازمند است تا بتواند میان اصطلاح و پدیده‌های جهان هستی ارتباط برقرار کند و ارجاعات را ایجاد نماید. اصطلاح‌نامه معمولاً با اصطلاحات و مفاهیم سروکار دارد در حالی که هستی‌شناسی با مفاهیم و ارجاع به مصداق‌ها در جهان هستی پیوند می‌یابد. از نظر معناشناختی روابط اصطلاح‌نامه غالباً در حد ارائه ویژگی‌های روابط بین اصطلاحات از جمله اعم و اخص (سلسله‌مراتبی) است. آنچه در مورد یک اصطلاح‌نامه باید بدانیم این است که یک اصطلاح به صورت گرهی در

¹ Meaning triangle

² Ogden & Richards

اصطلاح‌نامه و برای تفکیک معنایی آن از دیگر اصطلاحات و رفع ابهامات معنایی آن ایجاد شده است و به صورت اعم یا اخص تر از دیگر اصطلاحات نشان داده می‌شود.

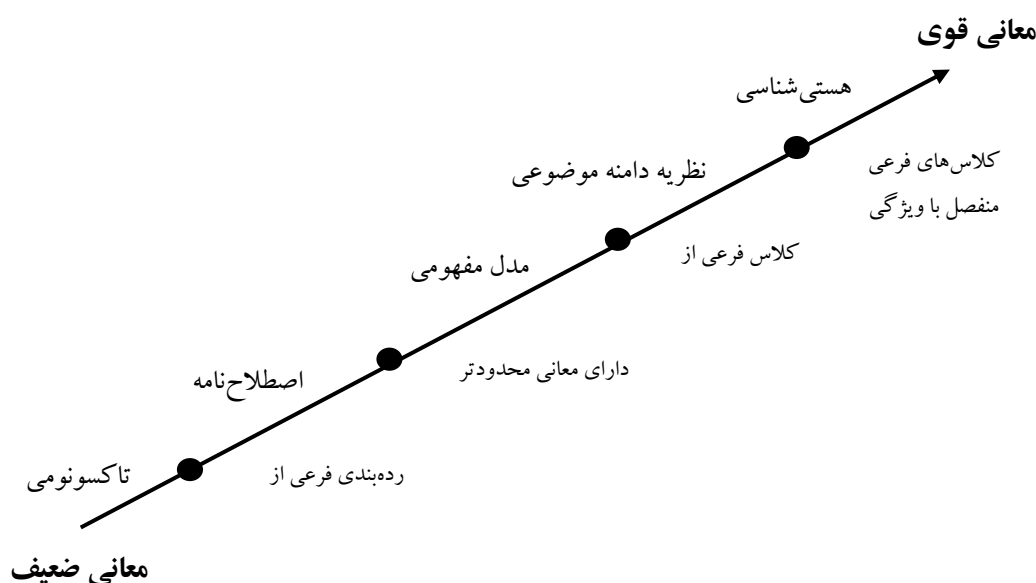
هستی‌شناسی سعی در بازنمایی معنایی مفاهیم و روابط پیچیده میان آن‌ها دارد و بدین جهت اجزاء، عناصر، ویژگی‌ها، ارزش گذاری، زنجیره‌های معنایی و قواعد حاکم بر آن‌ها را تعیین می‌کند. بنابراین هدف هستی‌شناسی از اصطلاح‌نامه متمایز است. هستی‌شناسی در صدد اکتساب و بازنمایی معانی یک دامنه موضوعی است که با پدیده‌های جهان هستی منطبق باشد، زیرا هدف آن شبیه سازی مفاهیم ذهنی انسان و ساختار دامنه موضوعی آن در جهان هستی است. البته هستی‌شناسی نیز مانند اصطلاح‌نامه نیازمند گسترش، روزآمدسازی، ویرایش و کاربرد مجدد است و علاوه بر دارا بودن کاربردهای اصطلاح‌نامه، هستی‌شناسی باید به معناشناسی دقیق، پیچیده و غنی مفهومی بپردازد. شکل ۲-۲ به خوبی تفاوت میان این دو را نشان می‌دهد (ساوسا ۲۰۱۳).



شکل ۲-۲. اصطلاح‌نامه در مقابل هستی‌شناسی
(دکونتا و همکاران ۲۰۰۳)

دکونتا ضمن مقایسه هستی‌شناسی‌ها با اصطلاح‌نامه و تاکسونومی تشریح می‌کند که تاکسونومی‌ها دارای تقسیم‌بندی سلسله‌مراتبی و ساختار پدر و فرزند 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 این جهت نسبت به تاکسونومی‌ها تمایز معنایی دقیق‌تری میان مفاهیم قائل می‌شوند. هستی‌شناسی‌ها در یک مرحله بالاتر با ارائه الگوی مفهومی^۱، تقسیم‌بندی معناشناختی دقیق‌تری میان مفاهیم ایجاد می‌کنند که بر یک نظریه منطقی^۲ استوار است و روابط ریزتر و دقیق‌تر معنایی را نمایش می‌دهد. بدین ترتیب می‌توان تاکسونومی‌ها را اسکلت معنایی در هستی‌شناسی‌ها دانست. هستی‌شناسی‌ها به‌طور خاص مفاهیم را از واژگان جدا نموده و از این راه ساختار درک انسان از دامنه‌ها را بهتر منعکس می‌کنند. در هستی‌شناسی‌ها، معانی از طریق اطمینان از این‌که هر مفهوم درون دامنه به‌صورت انحصاری و دقیق تعریف شده و همچنین از طریق تعیین و تخصیص تفصیلی روابط بین مفاهیم شرح داده شده‌اند. شکل ۲-۳ به درک این طیف معنایی کمک می‌کند.



شکل ۲-۳. طیف هستی‌شناسی (دکوتا و همکاران ۲۰۰۳)

همان‌طور که در شکل ۳-۳ نشان داده شده، در مقایسه با دیگر ابزارهای رده‌بندی معناشناختی، هستی‌شناسی‌ها دانش پیچیده‌تر و با غنای بیشتری را همراه با جزئیات و مشخصات معنایی ارائه می‌کنند.

۲-۵. طبقه‌بندی هستی‌شناسی‌ها

تقسیم‌بندی‌های متعددی از هستی‌شناسی‌ها بر اساس مواردی مانند اهداف، کاربردها، روش‌های ایجاد، زبان‌ها، سطح بازنمون دانش و از این دست صورت گرفته است که در این بخش به برخی از

^۱ Conceptual model

^۲ Logical theory

آنها که بیشتر مورد توجه پژوهش‌ها بوده اشاره می‌شود.

ون‌هیست و همکاران^۱ (۱۹۹۶) هستی‌شناسی‌ها را در قالب دو نوع طبقه‌بندی اصلی به انواع تقسیم می‌کنند که بیشتر بر محوریت کاربرد و دامنه موضوعی هستی‌شناسی‌ها استوار است. انواع هستی‌شناسی را بر اساس سطح عمومیت (کلی بودن) آنها به صورت زیر طبقه‌بندی می‌کند:

□ هستی‌شناسی‌های سطح بالا که مفاهیم کلی مانند فضا، زمان، ماده، شیء، رویداد، عمل و از

این دست را توصیف کرده و مستقل از یک مسئله یا دامنه خاص هستند؛

□ هستی‌شناسی‌های دامنه که واژگان مربوط به یک دامنه موضوعی خاص (مانند پزشکی یا

اتومبیل) را از طریق تخصیص اصطلاحات معرفی شده در هستی‌شناسی‌های سطح بالا

توصیف می‌کنند؛

□ هستی‌شناسی‌های وظیفه‌گرا^۲ که واژگان مربوط به یک حوزه کاری یا فعالیت کلی (مانند

تشخیص یا فروش) را توصیف می‌کنند؛

□ هستی‌شناسی‌های کاربردی که مفاهیم وابسته به یک دامنه و فعالیت خاص را توصیف کرده

و اغلب نوع ویژه‌ای از هر دوی هستی‌شناسی‌ای ذکر شده هستند (گوارینو ۱۹۹۸، ۸-۹).

با توجه به نوع زبان، سوا (۲۰۰۰) هستی‌شناسی‌ها را در سه مقوله طبقه‌بندی می‌کند:

□ هستی‌شناسی‌های غیررسمی که با فهرستی از انواع تعریف شده و نشده به زبان طبیعی مشخص

می‌شوند. این هستی‌شناسی‌ها شامل تمام اصطلاح‌شناسی یک حوزه و دامنه خاص شده و

مفاهیم و روابط آن را طبقه‌بندی می‌کند؛

□ هستی‌شناسی‌های رسمی که با مجموعه‌ای از نام‌ها برای مفاهیم و انواع روابط سازمان‌یافته در

یک نظام بر اساس انواع اصلی و زیر شاخه‌های آن مشخص می‌شوند. این هستی‌شناسی‌ها

توسط ماشین پردازش می‌شوند و معمولاً از زبان هستی‌شناسانه مانند زبان هستی‌شناسی وب

^۳OWL و ^۴DAML+OIL برای کدگذاری آن استفاده می‌شود؛

^۱ Van Heijst & et al.

^۲ Task ontologies

^۳ Web Ontology Language

^۴ DARPA Agent Mark-up Language + Ontology Inference Layer

□ هستی‌شناسی دامنه این هستی‌شناسی‌ها مربوط به یک حوزه خاص از دانش بوده و حاوی مفاهیم اصلی و روابط آن‌ها هستند. هستی‌شناسی دامنه راهکاری برای بازنمون مفاهیم خاص در هر دامنه است (گومز پرز ۲۰۰۴).

همچنین رُسی و همکاران (۲۰۱۱) هستی‌شناسی‌ها را بر اساس بیان و گویایی و رسمیت زبان مورد استفاده به هستی‌شناسی‌های اطلاعاتی، زبان‌شناسی / اصطلاح‌شناسی، نرم‌افزاری و رسمی طبقه‌بندی و بر اساس دامنه موضوعات توصیف شده در هستی‌شناسی به محلی / کاربردی، دامنه، مرجع هسته، عمومی، پایه / سطح عالی / سطح بالا تقسیم می‌کنند که به همراه دیگر طبقه‌بندی‌های ذکر شده در جدول ۲-۳ قابل مشاهده است (هماوندی ۱۳۹۹).

جدول ۲-۲. رویکردهای مختلف در طبقه‌بندی انواع هستی‌شناسی

پژوهشگران	انواع طبقه‌بندی‌های هستی‌شناسی
ون هیست و همکاران (۱۹۹۶)	۱- شامل سه مقوله ۲- شامل چهار مقوله هستی‌شناسی‌های اصطلاحات به عنوان واژه‌نامه هستی‌شناسی‌های اطلاعات به عنوان طرح‌های پایگاه داده هستی‌شناسی‌های مبتنی بر مدل‌سازی دانش هستی‌شناسی‌های کاربردی
گوارینو (۱۹۹۸)	هستی‌شناسی‌های سطح بالا هستی‌شناسی‌های دامنه هستی‌شناسی‌های وظیفه‌گرا هستی‌شناسی‌های کاربردی
سوا (۲۰۰۰)	هستی‌شناسی‌های غیررسمی هستی‌شناسی‌های رسمی هستی‌شناسی‌های دامنه
رُسی و همکاران (۲۰۱۱)	۱- بر اساس زبان مورد استفاده: هستی‌شناسی‌های اطلاعاتی زبان‌شناسی / اصطلاح‌شناسی نرم‌افزاری و رسمی طبقه‌بندی ۲- و بر اساس دامنه موضوعات توصیف شده: محلی / کاربردی دامنه مرجع هسته عمومی پایه / سطح عالی / سطح بالا

۶-۲. ساختار هستی‌شناسی

همان‌طور که در بخش تعاریف هستی‌شناسی نیز اشاره شد، این ابزار برای استفاده در سیستم‌های بازیابی اطلاعات باید ماشین‌خوان باشد و لازمه این ویژگی، قابلیت رسمی‌بودن هستی‌شناسی است. برای این منظور لازم است در تعریف مفاهیم هستی‌شناسی از زبان‌های قابل فهم و قابل استدلال برای ماشین استفاده شود. برای انجام این مهم، زبان‌های مختلفی ایجاد شده است که در تمام آن‌ها هستی‌شناسی دارای اجزای زیر است:

کلاس‌ها یا موجودیت‌ها^۱: که عبارتند از گروهی از موجودات یا اشیاء در یک دامنه موضوعی که دارای ویژگی‌های مشابهی هستند و می‌توانند مفاهیم را بازنمون کنند. کلاس‌های هستی‌شناسی با کلاس‌های زبان‌های شیء‌گرا^۲ اندکی متفاوتند. مطابق تعریف فوق، کلاس در هستی‌شناسی، یک مجموعه است حال آنکه در زبان‌های شیء‌گرا کلاس یک قالب برای ساخت اشیاء است.

نمونه‌ها یا مصادیق^۳: به هر یک از اعضای یک کلاس، یک «نمونه» گفته می‌شود. یک شیء می‌تواند به چندین کلاس تعلق داشته باشد یا به عبارتی دیگر نمونه چندین کلاس باشد. یک شیء در صورت داشتن ویژگی‌های خاص یا تأمین شروط ویژه می‌تواند به عضویت یک کلاس درآید یا نمونه آن کلاس گردد.

رابطه یا خصیصه^۴: یک «رابطه»، همان‌طور که از اسمش برمی‌آید، نمونه‌هایی از یک کلاس را به نمونه‌هایی از یک کلاس دیگر یا به مقادیر ثابت، مرتبط می‌کند. اغلب به رابطه، «خصیصه» نیز گفته می‌شود. از انواع رابطه‌ها در هستی‌شناسی می‌توان به has-a، Is-a و has-child اشاره نمود. این روابط در واقع نشان‌دهنده هم‌ارزی و مترادف، والد و فرزندی و کل و جزء هستند.

محدودیت‌ها و قواعد^۴: که شامل اطلاعاتی در مورد موجودیت‌ها و قواعد مرتبط با آن‌ها هستند و چگونگی استفاده از روابط را تعیین می‌کنند.

اصول: از این عنصر برای تأیید صحت عبارت‌ها و تعاریف استفاده می‌شود و برای تفسیر دانش جدید بسیار کمک‌کننده است.

۷-۲. معیارهای طراحی هستی‌شناسی

مرحله مهم و پایه برای ایجاد هستی‌شناسی، طراحی آن است. این مرحله به ما امکان می‌دهد تا پیش از اجرا در مورد مسائل و ابعاد مختلف دخیل در پیاده‌سازی هستی‌شناسی تصمیم‌گیری کنیم. بر

^۱ Class or Entities

^۲ Instances or Individual

^۳ Properties

^۴ Restrictions and rules

همین اساس نحوه پرداختن به این مرحله در کیفیت و خروجی هستی‌شناسی تأثیر به‌سزایی دارد. گروبر (۱۹۹۳) عقیده دارد برای طراحی هستی‌شناسی به راهنما جهت ارزیابی آن نیاز است. در ادامه او معیارهای پنج‌گانه‌ای را ارائه می‌دهد که باید در طراحی هستی‌شناسی با هدف اشتراک‌گذاری دانش و میان‌کنش‌پذیری بین برنامه‌ها رعایت شود:

۱. شفافیت: هستی‌شناسی‌ها باید به خوبی معانی اصطلاحات را منتقل کنند. تعاریف باید عینی باشند. به‌عنوان مثال زمانی که انگیزه تعریف یک مفهوم، ریشه در موقعیت‌های اجتماعی یا ملزومات رایانشی داشته باشد، باید تلاش شود که ارائه تعریف مستقل از بافت‌های ذکر شده باشد. وقتی یک تعریف می‌تواند بر اساس اصول منطقی بیان شود باید تا حد امکان بر همین اساس عمل نمود. همچنین تا حد امکان تعاریف به صورت کامل ارائه و به‌وسیله زبان طبیعی مستند شوند؛

۲. انسجام: هستی‌شناسی باید منسجم باشد به این معنی که باید بر استنتاج‌های منطبق و سازگار با تعاریف مبتنی باشد. حداقل تعاریف اصول باید از نظر منطقی سازگار بوده و تناقض نداشته باشند. انسجام حتی باید در مورد تعاریف غیررسمی مفاهیم توصیف شده به زبان طبیعی هم رعایت شود؛

۳. گسترش‌پذیری: طراحی هستی‌شناسی باید کاربردها و واژگان مشترک را پیش‌بینی کند. به این معنی که امکان تعریف اصطلاحات جدید برای موارد خاص بر اساس واژگان موجود هستی‌شناسی وجود داشته باشد بدون این که به بازنگری تعاریف موجود نیاز باشد؛

۴. حداقل سوگیری در کدگذاری: مفهوم‌سازی باید در سطح دانش تخصیص یابد و از کدگذاری در سطح نمادها مستقل باشد. سوگیری در کدگذاری زمانی ایجاد می‌شود که انتخاب بازنمون‌ها صرفاً بر مبنای سادگی نشانه‌گذاری یا اجرا انجام می‌شود. این مسئله از این جهت اهمیت دارد که دانش بازنمون شده ممکن است در سیستم‌ها یا شبکه‌های بازنمون دیگر اجرا شود؛

۵. حداقل تعهدات: هستی‌شناسی هنگام طراحی باید به حداقل تعهدات برای پشتیبانی از اشتراک‌گذاری دانش بسنده کند. به این معنی که تا حد امکان ادعای خود را در گستره پوشش محدود کند چرا که تعهدات آن به استفاده پیوسته از واژگان بستگی دارد. در همین راستا لازم است در مرحله طراحی به یک نظریه یا مدل محدود شود و بر اساس آن اصطلاحات ضروری برای برقراری جریان دانش در آن حوزه یا دامنه را ارائه کند. بهتر است هستی‌شناسی واژگان مورد استفاده و رایج در یک دامنه را توصیف کند تا این که به دانش

مورد نیاز برای حل یک مسئله یا پاسخ به سؤالات اختیاری در یک دامنه بپردازد.

۸-۲. زبان‌های هستی‌شناسی

هستی‌شناسی‌ها را می‌توان زبان بازنمایی اطلاعات (شامل واژگان، نحو و معنا‌شناسی یک متن) دانست. علاوه بر این، در هستی‌شناسی‌ها ارائه محتوا صورت می‌گیرد و از این رو به زبان‌هایی برای بازنمایی محتوا نیازمندند، که معمولاً «زبان بازنمایی دانش»^۱ نامیده می‌شوند. زبان‌های متعددی برای بازنمایی دانش در وب معنایی به وجود آمده است که از جمله آن‌ها می‌توان به کا. ال-وان^۲، آنتولینگوا^۳، کلاسیک^۴، لوم^۵، قالب تبادل دانش^۶، سایک. ال^۷، یو.ام. ال^۸، دی.ای.ام. ال اوپل^۹، آر.دی.اف.اس و زبان هستی‌شناسی وب^{۱۰} اشاره نمود. این زبان‌ها در دسترسی به محتوای وب و ارائه فراداده به منظور اشتراک‌گذاری، تبادل، پردازش و فهم محتوا توسط رایانه‌ها لازم هستند. در این بخش به مرور کوتاهی بر بعضی از پرکاربردترین زبان‌های ساخت هستی‌شناسی بسنده می‌شود. به نقل از ساوسا (۲۰۱۳) زبان‌های به کار رفته برای بازنمون دانش در هستی‌شناسی باید دارای قابلیت‌هایی باشند:

- معقول و فشرده بودن؛
- معناشناسی قوی و دقیق برای بازنمون دقیق مفاهیم؛
- قدرت بیان کافی به منظور بازنمایی دانش بشر (حوزه موضوعی)؛
- سازوکار استدلال قابل فهم، قوی و کارآمد و؛
- قابلیت استفاده برای ساخت پایگاه‌های دانش عظیم.

هر چند هم‌اکنون زبان‌های مختلفی مانند اچ.تی.ام. ال^{۱۱} و ایکس.ام. ال^{۱۱} برای نمایش محتوای صفحات وب وجود دارند، اما این زبان‌ها دارای کاستی‌هایی هستند که نمی‌توانند به منظور بازنمون هستی‌شناسی‌ها مورد استفاده قرار گیرند. عنا صر فراداده‌ای زبان اچ.تی.ام. ال که برای کاربران انسانی قابل فهم هستند، برای نظام‌های رایانه‌ای و نرم‌افزارهای کاربردی معنی خاصی ندارند. ایکس.ام. ال روز

¹ Knowledge Representation Language

² KL-ONE

³ Ontolingua

⁴ Classic

⁵ LOOM

⁶ KIF (Knowledge Interchange Format)

⁷ CycL

⁸ Uniformed Markup Language

⁹ OWL (Ontology Web Language)

¹⁰ HTML (Hypertext Markup Language)

¹¹ XML (Extensible Markup Language)

به روز بیشتر پذیرفته می‌شود و به سرعت جایگزین اچ.تی.ام.ال تحت عنوان زبان وب می‌گردد اما چارچوب کلی آن بیشتر در ارتباط با ساختار فیزیکی مدارک وب است. علاوه بر این، برچسب‌های ایکس.ام.ال مدل‌سازی معنایی دقیقی ندارند که بتوانند از تفسیرهای رایانه‌ای حمایت کنند. بر همین اساس، پیش‌نیاز کاربرد هستی‌شناسی‌ها در وب معنایی توسعه استاندارد برای تعریف و مبادله هستی‌شناسی یا به عبارتی زبان‌های بازنمون هستی‌شناسی است. زبان‌هایی مانند آر.دی.اف و آر.دی.اف.اس را که با همکاری کنسرسیوم وب جهانی^۱ توسعه یافتند و الگوی زبانی توصیف منابع هستند و در توسعه زبان‌های دیگری چون اوپل و دی.ای.ام.ال اوپل - که با همکاری دارپا انجام گرفته - نقش داشته‌اند می‌توان به عنوان نمونه‌هایی از این زبان‌ها نام برد.

به عنوان جدیدترین زبان توسعه یافته برای وب معنایی می‌توان به زبان هستی‌شناسی وب اشاره داشت که در سال ۲۰۰۴ توسط کنسرسیوم وب جهانی توسعه پیدا کرد و در طول زمان نسخه‌هایی روزآمد از آن نیز ارائه شد که هم‌اکنون نسخه دوم زبان هستی‌شناسی وب^۲ در دسترس قرار دارد. این زبان را می‌توان از پرکاربردترین زبان‌ها در پیاده‌سازی هستی‌شناسی‌ها دانست. زبان یاد شده بر مبنای ایکس.ام.ال و آر.دی.اف / آر.دی.اف.اس ایجاد و بر اساس نحو آن‌ها عمل می‌کند و طراحی آن بر بازنمون نحوه ارتباط اشیاء در یک مقوله خاص تمرکز دارد. این زبان از منطق توصیفی و با هدف افزایش قدرت استدلال و گویایی در توصیف محتوای وب معنایی ایجاد شده است. همچنین زبان هستی‌شناسی وب سازگاری بالایی با دیگر زبان‌ها، چارچوب‌ها و استانداردها دارد. یکی از مزایای این زبان نسبت به چارچوب توصیف منبع، تفاوت آن در تفسیرپذیری توسط رایانه به دلیل گستره وسیع واژگان و نحو آن است. علاوه بر این، روابط در این زبان غنی‌تر و کامل‌تر هستند (ساوسا ۲۰۱۳). بر اساس مبنای شکل‌گیری، کالیباتین و واسیلکس^۳ (۲۰۱۱) انواع زبان‌های یاد شده را در قالب شکل ۲-۴ تقسیم‌بندی کردند. بر همین اساس زبان‌های سنتی شامل زبان‌های مبتنی بر منطق استدلالی مرتبه اول مانند قالب تبادل دانش و سایک.ال، زبان‌های مبتنی بر چارچوب مانند آنتولینگوآ، اف - لاجیک^۴ و ا.سی.ام.ال^۵، زبان‌های مبتنی بر منطق توصیفی^۶ مانند لوم و دیگر زبان‌ها هستند. گروه دوم زبان‌های مبتنی بر استانداردهای وب هستند که می‌توان به زبان هستی‌شناسی وب، به عنوان مهم‌ترین آن‌ها اشاره داشت (هماوندی ۱۳۹۹).

^۱ W3C (World Wide Consortium)

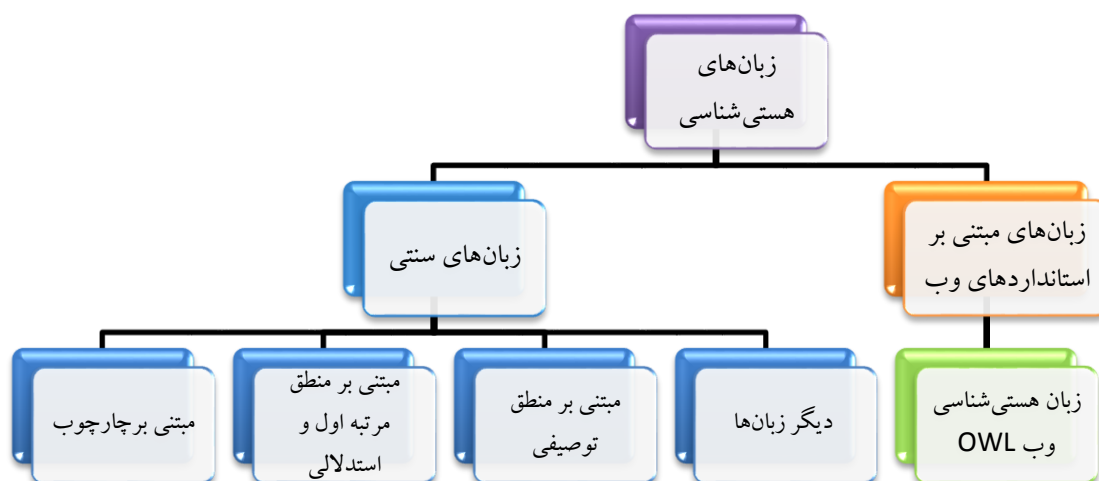
^۲ OWL2

^۳ Kalibatiene & Vasilecas

^۴ F-Logic (frame Logic)

^۵ OCML (Operational Conceptual Modeling Language)

^۶ DL (Description logic)



شکل ۲-۴. زبان‌های هستی‌شناسی

مقایسه زبان‌های بازنمون هستی‌شناسی نشان‌دهنده تنوع و تفاوت آن‌ها در امکانات، توانایی‌ها و خصوصاً سطح گویایی آن‌هاست. در این میان باید توجه داشت که این زبان‌ها مراحل رشد اولیه خود را پشت سر می‌گذارند و دائماً در حال تحول هستند.

۲-۹. زبان‌های بازنمایی هستان‌شناسی

بسیاری از زبان‌های بازنمایی هستی‌شناسی‌ها در دهه ۹۰ میلادی پیشنهاد شده‌اند. این زبان‌ها مبتنی بر هوش مصنوعی و عمدتاً منطق مرتبه اول بوده‌اند. پیدایش اینترنت، منجر به خلق زبان‌های هستی‌شناسی بر پایه وب شد (گومز-پرز و همکاران ۲۰۰۴)^۱. به‌منظور کدگذاری و بازنمون هستی‌شناسی‌ها در وب، لازم است از زبان‌های خاص بازنمون هستی‌شناسی مانند آ.دبلیو.ال. و آر.دی.اف. استفاده شود. هم‌اکنون زبان‌های مختلفی مانند اچ.تی.ام.ال.^۲، ایکس.ام.ال.^۳ و غیره برای نمایش محتوای صفحات وب وجود دارد. این زبان‌ها کاستی‌هایی دارند که آن‌ها را برای بازنمون هستی‌شناسی نامناسب کرده است. عناصر فراداده‌ای زبان اچ.تی.ام.ال. که برای کاربران انسانی قابل فهم هستند، برای نظام‌های رایانه‌ای و نرم‌افزارهای کاربردی معنای خاصی ندارند. ایکس.ام.ال. هم که به سرعت رشد کرده و جایگزین اچ.تی.ام.ال. می‌شود، بیشتر در ارتباط با ساختار فیزیکی مدارک وب است. هم‌چنین برچسب‌های ایکس.ام.ال. مدل‌سازی معنایی دقیقی ندارند که بتوانند از تفسیرهای رایانه‌ای به‌طور کامل حمایت کنند.

زبان هستی‌شناسی به‌منظور کارایی در مدل‌سازی نظام‌ها باید دارای ویژگی‌ها به شرح زیر باشد:

^۱ Gómez-Pérez, Fernandez-Lopez and Corcho 2004

^۲ HTML (Hypertext Markup Language)

^۳ XML (Extensible Markup Language)

- ساختار نحوی منطقی و فشرده؛
 - ساختار معنایی که امکان بیان دقیق حوزه بازنمون‌شده را داشته باشد؛
 - قابلیت بیان کافی برای بازنمایی دانش؛
 - سازوکار استدلالی قدرتمند، قابل فهم و مؤثر؛
 - دارای قابلیت کاربردی در ایجاد ساختارهای عظیم دانش‌بنیان.
- زبان هستی‌شناسی همچنین باید معانی و مفاهیم را به صورتی ارائه دهد که قابل خواندن توسط ماشین باشد. به عبارتی، نه تنها باید قابلیت تعیین و تصریح واژگان را داشته باشد، بلکه باید آن‌ها را به گونه‌ای به صورت رسمی توصیف کند، که امکان به کارگیری در فرایندهای استنتاج خودکار را پیدا کنند (گوتیرز-پولیدو^۱ و دیگران ۲۰۰۶).
- در ادامه به برخی از زبان‌های بازنمون هستی‌شناسی‌ها اشاره می‌شود:
- ۹-۱. کی.آی.اف. (قالب تبادل دانش)^۲: این زبان در طی پروژه آنتولینگوا^۳ که یک هستی‌شناسی ساز اشتراکی برای استخراج هستی‌شناسی‌هایی با رسمی سازی‌های متفاوت است، خلق شد. کی.آی.اف.، یک زبان مبتنی بر منطق مرتبه اول می‌باشد که برای توصیف دانش و تبادل دانش میان نظام‌های مختلف به کار می‌رود. این زبان به منظور افزایش قابلیت پیاده‌سازی و خوانایی طراحی شده است. مبنای مفاهیم در این زبان، مصور ساختن جهان در قالب اشیاء و روابط میان آن‌ها می‌باشد. اشکال عمده این زبان، رسا بودن بالای آن است که منجر به ابهام در ایجاد نظام‌های تأییدکننده می‌شود. نظام‌های به وجود آمده با این زبان در مقایسه با نظام‌هایی که از زبان‌های دارای محدودیت بیشتری استفاده می‌کنند، کارایی کمتری دارند (کالیاتین و واسیلکاس ۲۰۱۱).
- ۹-۲. آنتولینگوا: این زبان که در سال ۱۹۹۲ میلادی در دانشگاه استنفورد به وجود آمد، مبتنی بر کی.آی.اف. می‌باشد و در میان زبان‌های بازنمون هستی‌شناسی‌ها، از همه رساتر است. آنتولینگوا یک محیط طبقه‌بندی شده اشتراکی برای مرور، ایجاد، ویرایش، تغییر و استفاده از هستی‌شناسی‌ها فراهم می‌کند. این زبان پارادایم‌های قالب را با گزاره‌های منطق مرتبه اول ترکیب می‌کند (به نقل از ساو سا ۲۰۱۳).
- ۹-۳. لوم: نظام بازنمایی دانش لوم در سال ۱۹۸۶ میلادی در مؤسسه علوم اطلاعاتی^۴ در دانشگاه

^۱ Gutierrez-Pulido

^۲ KIF (Knowledge Interchange Format)

^۳ Ontolingua Project

^۴ ISI (Information Sciences Institute)

کالیفرنای جنوبی به وجود آمد. لوم، طبقه‌بندی خود کار مفاهیم را ایجاد کرده و قابلیت بازنمون مفاهیم، طبقه‌بندی مفاهیم، روابط چندگانه^۱، توابع، اصول موضوعی و قواعد ساخت را دارد. این زبان، یک زبان برنامه‌نویسی سطح بالا و محیطی برای ساخت نظام‌های تخصصی و سایر برنامه‌های کاربردی است و برای مدل سازی اشیاء و روابط از زبان توصیفی و برای تخصیص محدودیت‌ها برای مفاهیم و روابط، از زبان تأکیدی استفاده می‌کند. در مقایسه با سایر زبان‌ها، لوم تنها زبانی است که تعداد ویژگی‌های بیشتری در حوزه‌های دانش را تحت پوشش قرار می‌دهد (لی^۲ ۲۰۱۲).

۲-۹-۴. اُ.سی.ام.ال. (زبان مدل سازی مفهومی عملیاتی)^۳: این زبان در سال ۱۹۹۳ میلادی در مؤسسه ر سانه‌های دانش^۴ دانشگاه اُپن^۵ برای ایجاد و توسعه هستی‌شناسی‌های اجرایی و مدل‌های موجود در روش‌های حل مسئله به وجود آمد. اُ.سی.ام.ال. یک زبان مبتنی بر ساختار می‌باشد که مکانیسم‌هایی را برای بیان مواردی مانند روابط، توابع، قواعد (زنجیره پس رو^۶ و زنجیره پیش رو^۷)، کلاس‌ها و نمونه‌ها فراهم می‌کند. اُ.سی.ام.ال.، از تعاریف منطق مرتبه اول پشتیبانی کرده و به کاربر قدرت متمایز کردن تعاریف کاربردی از تعاریف غیر کاربردی را می‌دهد.

۲-۹-۵. اِف-لاجیک (منطق ساختار)^۸: در سال ۱۹۹۵ میلادی در دانشگاه کارلسروهه^۹ آلمان به وجود آمد. مزیت اصلی این زبان، گسترش‌پذیری و توانایی بازنمایی مستقیم مفاهیم اصلی مشتق از زبان برنامه‌نویسی شیء‌گرا^{۱۰} و زبان‌های مبتنی بر ساختار هستند. هدف اصلی این زبان، تلفیق مدل سازی مفهومی اجزای هستان‌شناسی (کلاس‌ها، ویژگی‌ها، محدودیت‌های حوزه، توارث و اصول موضوعی) در قالب یک چارچوب منطقی منسجم می‌باشد. اِف-لاجیک با فراهم آوردن کلاس‌ها، ویژگی‌ها با توضیحات دامنه و حوزه، سلسله‌مراتب با مجموعه‌ای از زیر کلاس‌ها و توارث ویژگی‌ها، و اصول موضوعی، توصیف روابط میان اجزا و نمونه‌های هستی‌شناسی را امکان‌پذیر می‌کند.

۲-۹-۶. ایکس.ا.ال. (زبان هستی‌شناسی مبتنی بر ایکس.ام.ال.)^{۱۱}: یک زبان واسطه برای تبادل تعاریف هستی‌شناسی در میان اعضای حوزه می‌باشد. ترکیب آن مبتنی بر ایکس.ام.ال. و معانی آن

¹ N-ary Relations

² Li

³ OCML (Operational Conceptual Modeling Language)

⁴ KMI (Knowledge Media Institute)

⁵ Open University

⁶ Backward Chaining

⁷ Forward Chaining

⁸ F-Logic (frame Logic)

⁹ Karlsruhe University

¹⁰ Object Oriented Programming Language

¹¹ XOL (XML-based Ontology Language)

مبتنی بر ا.کی.بی.سی.^۱ می‌باشد. این زبان در سال ۱۹۹۹ میلادی به عنوان زبانی برای تبادل مدل‌های اطلاعاتی رسمی در حوزه انفورماتیک زیستی پیشنهاد شد. ایکس.ا.ال.^۲، یک زبان با محدودیت‌های بسیار است که فقط مفاهیم، طبقه‌بندی مفاهیم و روابط دودویی را تحت پوشش قرار می‌دهد و هیچ نظام استنتاجی به آن اضافه نشده است.

۲-۹-۷. شو (توسعه هستی‌شناسی اچ.تی.ام.ال. ساده):^۳ این زبان ابتدا به عنوان افزونه‌ای از اچ.تی.ام.ال. با هدف ادغام دانش معنایی ماشین خوان در اچ.تی.ام.ال. و یا سایر مدارک وب در سال ۱۹۹۶ میلادی در دانشگاه مریلند^۴ توسعه پیدا کرد. هدف از پیدایش این زبان، امکان پذیر ساختن جمع‌آوری اطلاعات معنادار در مورد صفحات و مدارک وب، بهبود سازوکارهای جستجو و گردآوری دانش می‌باشد. دستیابی به این اهداف طی دو مرحله عملی می‌شود: الف- تعریف یک هستی‌شناسی که رده‌بندی معتبری از اشیا را تشریح کرده و روابط معتبر میان آن‌ها را توصیف کند. ب- حاشیه‌نویسی صفحات اچ.تی.ام.ال. برای توصیف خودشان و یا سایر صفحات. شو می‌تواند از ترکیب اچ.تی.ام.ال. و یا ایکس.ام.ال. استفاده کند. این زبان قادر به بازنمایی مفاهیم، طبقه‌بندی آن‌ها، روابط چندگانه، نمونه‌ها و قواعد استنتاجی می‌باشد.

۲-۹-۸. آر.دی.اف. (چارچوب توصیف متن) و آر.دی.اف.اس.: یک مدل استاندارد برای تبادل داده در وب است. آر.دی.اف.، به وسیله جملاتی که شامل یک فاعل^۵ (منبع)، یک گزاره^۶ (ویژگی) و یک مفعول^۷ که می‌تواند داده یا مقدار باشد، بازنمایی می‌شود. بازنمایی ماشین خوان آر.دی.اف.، مبتنی بر ترکیب ایکس.ام.ال. است. برخی از عناصر آن، rdf:RDF و rdf:Description هستند که مناسب بازنمایی فراداده هستند ولی برای بازنمایی هستی‌شناسی‌ها کافی نیستند. همین کاستی موجب پیدایش طرح‌واره آر.دی.اف.^۷ شد. این زبان، زبانی برای توصیف واژگان با استفاده از ویژگی‌ها و کلاس‌ها می‌باشد. طرح‌واره آر.دی.اف. تعداد محدودی اجزاء از قبل تعریف شده را فراهم کرده و به کلاس‌ها و ویژگی‌های جدید امکان توصیف شدن می‌دهد. آر.دی.اف. و طرح‌واره آر.دی.اف.، با هم به کار گرفته می‌شوند و با نام آر.دی.اف.اس. شناخته می‌شوند. این زبان، مجموعه‌ای از اجزاء برای مدل‌سازی هستان‌شناسی‌های ساده فراهم می‌کند و از ساختار منطقی پشتیبانی نمی‌کند.

^۱ OKBC (Open Knowledge Base Connectivity)

^۲ SHOE (Simple HTML Ontology Extension)

^۳ University of Maryland

^۴ Subject

^۵ Predicate

^۶ Object

^۷ RDF-Schema

۹-۹-۲. **اُیل (لایه استنتاجی هستان شناسی)**^۱: از آن‌جا که آر. دی. اف. فاقد رسمی سازی لازم برای پشتیبانی از سازوکارهای استنتاجی بود، اُیل در طی پروژه آنتونالج^۲ و برای بیان شبکه معنایی منابع ایجاد شد. شبکه معنایی و سازوکار استنتاجی بر پایه منطق توصیفی بنا شده است. اُیل، منطق توصیفی (معانی رسمی و پشتیبانی استنتاجی)، نظام‌های مبتنی بر ساختار (برای مدل‌سازی) و زبان‌های وب (مبتنی بر ایکس.ام.ال. و آر.دی.اف.) را دربرمی‌گیرد. این زبان، به‌منظور توسعه یک زبان هستی‌شناسی برای وب که دارای شبکه معنایی واضح و پشتیبانی منطقی پیچیده باشد، به‌وجود آمد. هستی‌شناسی اُیل دارای سه سطح می‌باشد: الف- محمل هستی‌شناسی (اطلاعات در مورد هستی‌شناسی)، تعریف هستان‌شناسی (تعاریف اصطلاحات هستی‌شناسی)، ج- سطح اشیاء (نمونه‌ها) (گومز-پرز و همکاران ۲۰۰۴).

۹-۱۰-۲. **دی.ای.ام.ال. (زبان نشانه‌گذاری واسطه دارپا)**^۳: هم‌زمان با توسعه اُیل، آژانس پروژه‌های تحقیقاتی پیشرفته دفاعی^۴، با گسترش آر.دی.اف از طریق افزودن قابلیت خوانایی بیشتر، زبان دی.ای.ام.ال. را توسعه می‌داد. هدف از ایجاد این زبان، سهولت تعامل میان واسطه‌های نرم‌افزاری می‌باشد. از آن‌جا که دی.ای.ام.ال. بسیاری از خصوصیات اُیل را به ارث برده است، ویژگی‌های هر دو زبان یکسان هستند.

۹-۱۱-۲. **دی.ای.ام.ال. پلاس اُیل**^۵: این زبان حاصل ترکیب دی.ای.ام.ال. و اُیل می‌باشد و برای استفاده در وب طراحی شده است. دی.ای.ام.ال. پلاس اُیل با افزودن خصوصیت انسجام منطق توصیفی به ایکس.ام.ال. و آر.دی.اف به‌وجود آمده است. رویکرد آن شیء‌گرا بوده و با استفاده از کلاس‌ها و ویژگی‌ها، ساختار را توصیف می‌کند. این زبان، سازوکار بازنمون صریح خدمات، فرایندها و مدل‌های تجاری را فراهم می‌کند (کالیباتین و واسیلکاس ۲۰۱۱). دی.ای.ام.ال. پلاس اُیل از بسیاری جهات به اُیل شباهت دارد ولی بیشتر با آر.دی.اف سازگار است. اگرچه سازگاری با آر.دی.اف دارای مزایایی از قبیل استفاده مجدد از زیرساخت‌های موجود مبتنی بر آر.دی.اف و قابلیت انتقال هستی‌شناسی‌های مبتنی بر این زبان می‌باشد، اما با توجه به این که آر.دی.اف بر خلاف ایکس.ام.ال. برای تصریح دقیق ساختار نحوی طراحی نشده است، استفاده از آن کمی دشوار است.

۹-۱۲-۲. **اُ.دبلیو.ال. (زبان هستی‌شناسی وب)**: این زبان به پیشنهاد کنسرسیوم وب جهان‌گستر در سال ۲۰۰۱ به‌عنوان استاندارد جاری برای کدگذاری هستی‌شناسی‌ها تعیین شده است. اُ.دبلیو.ال.، حاصل

¹ OIL (Ontology Inference Layer)

² On-To-Knowledge

³ DAML (DARPA Agenet Markup Language)

⁴ DARPA (Defence Advanced Research Project Agency)

⁵ DAML+OIL

بازنگری در دی.ای.ام.ال. پلاس ایل می‌باشد. این زبان، زبان وب معنایی بوده و برای بازنمون دانش غنی و پیچیده در مورد اشیاء، گروه‌های اشیاء و روابط میان آن‌ها طراحی شده است. اُ.دبلیو.ال.، یک زبان محاسباتی مبتنی بر منطق است و به همین دلیل برنامه‌های کامپیوتری می‌توانند از آن به‌منظور افزایش یکپارچگی دانش یا تبدیل دانش ضمنی به دانش صریح استفاده کنند. این زبان، قابلیت‌های متعددی برای بیان معنا و مفهوم و نیز بازنمایی محتوای قابل تفسیر توسط ماشین بر روی وب را دارد و برای به‌کارگیری توسط برنامه‌های کاربردی که به‌جای بازنمایی صرف اطلاعات برای کاربران، نیازمند تحلیل محتوای در وب هستند، طراحی شده است (به نقل از ساو سا ۲۰۱۳). این زبان نسبت به ایکس.ام.ال. و آر.دی.اف. امکانات بیشتری برای بیان مفاهیم و معانی دارد و به‌دلیل قابلیت نمایش محتوای میان‌کنش‌پذیر رایانه‌ها در وب، برتر از سایر زبان‌ها است.

به‌دلیل پیچیدگی‌هایی که در ساختار اُ.دبلیو.ال. وجود دارد، سه زیر شاخه برای آن طراحی شده است (شکل ۲-۵):

زیرشاخه اول، اُ.دبلیو.ال. لایت^۱ می‌باشد که از سلسله‌مراتب طبقاتی و ویژگی‌های آن محدودیت کمتری دارند. مزیت این امر، سهولت درک (برای کاربران) و سهولت به‌کارگیری (برای تولیدکنندگان ابزار) می‌باشد. ایرادی که بر این نسخه وارد است، قابلیت بیان و بازنمایی محدود آن است.

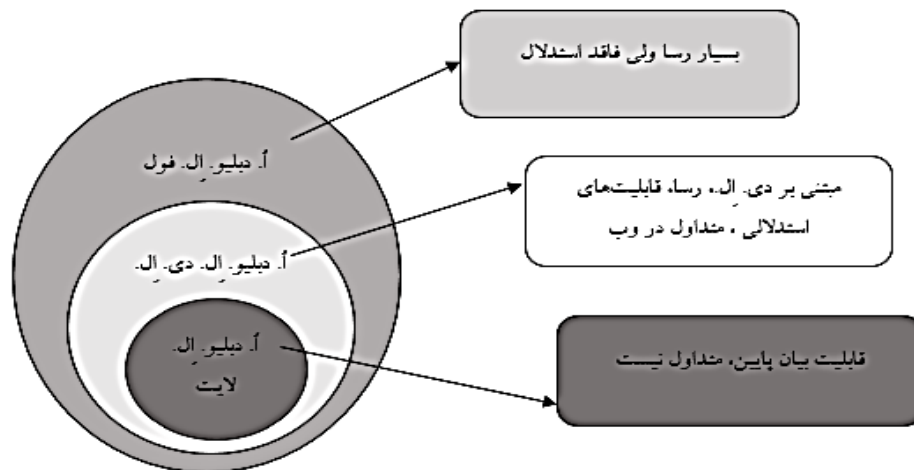
زیرشاخه دوم، اُ.دبلیو.ال. دی.ال.^۲ می‌باشد. این نسخه ساختار کامل زبان اُ.دبلیو.ال. به همراه تمام محدودیت‌ها مانند تفکیک نوع (یک کلاس نمی‌تواند هم‌زمان یک شخص یا ویژگی هم باشد). را شامل می‌شود و با منطق توصیفی مطابقت دارد. مزیت این نسخه، قابلیت پشتیبانی از استدلال است. در این نسخه سازگاری کامل با آر.دی.اف. لحاظ نشده است.

زیرشاخه سوم، اُ.دبلیو.ال. فول^۳ می‌باشد. این نسخه کامل‌ترین نسخه اُ.دبلیو.ال. است و بدون هیچ‌گونه تضمین محاسباتی از بیشترین میزان رسایی و آزادی نحوی آر.دی.اف. پشتیبانی می‌کند. این نسخه از لحاظ نحوی و هم از لحاظ معنایی با آر.دی.اف. سازگاری کامل دارد.

^۱ OWL Lite

^۲ OWL DL

^۳ OWL Full



شکل ۲-۵. زیرشاخه‌های ا. دبلیو. ال. (کالیباتین و واسیلکاس ۲۰۱۱)

۲-۱۰. مقایسه زبان‌های هستی‌شناسی

در جدول ۲-۳، با هدف به تصویر کشیدن شباهت‌ها و تفاوت‌های رایج‌ترین زبان‌های هستی‌شناسی، به مقایسه آن‌ها پرداخته شده است. برای هر سلول علامت ✓ نمایانگر پشتیبانی زبان از عنصر موردنظر و علامت ✕ نمایانگر عدم پشتیبانی از عنصر موردنظر است. با توجه به اطلاعات جدول، وجود برخی تفاوت‌ها میان زبان‌های سنتی هستی‌شناسی و زبان‌های نشانه‌گذاری آن مشخص می‌شود. همچنین، برخی از آن‌ها هستی‌شناسی‌های سنگین‌وزن و سبک‌وزن را بازنمایی می‌کنند. زبان هستی‌شناسی سنگین‌وزن، بازنمونی از اصول موضوعی رسمی، توابع و سایر اجزا می‌باشد حال آن‌که زبان هستی‌شناسی سبک‌وزن، مفاهیم، طبقه‌بندی مفاهیم و روابط آن‌ها را بازنمایی می‌کند. اجزای بازنمون دانش می‌توانند توسط زبان‌های سنتی مانند آنتولینگوا، لوم و ا.سی.ام.ال. مدل‌سازی شوند. بسیاری از زبان‌های هستی‌شناسی، به استثنای آر.دی.اف.اس. و شو، قابلیت بازنمایی مفاهیم و توصیف ویژگی‌های آن‌ها را دارند. در حقیقت اصول موضوعی تضاد، اجتماع و افتراق توسط بسیاری از زبان‌ها مانند آنتولینگوا، ایل، دی.ای.ام.ال. پلاس ایل و ا.دبلیو.ال. فراهم شده‌اند. رابطه دوتایی بین مفاهیم در تمام زبان‌ها بازنمایی می‌شوند در حالی که روابط معنایی سلسله‌مراتبی در ا.کی.بی.سی.، اف-لاجیک، شو و ایکس.ا.ال. قابل بازنمایی نیستند. ا.دبلیو.ال. قابلیت توصیف کلاس محدودیت را دارد. کلاس‌های بدون نام می‌توانند بر مبنای محدودیت‌های مقدار برای یک ویژگی خاص کلاس، توصیف شوند (ساوسا ۲۰۱۳).

جدول ۲-۳. مقایسه زبان‌های بازنمایی هستان‌شناسی‌ها

اصر		مفهوم		صفت				نمونه		اصول موضوعی				روابط معنایی			
زبان‌ها	تعریف	نام	تعریف	نام	نوع پایه	ویژگی نمونه	نام	نمونه مفهومی	رابطه	اجتماع	افتراق	تضاد	پوشش	دودویی	چندگانه	نوعی از	سلسله‌مراتبی
کی. آی. اف.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
آنتولینگوا	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
لوم	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓
اُ. کی. بی. سی.	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	✓	×	×	×
اُ. سی. ام. ال.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	×	✓
اف-لاجیک	✓	✓	✓	✓	✓	✓	✓	×	✓	×	×	×	×	✓	×	×	×
شو	✓	✓	×	✓	✓	✓	✓	×	✓	×	×	×	✓	✓	✓	×	×
ایکس. اُ. ال.	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	✓	×	×	×
آر. دی. اف. اس	✓	✓	×	×	×	✓	✓	✓	✓	×	×	×	✓	✓	×	×	✓
اُیل	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	×	×	✓
دی. ای. ام. ال. پلاس اُیل	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	×	×	✓
اُ. دبلیو. ال.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	×	✓

موضوعی، روابط دوتایی و نمونه‌ها تنها اجزایی هستند که در تمام زبان‌های معرفی شده امکان بازنمایی دارند. برخی تفاوت‌ها در زبان‌ها برای بازنمایی اصول موضوعی مفهومی وجود دارد. آنتولینگوا، لوم، اُ. سی. ام. ال.، اویل، دی. ای. ام. ال.، پلاس اُیل و اُ. دبلیو. ال. رساتر هستند و امکان ایجاد تجزیه‌های جامع و متضاد زیرکلاس‌های یک مفهوم را فراهم می‌کنند.

در آنتولینگوا و شو، امکان ایجاد روابط چندگانه اختیاری وجود دارد. در مابقی زبان‌ها، این روابط از طریق تجزیه آن‌ها به روابط دوتایی قابل بازنمایی هستند.

توابع، به سادگی در آنتولینگوا، لوم، اُ. سی. ام. ال.، اویل، دی. ای. ام. ال.، پلاس اُیل و اُ. دبلیو. ال. توصیف می‌شوند ولی در اف-لاجیک تنها از طریق تعریف یک رابطه و یک اصل موضوعی اضافی، که تعداد مقادیری که می‌تواند داشته باشد را محدود می‌کند، امکان پیدایش دارد.

اصول موضوعی رسمی، قدرتمندترین ابزار بازنمون دانش در هستی‌شناسی‌ها هستند و معمولاً برای بازنمایی قسمت‌هایی از دانش به کار می‌روند که امکان بازنمایی توسط سایر ابزار ابتدایی زبان را

ندارند. اصول موضوعی رسمی در آنتولینگوا، لوم، اُ.سی.ام.ال. و اف-لاجیک قابل توصیف هستند. قواعد، فقط در لوم و اُ.سی.ام.ال. قابل توصیف هستند و فرایندها فقط در آنتولینگوا، لوم و، اُ.سی.ام.ال. توصیف می‌شوند.

در نهایت، هر یک از زبان‌های هستی‌شناسی، مزایا و معایبی دارند و به همین دلیل برای انتخاب زبان مناسب باید هدف و کاربرد هستی‌شناسی مورد توجه قرار داده شود. از آن‌جا که زبان هستی‌شناسی پایه و اساس نظام‌های دانش مبتنی بر هستی‌شناسی هستند، زبان انتخاب شده نه تنها باید قابلیت بالایی برای بیان و به‌کارگیری داشته باشد، بلکه ساختار آن باید برای رایانه‌ها، پردازشگرها و برنامه‌های کاربردی قابل فهم باشد (ترابی ۱۴۰۰).

۲-۱۱. ابزارهای ساخت هستی‌شناسی

هستی‌شناسی‌ها برای ایجاد علاوه بر زبان‌ها نیاز به محیط‌هایی برای اجرا دارند که از مراحل مختلف ساخت آن‌ها پشتیبانی کند و اجازه ویرایش، روزآمدسازی و استفاده پایدار و مجدد از آن‌ها را بدهد. از دیدگاه گومز-پرز و همکاران (۲۰۰۴) ساخت هستی‌شناسی‌ها پیچیده و زمان‌بر است بنابراین طراحان و سازندگان حتی با مهارت‌های بالادر زبان‌های هستی‌شناسی، برای پیاده‌سازی آن نیاز به ابزارهایی دارند. در میانه دهه ۱۹۹۰ میلادی اولین محیط طراحی هستی‌شناسی به‌وجود آمد که در واقع رابط کاربری را برای انجام فعالیت‌های اصلی فرایند ایجاد هستی‌شناسی از جمله مفهوم‌سازی، اجرا، کنترل یکپارچگی و مستندسازی فراهم می‌نمود. مانند زبان‌های هستی‌شناسی، ابزارهای ساخت آن هم در حال توسعه و تکمیل شدن هستند. آنان ابزارهای به‌کار رفته در ساخت هستی‌شناسی را به گروه‌های زیر تقسیم کردند:

□ ابزارهای توسعه هستی‌شناسی: شامل ابزارهایی که برای ساخت یک هستی‌شناسی از پایه

به‌کار می‌روند. این ابزارهای علاوه بر قابلیت ویرایش و تورق، از مستندسازی و انتقال^۱ بین

زبان‌ها، قالب‌ها و نسخه‌های گرافیکی مختلف پشتیبانی می‌کنند؛

□ ابزارهای ارزیابی هستی‌شناسی: این ابزارها برای ارزیابی محتوای هستی‌شناسی‌ها و فناوری‌های

مرتبط با آن کاربرد دارند. این ابزارها برای کاهش مشکلات در هنگام ادغام و استفاده از

هستی‌شناسی‌ها و فناوری‌های وابسته به آن‌ها در سیستم‌های اطلاعاتی استفاده می‌شوند؛

¹ Export/ Import

- ابزارهای ادغام و تنظیم هستی‌شناسی: ابزارهایی هستند که برای حل مسائل مربوط به ادغام و یکپارچه‌سازی هستی‌شناسی‌های مختلف در یک دامنه به کار گرفته می‌شوند؛
- ابزارهای نشانه‌گذاری: از طریق این ابزارها کاربران می‌توانند مصادیق مربوط به مفاهیم و روابط را به هستی‌شناسی به صورت نیمه خودکار یا خودکار در صفحات وب اضافه و نگهداری کنند. این ابزارها غالباً در بستر وب معنایی ظاهر شده‌اند؛
- ابزارهای پرس‌وجو و موتورهای استنتاج هستی‌شناسی‌ها: این ابزارها اجازه انجام پرس‌وجو و را به سادگی داده و استدلال مربوط به آن‌ها را انجام می‌دهند. این ابزارها معمولاً تا حد زیادی به زبان به کار رفته در هستی‌شناسی بستگی دارند؛
- ابزارهای یادگیری هستی‌شناسی: این ابزارها می‌توانند به صورت خودکار یا نیمه خودکار هستی‌شناسی‌ها را از متون طبیعی مانند منابع نیمه ساختاریافته و پایگاه‌های داده به وسیله فنون تحلیل زبان و یادگیری ماشین استخراج کنند.

از جمله ابزارهای ذکر شده می‌توان به آنتولوژی سرور^۱، آنتو ساروس^۲، وب آنتو^۳، اویل ای دی^۴، کمپ تولز^۵، پروتزه^۶، وب پروتزه^۷ و گیت^۸ اشاره نمود که هر یک دارای ویژگی‌ها و کاربردهایی هستند که به فناوری‌ها و زبان‌های مورد استفاده در آن‌ها وابسته است. ساو سا (۲۰۱۳) ضمن مقایسه ابزارهای یاد شده عقیده دارد از منظر سهولت استفاده، پروتزه و وب آنتو نمایش گرافیکی برای ارائه داده‌ها دارند که استفاده از آن‌ها را آسان‌تر می‌کند. همچنین اغلب این ابزارها مبتنی بر منطق مرتبه اول هستند. ابزارهایی مانند آنتو ساروس و وب آنتو از لحاظ واژگانی قابلیت پرس‌وجو را ندارند در حالی که پروتزه این امکان را فراهم کرده است. پروتزه همچنین ویژگی‌هایی دارد که موجب برتری و استفاده بیشتر آن در مقایسه با دیگر ابزارها شده است. از جمله این ویژگی‌ها می‌توان به امکان نمایش کلاس‌ها، بخش‌ها، روابط و اصول اشاره کرد. همچنین رابط کاربری گرافیکی پروتزه استفاده از آن را آسان نموده است. از دیگر ویژگی‌های پروتزه پشتیبانی این ابزار از چندین زبان است که بر همین

^۱ Ontology server

^۲ OntoSaurus

^۳ WebOnto

^۴ OilEd

^۵ Camp tools

^۶ Protege

^۷ Web Protege

^۸ GATE

اساس می‌تواند خروجی‌هایی در قالب آر.دی.اف.اس، ایکس.ام.ال، اف-لاجیک، جاوا^۱ و کلیس^۲ ارائه کند. این ابزار همچنین رایگان، مستقل، منع باز و خوش ساختار است و موتور استدلالی قوی برای استنتاج و تفسیر پرسش‌ها دارد. علاوه بر این، افزونه‌های قابل گسترش آن قوی و کاربردی است که قابلیت جستجوی آسان در هستی‌شناسی را فراهم می‌کند.

۲-۱۲. رویکردهای طراحی هستی‌شناسی

مسئله دیگر در ساخت هستی‌شناسی، مرتبط با روش‌های مقوله‌بندی و توسعه مفاهیم است که در ادامه به آن‌ها پرداخته می‌شود. آسکولد و گرونیگر^۳ (۱۹۹۶) سه رویکرد برای مدل‌سازی هستی‌شناسی‌ها به بحث گذاشته‌اند:

□ رویکرد بالا به پایین^۴ که در آن فرایند توسعه با تعریف کلی‌ترین مفاهیم دامنه آغاز شده و در ادامه با خاص‌تر شدن مفاهیم ادامه پیدا می‌کند. این رویکرد منجر به کنترل بهتر سطوح پرداختن به جزئیات می‌شود. در هر صورت شروع با مفاهیم کلی می‌تواند باعث انتخاب و اعمال اختیاری مقوله‌های سطح بالا شود. این رویکرد نشأت گرفته از نظم طبیعی مقولات نیست بنابراین خطر پایداری پایین مدل‌های مبتنی بر آن وجود دارد که به نوبه خود می‌تواند منجر به دوباره کاری و انجام تلاش‌های مجدد شود. همچنین تأکید بر مقولات عام و کلی می‌تواند باعث استدلال متفاوت و از دست دادن ویژگی توارث مشترک در شبکه پیچیده مفاهیم مرتبط به هم شود؛

□ رویکرد پایین به بالا^۵ که با فرایند توسعه تعاریف خاص‌ترین کلاس‌ها و بدون توجه به سلسله‌مراتب آغاز شده و با گروه‌بندی این کلاس‌ها در مفاهیم کلی‌تر ادامه پیدا می‌کند. این رویکرد منجر به ارائه سطح بالایی از جزئیات می‌شود و تلاش‌های فزاینده در این جهت اشتراک بین مفاهیم مرتبط را با مشکل مواجه کرده و خطر عدم همخوانی مقولات را افزایش می‌دهد که به نوبه خود منجر به دوباره کاری و نیاز به تلاش‌های چندباره خواهد شد؛

□ رویکرد میانه^۶ که در آن بر هماهنگی و ایجاد تعادل به لحاظ سطوح جزئیات تأکید می‌شود. این رویکرد که در آن از دانش از پیش موجود برای تعریف مفاهیم سطح بالا استفاده

¹ JAVA

² CIPS

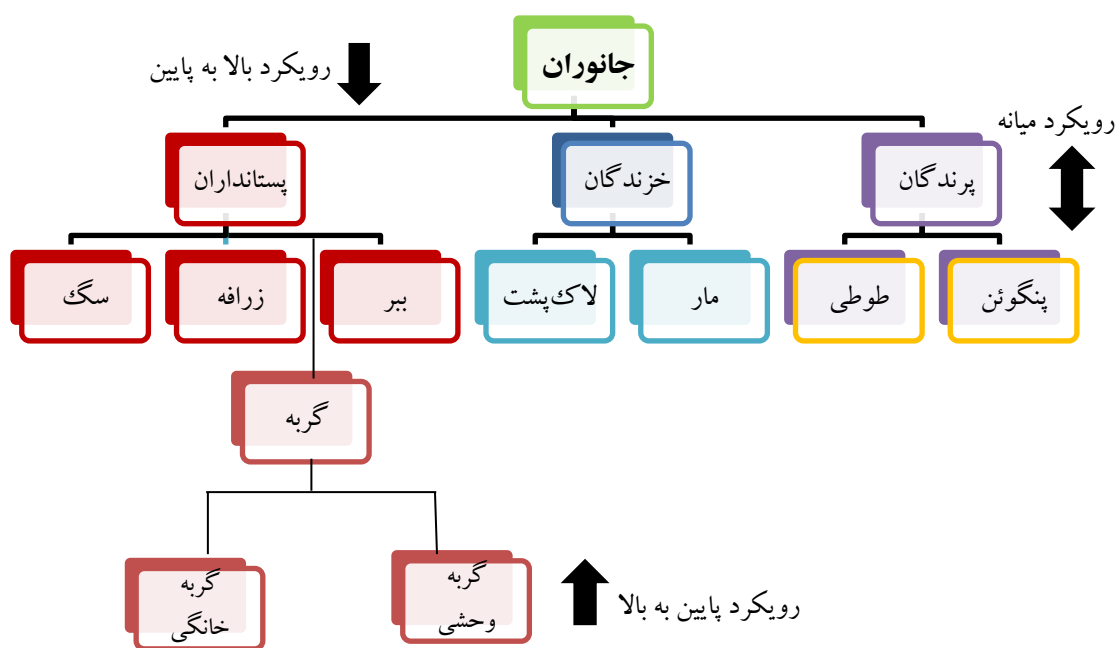
³ Uschold and Gruninger

⁴ Top-down approach

⁵ Bottom-up approach

⁶ Middle out approach

می‌شود؛ جزئیات تنها در مواقع لزوم و از طریق خاص نمودن مفاهیم پایه ارائه می‌شوند. بنابراین با شروع کردن از مهم‌ترین مفاهیم در وهله اول و تعریف مفاهیم سطح بالا از این جهت که این مقولات به صورت طبیعی و مطابق با نظم طبیعی موجود ارائه می‌شوند از تلاش‌های مجدد و دوباره کاری‌ها اجتناب می‌شود. بنابراین حاصل کار نسبت به دو رویکرد دیگر دارای ثبات بیشتر بوده و نیاز به دوباره کاری کمتری دارد (نوی و مک گینس^۱ ۲۰۰۱). شکل ۲-۶ با ذکر مثالی ساده نمونه‌ای از کاربرد رویکردهای متفاوت را نشان می‌دهد.



شکل ۲-۶. نمایش رویکردهای سه‌گانه طراحی هستی‌شناسی

در مجموع و با توجه به آنچه یاد شد، آسکولد و گرونینگر (۱۹۹۶) معتقدند در زمینه توسعه هستی‌شناسی، رویکرد بالا به پایین در کنترل جزئیات عملکرد بهتری دارد. هرچند این رویکرد از مفاهیم و سطوح بالای دانش آغاز می‌شود و برخی مفاهیم تصادفی از پیش تعریف شده در سطوح بالا را دربرمی‌گیرد، که این مسئله منجر به ثبات کمتر در هستی‌شناسی طراحی شده می‌شود. در مثال ذکر شده این رویکرد از طبقه اصلی یعنی جانوارن شروع به مدل‌سازی مفاهیم نموده و به سمت مصادیق یعنی حیوانات هر طبقه پیش می‌رود. از سوی دیگر رویکرد پایین به بالا به تلاش بیشتری احتیاج دارد و فهم اشتراک بین مفاهیم مرتبط در آن مشکل است. علاوه بر آن، در این رویکرد خطر عدم یکپارچگی در طول فرایند توسعه و در نهایت دوباره کاری وجود دارد. در این رویکرد همان‌طور که در شکل ۳-۶ نیز قابل مشاهده است حرکت از مفاهیم پایین مانند نام حیوانات به سمت طبقات اصلی

¹ Noy and McGuinness

مانند پرندگان، خزندگان و پستانداران انجام می‌شود. در رویکرد میانه یا ترکیبی مفاهیم مهم‌تر شناسایی و بدین ترتیب طبقه‌بندی‌های بالا در آن با ثبات بیشتری و با نیاز به تلاش و دوباره کاری کمتر ظاهر می‌شوند. برای نمونه، طبقات مهم جانوران شامل پرندگان، خزندگان و پستانداران به‌عنوان مهم‌ترین مفاهیم این سلسله‌مراتب نقطه شروع هستند که سطوح بالای آن‌ها شامل جانوران و سطوح پایین شامل مصادیق از آن‌ها نشأت می‌گیرند.

۲-۱۳. روش‌شناسی‌های مهندسی هستی‌شناسی

تعداد قابل توجهی از پژوهش‌ها اصول طراحی، روش‌ها و ابزارهای ساخت و ارزیابی هستی‌شناسی‌ها را مورد مطالعه قرار داده‌اند. روش‌شناسی‌ها در واقع راهنمایی برای توسعه هستی‌شناسی‌ها و انتخاب فنون مناسب برای هر مرحله از فرایند ساخت هستند. از دهه ۱۹۹۰ میلادی روش‌شناسی‌های متعددی برای ساخت هستی‌شناسی‌ها ارائه شد. اغلب آن‌ها رویکردهای متفاوتی دارند، بعضی برای ایجاد هستی‌شناسی از پایه طراحی شده‌اند و برخی برای استفاده مجدد از هستی‌شناسی‌های موجود. فرایند ساخت هستی‌شناسی که به‌طور گسترده به‌عنوان مهندسی هستی‌شناسی شناخته شده بیش از آنکه مجموعه فعالیت‌های مهندسی باشد، ساخت یک مصنوع است. علی‌رغم وجود روش‌های متعدد، در این زمینه فقدان راهنماها و روش‌های مورد توافق که ساخت هستی‌شناسی را پشتیبانی کنند، از دغدغه‌های پژوهشگران این حوزه است (ساو سا ۲۰۱۳). در ادامه به برخی از این روش‌ها اشاره می‌شود.

روش سایک^۱

روش سایک^۲ توسط لنات^۳ در شرکت فناوری رایانه و میکروالکترونیک^۴ در دهه ۱۹۸۰ در پروژه‌ای برای ایجاد یک پایگاه دانش ارائه شد. این روش برای اجرا از زبان سایک.ال استفاده کرد. بر همین اساس، لنات مراحل را برای طراحی هستی‌شناسی ارائه داد:

فراهم‌آوری دانش: در این مرحله مقالات و قطعات دانش که می‌توانند به‌صورت دستی کدگذاری شوند فراهم‌آوری می‌شوند. این مرحله بیشتر به‌صورت دستی انجام می‌شود تا با فنون یادگیری ماشینی و سیستم‌های پردازش زبان طبیعی.

جستجو و بازنمون دانش فراهم شده: در این مرحله دانش پایه مورد نیاز کاربران یک حوزه براساس

^۱ CYC

^۲ Cyc project

^۳ Doug Lenat

^۴ MCC (Microelectronics and Computer Technology Corporation)

مرحله قبل جستجو و بازنمون می‌شود.

ارزیابی دانش فراهم‌شده: در این مرحله محتوای مقالات و متون فراهم‌آوری شده بررسی و پالایش می‌شود.

تشخیص سؤالات کاربران: در این مرحله برای این که هستی‌شناسی قادر به پاسخگویی به کاربران باشد، سؤالات احتمالی آن‌ها از طریق مطالعه متون حوزه مربوطه مشخص می‌شوند.

کدگذاری: این مرحله شامل کدگذاری دانش ذخیره شده در پایگاه دانش سایک توسط ابزارهای این کار است (سایک ۲۰۱۶).

روش آسکولد و کینگ^۱

می‌توان گفت اولین روش ایجاد هستی‌شناسی در سال ۱۹۹۵ توسط آسکولد و کینگ ارائه شد که در سال ۱۹۹۶ توسط آسکولد و گرونینگر گسترش پیدا کرد. آن‌ها با ذکر این نکته که روش یاد شده دارای کاستی‌هایی در برقراری روابط در بین مراحل است، آن را به‌عنوان راهنمایی برای طراحی و اجرای هستی‌شناسی معرفی کردند (آسکولد و گرونینگر ۱۹۹۶، نقل در ساو سا ۲۰۱۳). روش یاد شده شامل مراحل زیر است:

گام اول: شامل هدف‌گذاری ساخت هستی‌شناسی، دامنه و کاربران و توسعه‌دهندگان آن است.

گام دوم: که شامل ساخت هستی‌شناسی است و بر اساس مراحل زیر صورت می‌گیرد:

□ فراهم‌آوری دانش: در این مرحله دانش مورد نیاز برای ساخت هستی‌شناسی از طریق تعیین مفاهیم و روابط دامنه موضوعی و ایجاد تعاریف صحیح مفاهیم و روابط مورد نظر فراهم‌آوری می‌شود.

□ تشخیص اصطلاحات: در این مرحله اصطلاحات مرتبط با مفاهیم و روابط بین آن‌ها تعیین و بر اساس آن‌ها در مورد مفاهیم اجماع صورت می‌گیرد.

□ کدگذاری: در این مرحله دانش فراهم‌شده که به‌صورت مفاهیمی مورد اجماع درآمده از طریق زبان‌های رسمی کدگذاری می‌شود.

□ ادغام هستی‌شناسی‌های موجود: این مرحله به استفاده از هستی‌شناسی‌های موجود برای فراهم‌آوری دانش پایه یا حتی کدگذاری آن اشاره دارد.

¹ Uschold & King

گام سوم: در این مرحله هستی‌شناسی ایجاد شده مورد ارزیابی قرار می‌گیرد.

گام چهارم: مستندسازی فرایند ساخت هستی‌شناسی به‌عنوان راهنماهای ایجاد آن در این مرحله صورت می‌گیرد.

روش توو^۱

در سال ۱۹۹۵ به نقل از گومز پرز و همکاران (۲۰۰۴) رویکردی رسمی برای ایجاد و ارزیابی هستی‌شناسی ارائه دادند. این روش بر اساس منطق مرتبه اول است و برای ساخت هستی‌شناسی سازمان مجازی تورنتو یا توو استفاده شد. توو شامل مجموعه‌ای از هستی‌شناسی‌ها برای توصیف ویژگی‌های پروژه‌های تجاری است. این روش شامل مراحل زیر است:

تشخیص سناریوهای مرتبط با انگیزه: در این مرحله سناریو انگیزشی که شامل مسائل طرح نشده و شکاف‌ها در هستی‌شناسی‌های موجود است ایجاد می‌شود. این سناریو تأثیر به‌سزایی در هدایت مسیر طراحی هستی‌شناسی داشته و راه‌حل‌های ممکن را برای مسائل مورد نظر فراهم می‌کند. منظور از راه‌حل معانی غیررسمی اشیاء و روابط بین آنهاست.

ساخت پرسش در ارتباط با سناریوی مد نظر: در این مرحله لازم است تا سؤالاتی در مورد کفایت و شایستگی سناریوی مورد نظر مطرح شود. این سؤالات در هستی‌شناسی با اصطلاحات رسمی بازنمون می‌شوند و تعهد هستی‌شناسی را در توسعه هستی‌شناسی مد نظر ارزیابی می‌کنند.

استفاده از زبان رسمی برای تخصیص اصطلاح‌شناسی در هستی‌شناسی: سؤالات غیررسمی شایستگی که در مرحله قبل بحث شد، با هدف استخراج محتوای هستی‌شناسی و تخصیص اصطلاحات به زبان رسمی است. بر این مبنا بازنمون رسمی مفاهیم موجودیت‌ها و روابط از طریق یک زبان هستی‌شناسی صورت می‌گیرد.

تعریف رسمی سؤالات شایستگی: در این مرحله سؤالات طرح شده در مراحل قبل به صورت رسمی تعریف می‌شوند.

استفاده از منطق مرتبه اول برای تشخیص اصول: در این مرحله اصول باید تعاریف مفاهیم و محدودیت‌ها را با استفاده از منطق مرتبه اول مشخص سازند.

تشخیص جامعیت: این مرحله شامل ایجاد شرایطی برای تشخیص کامل بودن هستی‌شناسی می‌شود. بدین معنا که موقعیت‌هایی تعریف می‌شود تا پاسخ‌های ارائه شده به سؤالات کامل شود (ساوسا

^۱ TOVE (Toronto Virtual Enterprise)

(۲۰۱۳).

روش کاکتوس^۱

به نقل از گومز پرز و همکاران (۲۰۰۴) وی در سال ۱۹۹۶ روشی استفاده شده برای ساخت هستی‌شناسی در حوزه شبکه‌های الکترونیکی را به‌عنوان بخشی از پروژه اسپریت کاکتوس مطرح کردند. در این پروژه روش پیشنهادی، طراحی هستی‌شناسی مبتنی بر یک پایگاه دانش کاربردی به‌وسیله پردازش انتزاع است که در واقع پیروی از رویکرد پایین به بالاست. این روش‌شناسی در ضمن پروژه‌ای با هدف امکان‌سنجی استفاده مجدد از دانش در سیستم‌های فنی پیچیده و حمایت هستی‌شناسی از آن بود. این روش‌شناسی قابلیت استفاده مجدد از هستی‌شناسی‌های موجود یا تلفیق چند هستی‌شناسی برای رسیدن به برنامه کاربردی را تحت پوشش قرار می‌دهد. در این روش‌شناسی سه فرایند زیر برای رسیدن به هستی‌شناسی مورد نظر انجام می‌شود:

فرایند اول: توصیف مشخصات برنامه کاربردی؛

فرایند دوم: طراحی مقدماتی هستی‌شناسی بر اساس هستی‌شناسی سطح بالای مرتبط و؛

فرایند سوم: شامل اصلاح و ساختاردهی هستی‌شناسی است.

این روش درواقع پایگاه‌های دانش موجود را برای ایجاد و استقرار هستی‌شناسی‌ها تجهیز می‌نماید. این هستی‌شناسی بر اساس الگوهای دانش موجود ایجاد می‌شود و رویکردی از پایین به بالا دارد.

روش سن‌سس^۲

همان‌طور که به نقل از گومز پرز و همکاران (۲۰۰۴) اشاره می‌کنند، این روش برای حمایت از ایجاد هستی‌شناسی مبتنی بر حوزه از تولید شالوده آن تا برگرفتن هستی‌شناسی از هستی‌نگاری‌های کلان به‌منظور کمک به ایجاد اسکلت هستی‌شناسی‌های جدید از یک هستی‌شناسی بزرگ طراحی شده است. فرایند اصلی در این هستی‌شناسی اتصال دادن مفاهیم دامنه به هستی‌شناسی سن‌سس است که فرایندهای زیر را دربرمی‌گیرد:

تشخیص اصطلاحات هسته: در این مرحله اصطلاحات کلیدی مربوط به یک دامنه موضوعی مشخص می‌شوند.

برقراری ارتباط بین اصطلاحات هسته و سن‌سس به صورت دستی: این برقراری ارتباط توسط ابزار

¹ KACTUS

² SENSUS

آنتوساروس^۱ و به‌منظور گسترش دامنه پوشش هستی‌شناسی انجام می‌شود.

اضافه‌نمودن ریشه‌ها و مسیرها: این مرحله با جمع‌آوری همه مفاهیمی که باید به ریشه‌های سن سس مرتبط شوند انجام می‌گیرد.

اضافه‌نمودن اصطلاحات جدید دامنه: اصطلاحاتی که مرتبط با دامنه هستند اما هنوز پوشش داده نشده‌اند در این مرحله به سایر اصطلاحات اضافه خواهند شد.

اضافه‌نمودن درخت‌های فرعی تکمیلی: در صورت مرتبط‌بودن گره‌ها، در این مرحله لازم است تا درخت‌های فرعی به هستی‌شناسی نهایی اضافه شوند (گومز پرز و همکاران ۲۰۰۴).

مت‌آنتولوژی^۲

فرایندهای ساخت هستی‌شناسی در رویکردها و روش‌های مختلف شامل مراحل است که علی‌رغم تفاوت‌های که دارند در مراحل اصلی غالباً از گام‌های مشترکی پیروی می‌کنند. یکی از پراستنادترین منابع در این زمینه که مورد استفاده بسیاری از پژوهش‌ها قرار گرفته است گام‌های ساخت مربوط به روش مت‌آنتولوژی است. در سال ۱۹۹۷ فرناندز-لوپز و همکاران چرخه‌ای برای ساخت هستی‌شناسی ارائه دادند که به‌عنوان روشی خوش ساخت برای ساخت هستی‌شناسی از پایه محسوب می‌شود. مت‌آنتولوژی بر مبنای تجارب حاصل از ایجاد هستی‌شناسی در حوزه شیمی ایجاد شد. بر همین اساس فرناندز و همکاران به چرخه ایجاد یک هستی‌شناسی شامل مراحل تخصیص، مفهوم‌سازی، فرموله‌سازی، یکپارچه‌سازی، اجرا و نگهداری اشاره می‌کنند. این چرخه به سازنده هستی‌شناسی امکان می‌دهد تا در صورتی که بعضی تعاریف از قلم افتاده یا اشتباه باشند، از یک مرحله به مرحله پیشین بازگردد. مت‌آنتولوژی چرخه ساخت هستی‌شناسی از پایه را بر اساس مراحل زیر ذکر می‌کند:

□ تعیین و تشخیص: تشخیص هدف و دامنه هستی‌شناسی. هدف از این مرحله پاسخ به این سؤال

است: هدف از ساخت این هستی‌شناسی چیست؟ و در مورد دامنه پاسخ به این سؤال که

کاربری و کاربران احتمالی این هستی‌شناسی چیست؟

□ مفهوم‌سازی: توصیف مفاهیم در قالب یک مدل مفهومی و ساخت یک هستی‌شناسی که

اهداف مشخص شده مرحله اول را تأمین کند. روش‌شناسی‌های مختلف برای این مرحله

مدل‌های مختلفی از مفهوم‌سازی را پیشنهاد می‌کنند. از مدل‌های سطحی و غیررسمی مانند

¹ OntoSarus

² Methontology

نقشه‌های ذهنی تا مدل‌های نیمه‌رسمی مانند دیاگرام‌های روابط دودویی یا واژه‌نامه مفاهیم (فرناندز لوپز و همکاران ۱۹۹۹) مدل مفهومی یک هستی‌شناسی شامل مفاهیم یک دامنه موضوعی و روابط بین آن‌هاست. روابط ایجاد شده پیوند بین گروه‌های مفاهیم را ارتقا می‌دهند؛

□ رسمی‌سازی^۱: تبدیل توصیفات مفهومی به مدل‌های رسمی که در واقع تبدیل خروجی مرحله دوم به فرمی رسمی‌تر است. در این مرحله معمولاً مفاهیم از طریق اصول و قواعدی تعریف می‌شوند که تفاسیر ممکن از مفاهیم را محدود می‌سازند. معمولاً مفاهیم در این مرحله به صورت سلسله‌مراتبی و از طریق روابط ساختاری مانند نمونه‌ای از^۲ شامل کلاس‌ها، کلاس‌های فرعی، نمونه‌ها و بخشی از^۳ سازماندهی می‌شوند؛

□ اجرا^۴: پیاده‌سازی هستی‌شناسی رسمی شده با یک زبان بازنمون در این مرحله صورت می‌گیرد. برای این منظور مهندسان هستی‌شناسی زبان بازنمایی را انتخاب و مدل رسمی را به این زبان پیاده‌سازی می‌کنند؛

□ نگهداری^۵: این مرحله روی روزآمد سازی و اصلاح و ویرایش هستی‌شناسی تمرکز دارد؛

□ فراهم‌آوری دانش^۶: دانش مربوط به حوزه موضوعی هستی‌شناسی با استفاده از فنون انتخاب دانش از طریق خبرگان موضوعی یا روش‌های کتابشناختی و مبتنی بر متن در این مرحله انجام می‌شود. از جمله روش‌های مرسوم در این مرحله استفاده از تکنیک طوفان فکری، مصاحبه، پرسشنامه، تحلیل متن و روش‌های استقرائی است؛

□ ارزیابی^۷: در این مرحله کیفیت هستی‌شناسی ساخته شده از منظر فنی بررسی می‌شود؛

□ مستندسازی^۸: این مرحله شامل ارائه گزارش از تمامی مراحل و نحوه اجرای آن‌هاست. تشریح آنچه انجام شده و دلیل آن و همچنین مستندات مربوط به اصطلاحات ارائه شده در

¹ Formalization

² is-a

³ part of

⁴ Implementation

⁵ Maintenance

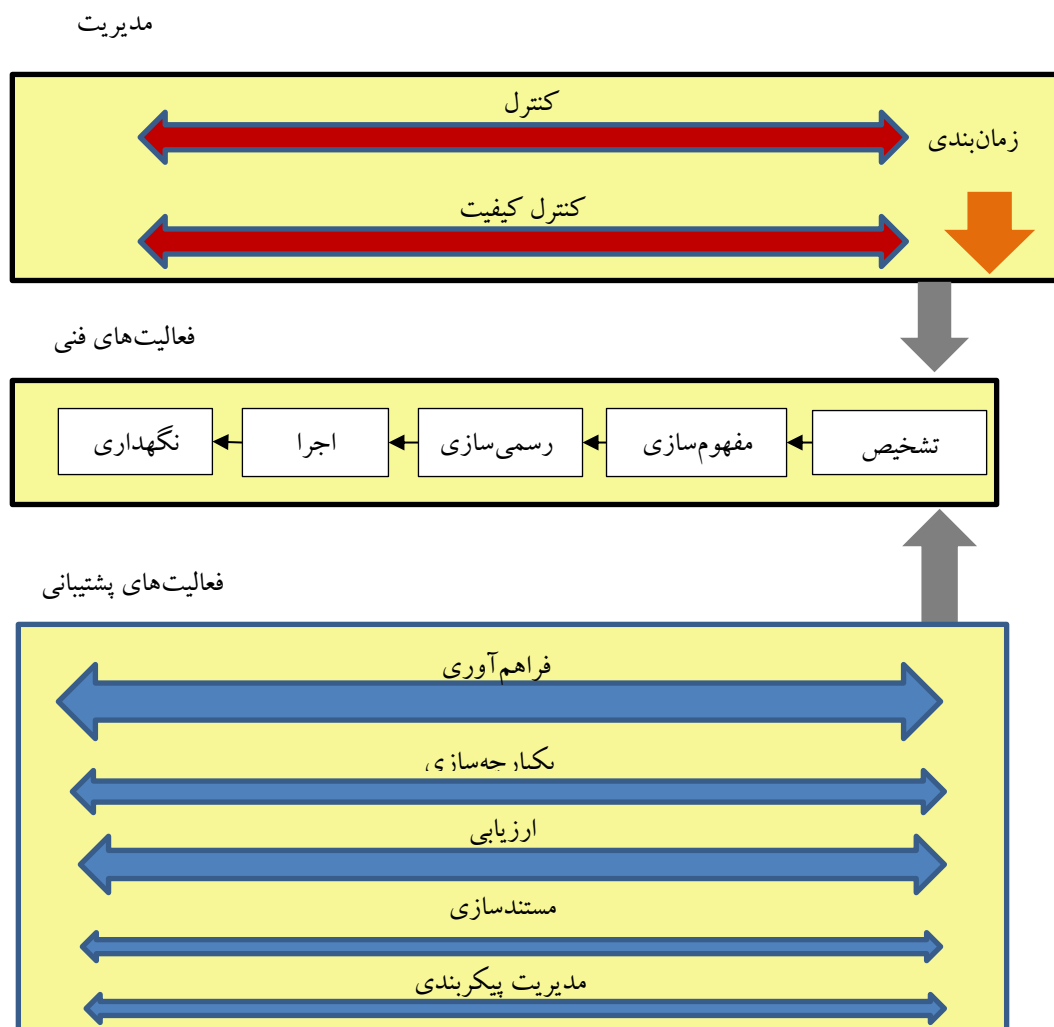
⁶ Knowledge acquisition

⁷ Evaluation

⁸ Documentation

هستی‌شناسی از ملزومات این مرحله است. مستندسازی هم به شفافیت هستی‌شناسی کمک می‌کند و هم در تسهیل نگهداری، استفاده و کاربرد مجدد آن نقش مهمی دارد.

□ شکل ۳-۷ به‌طور خلاصه مراحل ذکر شده را نمایش می‌دهد.



شکل ۲-۷. مراحل ساخت هستی‌شناسی (کرچو و همکاران ۲۰۰۳)

روش نوی و مک‌گینس^۱

نوی و مک‌گینس (۲۰۰۱) روش‌شناسی خود را برای ساخت هستی‌شناسی در قالب راهنمایی که بارها مورد استناد و استفاده پژوهش‌های این حوزه قرار گرفته است، ارائه می‌کنند. روش یاد شده شامل هفت گام عمده است که هر یک از آن‌ها نیز دارای مراحل هستی‌شناسی در راهنمای یاد شده به تفصیل مورد بحث قرار گرفته‌اند. در ادامه به گام‌های اصلی این روش‌شناسی ارائه می‌شود:

¹ Noy and Mac Guinness

گام اول: شناسایی حوزه و دامنه هستی‌شناسی که شامل پاسخ‌گویی به سؤالات زیر است:

- دامنه پوشش هستی‌شناسی چیست؟
- از این هستی‌شناسی برای چه منظوری استفاده خواهد شد؟
- این هستی‌شناسی باید پاسخگوی چه نوعی از سؤالات باشد؟
- چه کسانی از این هستی‌شناسی استفاده و نگهداری خواهند نمود؟
- سؤالات شایستگی که شامل طرح و پاسخ‌گویی به فهرستی از سؤالات برای اطمینان از جامعیت و تمامیت هستی‌شناسی با توجه به دامنه مورد نظر است.

گام دوم: در نظر داشتن استفاده مجدد از هستی‌شناسی‌های موجود: این گام به‌طور کلی شامل بررسی منابع و هستی‌شناسی‌های موجود و ارزیابی قابلیت استفاده از آن‌ها است.

گام سوم: ذکر اصطلاحات مهم در هستی‌شناسی است بدین معنی که فهرستی از اصطلاحات مهمی را که لازم است برای کاربران توضیح دهیم به همراه تعاریف آن‌ها یادداشت کنیم.

گام چهارم: تعریف کلاس‌ها و سلسله‌مراتب کلاس‌ها است که از طریق اتخاذ رویکردهای مختلفی در ایجاد سلسله‌مراتب مفاهیم که پیش‌تر به آن اشاره شد، قابل انجام است.

گام پنجم: تعریف ویژگی رده‌ها است که در آن ساختار کلاس‌ها، کلاس‌های فرعی و ساختار درونی مفاهیم توصیف می‌شوند.

گام ششم: تعریف چهریزه‌ها است که شامل تعیین نوع مصادیق، تعداد آن‌ها و ارزش‌های^۱ مجاز مثلاً حرفی یا عددی بودن و ... است که می‌توان به یک ویژگی تخصیص داد.

روش‌شناسی یوپن^۲

به نقل از گومز پرز و همکاران (۲۰۰۴) در سال ۲۰۰۵ این روش‌شناسی را بر اساس فرایند یکپارچه توسعه نرم‌افزار^۳ ارائه دادند. هرچند این روش هنوز به‌صورت کامل در پروژه‌های مختلف آزموده نشده و ابزارهای پشتیبانی آن نیز در مراحل مقدماتی هستند، از نظر مفهومی به‌خوبی برنامه‌ریزی شده و بر اساس فرایند یکپارچه توسعه نرم‌افزار و زبان مدل‌سازی یکپارچه^۴ پشتیبانی می‌شود. این روش شامل مراحل زیر است که به‌صورت زنجیره‌وار انجام می‌پذیرد:

^۱ Value

^۲ UPON (Unified Process for Ontology Building)

^۳ Unified Software Development Process

^۴ Unified Modeling Language

مرحله آغازین^۱: که شامل شناسایی شرایط و لوازم و تحلیل منابع موجود و روند اجرایی است.

توسعه و گسترش^۲: که در آن طراحی نظام مفهومی صورت می‌پذیرد.

ساخت^۳: که شامل پیاده‌سازی و رسمی‌سازی مفاهیم به زبان هستی‌شناسی است.

گذار و انتقال^۴: که شامل آزمون پوشش و کاربرد دامنه موضوعی هستی‌شناسی است.

این مراحل به صورت دور تسلسل انجام می‌پذیرند و پس از هر دور یک هستی‌شناسی کاربردی تولید می‌شود.

۲-۱۴. مقایسه روش‌شناسی‌های مهندسی هستی‌شناسی

اغلب روش‌شناسی‌های مورد بحث در این بخش مبتنی بر تجارب مربوط به هستی‌شناسی‌های سازمانی هستند. این روش‌ها اغلب مراحل مشترکی را برای مهندسی هستی‌شناسی ارائه می‌کنند. بر همین اساس مراحل اصلی این روش‌شناسی‌ها عبارتند از: هدف‌گذاری، فراهم‌آوری دانش و کدگذاری مفاهیم و روابط. توافقی در مورد روش‌های طراحی و ساخت هستی‌شناسی‌ها وجود ندارد، به این دلیل که این تصمیمات باید مبتنی بر کاربرد و هدف هستی‌شناسی گرفته شود (نوی و مک‌گینس ۲۰۰۱). البته فقدان توافق بر روی روش‌ها به معنی کم بودن کیفیت آن‌ها نیست. هدف و نیاز باید نقطه شروع برای ساخت یک هستی‌شناسی جدید یا استفاده مجدد از یک هستی‌شناسی موجود باشد (پیتو و مارتینز ۲۰۰۱، نقل در ساوسا ۲۰۱۳).

فرناندز لویز و گومز پرز روش‌شناسی‌های ساخت هستی‌شناسی را بر اساس کاربرد آن‌ها به سه گروه تقسیم می‌کنند:

□ روش‌شناسی‌های مستقل از کاربرد: که در آن‌ها فرایند ساخت هستی‌شناسی از کاربران مستقل

است مانند روش سایک، مت‌آنتولوژی و روش آسکولد و کینگ.

□ روش‌شناسی‌های وابسته به کاربرد: که در آن‌ها سناریوهای ساخت هستی‌شناسی در یک

فرایند خاص تعیین می‌شوند.

□ روش‌شناسی‌های نیمه‌مستقل: که در آن‌ها هستی‌شناسی بر اساس کاربردهای که از آن‌ها

انتظار می‌رود، ساخته می‌شوند مانند روش توو و سن سس (فرناندز لویز و گومز پرز ۲۰۰۲).

¹ Inception

² Elaboration

³ Construction

⁴ Transition

- در مجموع برای مقایسه روش‌شناسی‌های مختلف هستی‌شناسی توجه به چهار نکته لازم است:
۱. هیچ‌کدام از رویکردها تمام مراحل ساخت هستی‌شناسی را پوشش نمی‌دهند. اغلب آن‌ها روی ایجاد هستی‌شناسی به‌خصوص مرحله مفهوم‌سازی و کاربرد هستی‌شناسی تأکید داشته و توجه کمی به تلفیق، ارزیابی و مراحل دیگر هستی‌شناسی در آن‌ها وجود دارد. بنابراین برای ساخت هستی‌شناسی از پایه باید این مراحل به روش‌شناسی مورد نظر اضافه شوند؛
 ۲. به دلیل جدید بودن نسبی حوزه مهندسی هستی‌شناسی اغلب مسائل مدیریت و ارزیابی در روش‌های یاد شده مد نظر قرار نگرفته است؛
 ۳. در کاربرد بسیاری از روش‌شناسی‌ها و رویکردهای ساخت هستی‌شناسی موانعی وجود دارد که باعث شده برخی از آن‌ها در خارج از حوزه موضوعی یا نمونه آزمایشگاهی تعمیم‌پذیر نباشند؛
 ۴. اغلب روش‌شناسی‌ها ابزار خاصی برای پشتیبانی از بخش فناوری و پیاده‌سازی هستی‌شناسی ندارند.

۲-۱۵. روش‌های طراحی هستی‌شناسی

با در نظر گرفتن پژوهش‌های مختلفی که به ساخت هستی‌شناسی در حوزه‌های گوناگون دانش پرداخته‌اند می‌توان دریافت که رویکردهای کلی در طراحی هستی‌شناسی نیز در کنار روش‌های مهندسی که پیش از این به آن‌ها پرداخته شد بخش مهمی از توجه محققان را به خود اختصاص داده‌اند. در واقع روش‌های طراحی خط‌مشی کلی سازندگان هستی‌شناسی را در استفاده از روش‌ها، ابزارها و رویکردهای متنوع در ساخت هستی‌شناسی‌ها مشخص می‌سازند. شمس‌فرد و عبدالله‌زاده بارفروش (۱۳۸۱) این روش‌ها را به سه دسته کلی تقسیم کرده‌اند:

- ساخت دستی: در این روش، حجم عظیمی از دانش مفهومی به‌وسیله افراد در ماشین‌کد می‌شود و پایگاه‌های دانش بزرگ عمومی یا تخصصی ایجاد می‌شوند. سایک و میکروکاسموس^۱ (نیرنبرگ و همکاران ۱۹۹۵) نمونه‌هایی از هستی‌شناسی‌های ساخته شده با این روش هستند.

استفاده از ابزارهای مهندسی هستی‌شناسی: در سال‌های اخیر ابزارهایی برای پشتیبانی ساخت هستی‌شناسی ساخته شده‌اند. پروتزه (اریکسون و همکاران ۱۹۹۹)، آنتولینگوآ (فارکوهار و همکاران ۱۹۹۷)، آنتوساروس (اسوارتوت و همکاران ۱۹۹۷) و وب آنتو (دومینگ ۱۹۹۸) با فراهم کردن واسط

^۱ Mikrokosmos

کاربری مناسب، محیط را برای فراهم‌آوری دانش مفهومی فراهم نموده‌اند. دسته دیگری مانند دادل^۱ (یا ماگوچی ۲۰۰۱) و اسوتلن^۲ (چالندار و گرو ۲۰۰۰) داده‌ها و ساختارهای لازم برای ساخت هستی‌شناسی را از منابع ورودی استخراج و در اختیار سازنده هستی‌شناسی (انسان یا ماشین) قرار می‌دهند.

□ ساخت نیمه‌خودکار: برای کاهش مشکلات ساخت هستی‌شناسی و تولید نیمه‌خودکار آن‌ها دو راهکار پیشنهاد شده است:

الف - یکپارچه‌سازی و استفاده مجدد از هستی‌شناسی‌های موجود (چاپولسکی و همکاران ۱۹۹۷؛ نوی و موسن ۲۰۰۰؛ ریوتارو و همکاران ۲۰۰۱)؛

ب - یادگیری و ساخت خودکار هستی‌شناسی‌ها از روی منابع موجود، مانند سیستم‌های هستی. تکست تو آنتو^۳ (مید و استاب^۴ ۲۰۰۱) و سیندیکیت^۵ (هان و روماکر^۶ ۲۰۰۱)، آسیوم^۷ (فوره و همکاران^۸ ۱۹۹۱) (به نقل از ساوسا ۲۰۱۳).

ساخت دستی هستی‌شناسی فرایندی پرهزینه، وقت‌گیر و مستعد خطا و سوگیری است و در مقابل تغییرات غیرمنعطف است. به همین دلیل خودکار سازی عمل ساخت هستی‌شناسی نه فقط هزینه‌ها را کاهش می‌دهد بلکه به تولید هستی‌شناسی‌ها با انطباق بیشتر با کاربرد آن منجر می‌شود. با این وجود استفاده از ابزارهای مهندسی هستی‌شناسی که به صورت واسط کاربرد عمل می‌کنند، نیاز به سازنده انسانی را منتفی نمی‌کنند، بلکه فقط محیط را برای او مهیا می‌سازند. حرکت به سمت خودکار سازی فراهم‌آوری دانش از روی منابع و متون، پایگاه‌های داده و هستان‌شناسی‌های دیگر، مشکلات مهندسی هستان‌شناسی را محدود و هزینه‌های ساخت و استفاده مجدد از آن‌ها را کاهش می‌دهد. با این حال در روش‌های خودکار نیز نظارت عامل انسانی همچنان امری ضروری است. همچنین در بخش‌های مربوط به معنا و بافت زبانی نیز استفاده از روش‌های خودکار با مشکلات متعددی روبه‌رو است که همچنان حل نشده باقی مانده است. بنابراین قضاوت در این مورد که کدام یک از روش‌ها برای ساخت هستی‌شناسی مناسب است به اهداف، کاربردها، زبان، نیروی انسانی، زمان و بسیاری عوامل دیگر بستگی دارد.

¹ DODDLE2

² SVETLAN

³ Text-to-onto

⁴ Maedche and Staab

⁵ Syndikate

⁶ Hahn & Romacker

⁷ Asium

⁸ Faure & et al.

۲-۱۶. روش‌های ارزیابی هستی‌شناسی

از سال ۱۹۹۰ بر اساس کاربردها و اهداف مختلف، طیفی از روش‌ها برای مهندسی و طراحی هستی‌شناسی‌ها توسعه یافته است که همزمان با آن برخی رویکردهای ارزیابی آن‌ها نیز ارائه شده است.

آسکولد و گرونینگر (۱۹۹۶) شفافیت، حداقل سوگیری در کدگذاری، گسترش‌پذیری، همبستگی مفاهیم با یکدیگر و وجود حداقل الزامات هستی‌شناسانه را به عنوان اصول طراحی هستی‌شناسی‌ها که می‌تواند برای ارزیابی آن‌ها نیز به کار رود را تشریح می‌کنند.

فرناندز لویز^۱ (۱۹۹۹) بعضی از روش‌های طراحی هستی‌شناسی‌ها را بر اساس ۹ معیار ارزیابی مورد تجزیه و تحلیل قرار داده است. وی روش‌شناسی آسکولد و کینگ (۱۹۹۵) که در توسعه اینترپرایز^۲ استفاده شد؛ روش‌شناسی گرونینگر و فاکس (۱۹۹۴) که حاصل توسعه پروژه هستی‌شناسی توو^۳ بود؛ روش‌شناسی ارائه شده توسط برناراز و همکاران (۱۹۹۶) که مجموعه‌ای در پروژه کاکتوس است؛ چارچوب مت‌آنتولوژی (فرناندز و همکاران ۱۹۹۷) که امکان ایجاد هستی‌شناسی‌ها را در سطح دانش فراهم می‌کند و همچنین روش‌شناسی مبتنی بر هستی‌شناسی سن‌سس برای استفاده در پردازش زبان طبیعی که در انسیتو علم اطلاعات^۳ گسترش یافته را مورد بحث قرار داده و برای هر یک نقاط قوت و ضعفی قائل می‌شود.

علاوه بر آنچه یاد شد، همان‌طور که در پژوهش ساوسا (۲۰۱۳) ذکر شده، گومز پرز و همکاران (۲۰۰۴) رویکردهای مختلفی را شامل تقسیم ارزیابی هستی‌شناسی به اعتبار‌پذیری و تأیید‌پذیری مطرح می‌کند. تأیید‌پذیری صحت ساخت و همچنین درستی تعاریف آن را ارزیابی می‌کند. اعتبار‌پذیری این‌که آیا تعاریف هستی‌شناسی با توجه به اهداف ایجاد آن بازنمون دنیای واقعی هستند یا خیر را بررسی می‌کند. تأکید وی بر اهداف هستی‌شناسی برای اطمینان از تعریف درست مفاهیم است. بر اساس این روش، ارزیابی تأیید‌پذیری هستی‌شناسی به معیارهای زیر بستگی دارد:

□ سازگاری^۴ به این معنا که کلاس‌ها ضد و نقیض نبوده و با یکدیگر هم‌خوانی داشته و فاقد خطای سازگاری باشند؛

□ جامعیت^۵ به این معنا که در چه سطحی دنیای واقعی را بازنمون می‌کنند. بر همین اساس چنانچه هستی‌شناسی کل دامنه هدف را پوشش ندهد خطای جامعیت اتفاق می‌افتد؛

¹ Fernández-López

² Enterprise Ontology

³ ISI (Information Sciences Institute)

⁴ Consistency

⁵ Completeness

□ فشردگی^۱ که با سازگاری تمام اطلاعات قابل دسترس در هستی‌شناسی مرتبط است که عدم این سازگاری، خطای حشو نامیده می‌شود.

رویکرد اعتبار‌پذیری برای ارزیابی کیفیت هستی‌شناسی اهمیت دارد. بر این اساس معیارهای کیفی مورد بحث قرار می‌گیرند. کارشناسان دامنه موضوعی مورد نظر می‌توانند هستی‌شناسی‌ها را بر اساس معیارهای کیفی نظیر آنچه در ادامه ذکر خواهد شد، ارزیابی کنند:

۱. سازگاری: به این معنا که هیچ تناقضی بین مفاهیم هستی‌شناسی وجود ندارد و این مفهوم با موارد زیر آشکار می‌شود:

□ دارای چرخه بودن؛

□ خطای از هم گسستگی بخش‌ها؛

□ رده‌بندی نادرست.

۲. جامعیت: نحوه پوشش موضوعات مورد نظر توسط هستی‌شناسی است که فقدان آن توسط موارد زیر قابل تشخیص است:

□ تعریف غیردقیق مفاهیم؛

□ از قلم افتادن برخی مفاهیم؛

□ تعریف جزئی و ناکامل برخی مفاهیم؛

□ موجودیت‌های منفصل و از هم گسیخته؛

□ حشو و فزونی کلاس‌ها، روابط یا نمونه‌ها.

۳. فشردگی: که به معنای عدم ارائه اطلاعات غیرضروری در هستی‌شناسی است.

۴. شفافیت: چگونگی ارائه مفاهیم با معانی مؤثر در هستی‌شناسی است؛

۵. کلی بودن: نحوه کاربرد هستی‌شناسی در طیفی از اهداف در دامنه مورد نظر را مورد بررسی قرار می‌دهد؛

۶. قدرت و پایداری در گذر زمان: که با توانایی هستی‌شناسی در پشتیبانی از تغییرات آینده سنجیده می‌شود؛

۷. غنای معنایی داده‌ها: که غنا و تنوع مفهوم‌سازی در هستی‌شناسی را بررسی می‌کند؛

¹ Conciseness

□ پوشش موضوعی غنی در دامنه مورد نظر: که مشخص می‌کند هستی‌شناسی دامنه موضوعی را تا چه سطحی.

۱۷-۲. محدودیت‌های هستی‌شناسی موجود در حوزه علوم زیستی

به‌طور فزاینده‌ای تشخیص داده می‌شود که پیچیدگی حوزه مراقبت‌های بهداشتی و علوم زیستی نیاز به اجماع در مورد اصطلاحات و زبان مورد استفاده در مستندات و ارتباطات دارد. این نیاز با رشد داده‌های تولید شده توسط پژوهش‌های حوزه مراقبت از بیمار و تحقیقات علوم زیستی مشخص‌تر می‌شود. در حال حاضر، این داده‌ها نمی‌توانند به‌طور کامل برای ادغام، بازیابی یا قابلیت تعامل مورد استفاده قرار گیرند، زیرا نظام‌های ترمینولوژی و طبقه‌بندی (که اغلب تحت عنوان «واژگان زیست‌پزشکی» قرار دارند) از جهات مختلف ناکارآمد هستند. ناهمگونی آن‌ها منعکس‌کننده زمینه‌ها، وظایف و نیازهای جوامع مختلف بوده و مانعی جدی برای جمع‌آوری داده‌ها و قابلیت همکاری متقابل مورد نیاز در تحقیقات زیست‌پزشکی و مراقبت‌های بهداشتی ایجاد می‌کند (ترابی ۱۴۰۰).

در جدول زیر تعدادی از هستی‌شناسی‌های مورد استفاده در حوزه‌های مختلف علوم زیستی و با اهداف مختلف ارائه شده است (محمدی ۱۴۰۰):

جدول ۲-۴. هستی‌شناسی‌های موجود در حوزه علوم زیستی

مشخصات	هستی‌شناسی / اصطلاحنامه
هستی‌شناسی سطح بالا برای تسهیل قابلیت همکاری بین هستی‌شناسی آناتومی موجود برای گونه‌های مختلف	CARO - Common Anatomy Reference Ontology
هستی‌شناسی بیماری که با استفاده از نقشه‌برداری گسترده از اصطلاحات DO به MeSH، ICD، اصطلاح‌نامه NCI، SNOMED و OMIM از لحاظ معنایی بیماری و واژگان پزشکی را ادغام می‌کند.	DO - Disease Ontology
هستی‌شناسی دامنه‌ای که نشان‌دهنده یک کل منسجم از دانش صریح اظهاری در مورد آناتومی انسان است. چارچوب هستی‌شناختی آن را می‌توان در مورد همه گونه‌های دیگر اعمال و گسترش داد.	FMA - Foundational Model of Anatomy
تلاش مشترک برای رفع نیاز به توصیف‌های مداوم از محصولات ژن در سراسر پایگاه داده‌ها. GO اطلاعات مربوط به فرایندهای بیولوژیکی، اجزای سلولی و عملکردهای مولکولی را نشان می‌دهد.	GO - Gene Ontology
هستی‌شناسی که نشان‌دهنده ناهنجاری‌های فنوتیپی است که در بیماری انسان مشاهده می‌شود.	HPO - Human Phenotype Ontology
مجموعه‌ای از هستی‌شناسی‌های قابل تعامل که حوزه بیماری‌های عفونی را پوشش	IDO - Infectious

مشخصات	هستی‌شناسی / اصطلاح‌نامه
می‌دهند. در هسته اصلی مجموعه هستی‌شناسی بیماری‌های عفونی (IDO-Core) موجودیت‌های مرتبط با جنبه‌های پزشکی و کلینیکی اکثر بیماری‌های عفونی قرار دارد. پسوند‌های خاص زیر دامنه‌ای از IDO-Core، موجودیت‌های مربوط به پاتوژن‌ها یا بیماری‌های خاص را پوشش می‌دهد.	Disease Ontology
مفاهیم، تعاریف، اصطلاحات و منابع برای توصیف استاندارد آزمایش‌های microarray	MGED - Microarray Gene Expression Data Ontology
تلاش برای ارائه شرایط استاندارد برای حاشیه‌نویسی داده‌های فنوتیپی پستانداران.	MP - Mammalian Phenotype Ontology
یک هستی‌شناسی یکپارچه برای توصیف علوم زیستی و تحقیقات بالینی.	OBI - Ontology for Biomedical Investigations
یک هستی‌شناسی مرجع برای حاشیه‌نویسی نتایج آزمایشات بالینی.	OCI - Ontology for Clinical Investigations
یک هستی‌شناسی مبتنی بر مقالات درمان و بیماری و تشخیص و سرطان و سایر اشخاص آسیب‌شناختی. قبلاً OGMS با نام clinical phenotype ontology شناخته می‌شد.	OGMS - Ontology for General Medical Science
هستی‌شناسی که می‌تواند در ارتباط با سایر هستی‌شناسی‌ها یا هستی‌شناسی‌های آناتومیکی برای مراجعه به فنوتیپ‌ها استفاده شود.	PATO - Phenotypic Quality Ontology
هستی‌شناسی برای نشان دادن دانش واکسیناسیون و پشتیبانی از استدلال خودکار.	VO - Vaccine Ontology
IICD که توسط سازمان بهداشت جهانی نگهداری می‌شود، سیستمی از کدهای تشخیصی را برای طبقه‌بندی بیماری‌ها، از جمله طبقه‌بندی‌های انواع مختلفی از علائم، نشانه‌ها، یافته‌های غیرطبیعی، شکایات، شرایط اجتماعی و دلایل خارجی آسیب یا بیماری فراهم می‌کند.	ICD - International Classification of Diseases
بانک اطلاعاتی و استاندارد جهانی برای شناسایی مشاهدات آزمایشگاهی پزشکی، از جمله نام کد آزمایشگاه پزشکی، تشخیص پرستاری، مداخلات پرستاری، طبقه‌بندی نتایج و مجموعه داده‌های مراقبت از بیمار.	LOINC - Logical Observational Identifiers, Names, and Codes
اصطلاحات سلسله‌مراتبی برای فهرست‌بندی و فهرست‌بندی اطلاعات پزشکی مانند MEDLINE/ PubMed و سایر پایگاه‌های داده ملی کتابخانه پزشکی.	MeSH - Medical Subject Headings
لیست داروهای الکترونیکی متمرکز که مورد استفاده بیمارستان‌ها و کلینیک‌های VHA قرار دارد. از NDF برای بررسی تداخلات دارویی، مدیریت سفارشات و ارسال نسخه‌های سرپایی به داروخانه‌های منطقه‌ای خودکار پست الکترونیکی استفاده می‌کنند.	NDF - National Drug File
نام‌گذاری استاندارد برای داروهای بالینی تولیدشده توسط کتابخانه ملی پزشکی.	RxNorm
یک مجموعه سیستماتیک قابل پردازش با رایانه شامل کدها، اصطلاحات، مترادف‌ها و	SNOMED-CT -

هستی‌شناسی / اصطلاح‌نامه	مشخصات
Systemized Nomenclature of Medicine-Clinical Terms	تعاریف مورد استفاده در اسناد و گزارش‌های کلینیکی SNOMED CT که از اصطلاحات پزشکی تهیه می‌شود. اصطلاحات اصلی عمومی را برای سوابق سلامت الکترونیکی، از جمله یافته‌های بالینی، علائم، تشخیص، روش‌ها، ساختار بدن، ارگان‌سیستم‌ها و سایر بیماری‌ها، مواد، داروهای دارویی، دستگاه‌ها و نمونه‌ها ارائه می‌دهد.
UMLS - Unified Medical Language System	مجموعه‌ای از پرونده‌ها و ابزارهایی که توسط کتابخانه ملی پزشکی نگهداری می‌شود و بسیاری از واژگان و استانداردهای بهداشتی و زیست‌پزشکی را با هم جمع می‌کند تا قابلیت همکاری بین سیستم‌های رایانه‌ای را فراهم کند.
NCBI Taxonomy	بانک اطلاعاتی تاکسونومی محور و یک طبقه‌بندی اصطلاحی برای همه ارگان‌سیستم‌های موجود در پایگاه داده‌های توالی عمومی است. این در حال حاضر حدود ۱۰٪ از گونه‌های توصیف شده از زندگی در این سیاره را نشان می‌دهد.

این تفاوت در هدف، عملکرد و طراحی هستی‌شناسی‌ها، دانشمندان این حوزه را متوجه ضرورت‌هایی دیگر در حوزه‌های تخصصی نموده است که ایجاد هستی‌شناسی‌ها، با توجه به نیاز کاربران و مخاطبان و ساختارهای مفهومی حوزه‌های موضوعی است.

۲-۱۷. ۱. ضرورت وجود هستی‌شناسی‌های تخصصی

درمان پزشکی مبتنی بر شواهد^۱ (EBM) در خدمات الکترونیکی مراقبت‌های بهداشتی مدرن به‌طور فزاینده‌ای اهمیت یافته است. در واقع، مفهوم عمل مبتنی بر شواهد حتی در حوزه مهندسی نرم‌افزار گسترش می‌یابد. مراقبتی که مبتنی بر شواهد نباشد، از دیدگاه‌های حرفه‌ای، ایمنی و اقتصادی به‌طور فزاینده‌ای غیرقابل پذیرش می‌شود. دسترسی الکترونیکی به اطلاعات با کیفیت بالا می‌تواند دانش حرفه‌ای پزشکان را بهبود ببخشد و بسیار محبوب شده است. با این وجود، در تهیه اطلاعات با کیفیت بالا برای پشتیبانی از پزشکی مبتنی بر شواهد، مشکلاتی وجود دارد. اما در این مسیر مشکلاتی جهت دستیابی به شواهد وجود دارد که تحقق آن را دشوار ساخته است. از جمله در سیستم‌های اطلاعات سلامت، شیوه‌های مدیریت سنتی واژگان غالباً در حوزه‌های اصطلاحات خاص بی‌اثر هستند. افزایش تنوع در این محیط‌های زبانی، نشان‌دهنده دشواری رو به رشد در اتخاذ و استفاده از منابع مبتنی بر شواهد است. در چنین محیط‌هایی توسعه چارچوب زبان پزشکی یکپارچه و پذیرفته‌شده جهانی تاکنون اثبات نشده است. تحولات کنونی هم در سیستم‌های ملی و هم در سطح بین‌المللی حاکی از آن است که استانداردها در واقع قابل دستیابی هستند اما نیاز به همکاری و سازش بیشتر توسط ارائه‌دهندگان مراقبت‌های بهداشتی دارند.

¹ Evidence Based Medical

از طرفی با توجه به تکامل پذیرش پزشکی مبتنی بر شواهد در سراسر جهان، الزامات استانداردهای اصطلاحات در ارتباطات در سراسر سیستم‌های بهداشتی به‌عنوان یک نیاز اساسی مطرح شده است. مطالعات درباره جستجوی اطلاعات مربوط به سلامت در وب نشان می‌دهد که ترمینولوژی‌های غیرتخصصی تنها به مقداری جزئی در یافتن اطلاعات مفید سلامت مؤثر هستند و اثربخشی جستجو با استفاده از این ترمینولوژی‌ها در بسیاری از موارد جستجو در وب، غالباً اطلاعات گمراه‌کننده یا اطلاعات نامربوط را برای مصرف‌کننده اطلاعات سلامت بازیابی می‌کند. دسترسی به اطلاعات بهداشتی با استفاده از موتورهای جستجو و یک عبارت ساده جستجو کارآمد نیست، زیرا سطح بالایی از امکانات برای بازیابی و درک اطلاعات بهداشتی مبتنی بر وب لازم است. در حال حاضر، به نظر می‌رسد که وب معنایی‌تر شده است و تکنیک‌های گسترش جستجوی دانش‌بنیان، پذیرفته‌تر می‌شوند. تفسیر پرس‌وجوهای کاربر برای موتورهای جستجو سخت‌تر است زیرا اکثر آن‌ها از اصطلاحات عمومی یا عام‌تر استفاده می‌کنند. در این راستا روش‌های معنایی در بهبود فرایندهای جستجو بسیار موفق بوده‌اند. برای مدیریت کارآمد جستجو و ارائه اطلاعات مربوطه، باید اطلاعات در سیستم‌های جستجو جانمایی و تعریف شود. این کار را می‌توان با تکنیک‌های بازیابی اطلاعات مبتنی بر هستی‌شناسی انجام داد. هستی‌شناسی برای تفکیک مفاهیم در زبان طبیعی مفید است.

مهم‌ترین مزیت استفاده از هستی‌شناسی معنایی رسمی^۱ است. موتورهای جستجوگر مبتنی بر وب معنایی برای تقویت عملکرد فرایند بازیابی اطلاعات از هستی‌شناسی در یک حوزه خاص استفاده می‌کنند. از دیدگاه پژوهشی، توانایی استنباط حقایق اضافی بر اساس محتوای هستی‌شناسی می‌تواند مهم باشد. استفاده از هستی‌شناسی در علوم زیستی و پزشکی عمدتاً بر بازیابی اصطلاحات پزشکی متمرکز است. متخصصان پزشکی از آن‌ها برای بیان دانش در مورد علائم و درمان بیماری‌ها استفاده می‌کنند. شرکت‌های داروسازی از آن‌ها برای ارائه اطلاعات در مورد داروها، دوزها و آلرژی‌ها استفاده می‌کنند. از طرف دیگر، طبیعت غیرمتمرکز وب ساخت یک هستی‌شناسی واحد را دشوار می‌کند. اگرچه استفاده از یک هستی‌شناسی واحد می‌تواند وظیفه ادغام و تداخل معنایی را آسان‌تر کند، اما از منظر مقیاس‌پذیری، حفظ قوام جهانی با یک هستی‌شناسی عظیم غیرعملی است. بنابراین، با توجه به ناکارآمدی نظام‌های ترمینولوژی و طبقه‌بندی (که اغلب تحت عنوان «واژگان زیست‌پزشکی» قرار دارند) و ناهمگونی آن‌ها به دلیل متفاوت بودن زمینه‌ها، وظایف و نیازهای تخصصی حوزه‌های متفاوت که مانعی جدی برای جمع‌آوری داده‌ها و قابلیت همکاری متقابل مورد نیاز در تحقیقات زیست‌پزشکی و مراقبت‌های بهداشتی ایجاد می‌کند، هستی‌شناسی‌های تخصصی می‌توانند به‌عنوان

¹ Formalized Semantic

یکی از بزرگ‌ترین راه‌حل‌ها در جهت رفع این مشکلات مطرح شوند. به مدد این هستی‌شناسی‌ها، می‌توان انتظار داشت نظام‌های بازیابی اطلاعات تخصصی، جستجوی هدفمندتر و نتایج هدفمندتری را بازیابی کنند.

۲-۱۸. مقایسه اصطلاح‌نامه و هستی‌شناسی

با وجود این که برخی از نویسندگان، هستی‌شناسی را شکل توسعه‌یافته اصطلاح‌نامه‌ها می‌دانند اما طبق تعاریف ارائه شده در بخش پیشین، تفاوت‌هایی بین این دو ابزار معنایی شناسایی شده است، برخی از تفاوت‌ها و شباهت‌های آن‌ها در زیر آمده است:

الف. شباهت‌ها

طبق آنچه در بخش‌های پیشین از ویژگی‌های هستی‌شناسی و اصطلاح‌نامه ارائه شد می‌توان گفت هر دوی آن‌ها فهرستی از واژگان و روابط میان آن‌ها را ارائه می‌کنند. صرف نظر از نوع روابط و محدودیت‌هایی که بر این روابط در اصطلاح‌نامه‌ها اعمال می‌شود، وجه مشترک اصلی این دو ابزار معنایی مفاهیم و روابط بین آن‌هاست. در حوزه علم اطلاع‌رسانی، هر دو این ابزار، نظامی برای سازماندهی دانش محسوب می‌شوند که این نظام بر پایه مفاهیم و برای نمایش دانش موضوعی شکل می‌گیرند و به نوعی هر دو در سازماندهی منابع اطلاعاتی و بازیابی اطلاعات کاربرد مؤثری دارند. با وجود این که برخی نظام هر دوی این ابزار را نظامی سلسله‌مراتبی می‌دانند، اما بخشی (نه تمام ساختار) هر دوی ابزار سلسله‌مراتبی هستند و برای گروه‌بندی اصطلاحات و تعیین رده‌های فرعی و اصلی به کار می‌روند. ویژگی مشترک دیگر اصطلاح‌نامه‌ها و هستی‌شناسی‌ها این است که هر دو تصویری مفهومی از حوزه‌های دانش ارائه می‌دهند. به نحوی که می‌توان از این تصویر به منظور ارتباط مؤثر کاربر و حوزه استفاده نمود.

علاوه بر موارد ذکر شده، می‌توان در روش تدوین اصطلاح‌نامه‌ها و هستی‌شناسی‌ها نیز شباهت‌هایی یافت، مراحل مشترک ساخت این دو ابزار شامل مشخص کردن حوزه؛ جمع‌آوری واژگان و مفاهیم از متون، متخصصان و دیگران؛ تعیین روابط بر اساس قواعد، اصول و استانداردها است.

ب. تفاوت‌ها

چنانچه در بخش پیش گفته شد، اصطلاح‌نامه‌ها و هستی‌شناسی‌ها هر دو شبکه‌ای از اصطلاحات و روابط میان آن‌ها را به نمایش می‌گذارند، اما این دو نظام را مواردی از یکدیگر متمایز می‌سازد. طبق تعریفی که از هستی‌شناسی ارائه شد، یکی از ویژگی‌های آن، رسمی بودن است، که به معنای طراحی

آن برای درک از سوی ماشین است. بنابراین می‌توان تفاوتی در کاربران این دو ابزار قائل شد. به شکلی که اصطلاح‌نامه برای کاربران انسانی معنادار و قابل درک است در حالی که هستی‌شناسی برای پردازش توسط انسان و نیز ماشین کاربرد دارد. طبق آنچه در بخش‌های قبل نیز مؤکداً به آن اشاره شد، می‌توان تفاوت دیگری نیز بین این دو قائل شد به این شکل که اگرچه هر دو آن‌ها به مفاهیم و روابط بین آن‌ها می‌پردازند، اما اصطلاح‌نامه‌ها فاقد مشخصه‌های کامل و صریحی از مفاهیم هستند. به عبارت دیگر، مفاهیم در اصطلاح‌نامه‌ها به شکل دقیق تعریف نشده‌اند در صورتی که در هستی‌شناسی دانش مفهومی تعریف دقیقی از مفاهیم با استفاده از قواعد تعریف شده بر اساس یک حوزه ارائه می‌شود. در واقع در اصطلاح‌نامه ابهام‌های معنایی وجود دارد که به دلیل روابط محدود ایجاد شده است. اما در هستی‌شناسی با روابط معنایی واضح بین مفاهیم این ابهام مرتفع شده است. در اصطلاح‌نامه گاه روابط مربوط به نوع یا جنس، روابط نشان‌دهنده جزئی از کل و روابطی که برای نشان دادن نمونه یا مثال به کار می‌روند، می‌توانند در هم تلفیق شوند در صورتی که در برخی اصطلاح‌نامه‌ها، این روابط به جای هم به کار رفته‌اند. هستی‌شناسی به لحاظ ارائه روابط معنایی دقیق، از این انتقاد مبرا است.

اصطلاح‌نامه‌ها عمدتاً بر اساس منابع موجود شکل می‌گیرند و بیشتر درباره مدارک موجود هستند در حالی که هستی‌شناسی سعی در بازنمون گستره وسیعی از واژگان یک حوزه دارد. در واقع با قواعد وضع شده در هستی‌شناسی امکان کنترل یکنواخت بر مفاهیم متفاوت وجود دارد. مسئله متمایزکننده دیگری که بین اصطلاح‌نامه و هستی‌شناسی وجود دارد، پشتیبانی ماشینی است که در هستی‌شناسی‌ها به دلیل قابل درک بودن برای ماشین امکان‌پذیر است در حالی که در اصطلاح‌نامه‌ها به این دلیل که در نظام پایگاه داده‌ای خاص استفاده می‌شوند، امکان اشتراک در سایر نظام‌ها وجود ندارد. با وجود تمایزهای اشاره شده بین اصطلاح‌نامه و هستی‌شناسی، در کاربرد این دو در بازیابی اطلاعات در فرایندهای مختلف مشابهت‌هایی وجود دارد. اما، نوع و چگونگی استفاده از هستی‌شناسی در این فرایندها متفاوت است.

۲-۱۹. هستی‌شناسی در بازیابی اطلاعات

هستی‌شناسی با هدف اشتراک و استفاده مجدد از اطلاعات ذخیره شده ایجاد و رسمی‌سازی شده و موجب می‌شود هم انسان و هم رایانه بتوانند اطلاعات را تفسیر کنند. لازمه این که کاربران قادر باشند به جستجو در مجموعه‌های عظیم اطلاعاتی و نظام‌های برخوردار از قابلیت‌های جستجوی پیچیده (نه صرفاً آزاد) بپردازند، این است که نظام‌هایی با قابلیت‌های زیر وجود داشته باشند:

۱. بهبود تعامل پیشرفته کاربران با نظام سازماندهی دانش و در نتیجه بهبود فرمول‌بندی جستجو؛

۲. ارائه نظام هوشمند برای گسترش جستجو؛

۳. پشتیبانی از هوش مصنوعی و وب معنایی.

هستی‌شناسی‌ها به منظور مشارکت در جستجو می‌توانند نقش‌های مختلف و ویژه‌ای را ایفا کنند. این نقش‌ها ممکن است به صورت جداگانه یا گروهی و هماهنگ با یکدیگر ایفا شوند.

۱. به عنوان اساسی برای ساختاردهی و سازماندهی معنایی اطلاعات مخزن؛ این ساختار ممکن است حداقل به دو روش مختلف استفاده شود:

□ به عنوان چارچوبی مفهومی به منظور کمک به جستجوگر در تنظیم سؤالات پیرامون مخزن اطلاعات مورد (برای مثال: نظام طبقه‌بندی یاهو)؛

□ به عنوان واژگانی برای فراداده‌های مورد استفاده به منظور ایجاد فهرست راهنما و برجسب‌گذاری مستندات در مخزن (نیازمند همکاری متقابل کاربر نمی‌باشد).

۲. کمک به تنظیم سؤالات: در اینجا با توجه به مشارکت کاربر، هستی‌شناسی می‌تواند به دو شیوه مختلف استفاده شود:

□ واداشتن رابط کاربر گرافیکی نسبت به ایجاد اصطلاح سؤالات، برای مثال از طریق از بین بردن ابهامات؛

□ به کارگیری استنباط به منظور ارتقای سؤال (نیازمند همکاری متقابل کاربر نیست) (آسوشه و همکاران ۱۳۹۰).

برای درک عملکرد هستی‌شناسی در فرایند بازیابی اطلاعات می‌توان به آن‌چه که در این فرایند رخ می‌دهد و نیز عوامل و عناصر دخیل در هر مرحله از فرایند توجه کرد. مرور متون منتشر شده در زمینه هستی‌شناسی‌ها نشان می‌دهد که طی سالیان اخیر، این حوزه مورد توجه متخصصان علوم مختلف از جمله کامپیوتر، زبان‌شناسی و علوم کتابداری و اطلاع‌رسانی قرار گرفته است. هر یک از متخصصان سعی کرده‌اند از جنبه‌های مختلفی به بررسی عملکرد هستی‌شناسی‌ها در فرایند بازیابی اطلاعات بپردازند.

در میان فرایندهای مربوط به بازیابی اطلاعات، جستجو فرایند اصلی آن و هسته اصلی این پژوهش را تشکیل می‌دهد. فرایند جستجو شامل استراتژی سیستم برای بازیابی اسنادی است که با پرس‌وجو مطابقت دارند. یک سیستم بازیابی اطلاعات اسناد مرتبط را انتخاب می‌کند و بر اساس استراتژی، امتیازی که بسیار به نمایه‌سازی آن‌ها بستگی دارد رتبه‌بندی می‌کند. گسترش پرس‌وجو فرایند میانی است که پرس‌وجوی کاربر را بر اساس اطلاعات سیستمی درونی برای بهبود کیفیت نتایج مجدداً

فرمول‌بندی می‌کند. هدف اصلی گسترش پرس‌وجو بهبود پرس‌وجو با اضافه کردن اصطلاحات، عبارت‌ها، مثل مترادف‌ها، املاهای جایگزین و مفاهیم مرتبط نزدیک است.

در واقع در فرایند جستجو، نیاز اطلاعاتی کاربر به عبارت پرس‌وجو تبدیل می‌شود که رایانه می‌تواند آن را از طریق هستی‌شناسی‌ها تفسیر کند. سپس برای بهبود بازیابی، پرس‌وجوها گسترش داده می‌شوند. این فرایند بر اساس مفاهیم مرتبط در هستی‌شناسی‌های به اشتراک گذاشته شده انجام می‌شود. فرمول‌بندی پرس‌وجو می‌تواند با تعامل با یک یا چند هستی‌شناسی پشتیبانی شود که بر اساس حوزه اطلاعات مورد نیاز - که عنصر اصلی مدل بیتس است، انتخاب می‌شود، می‌توان معنای پرس‌وجوی کاربر را با توسعه پرس‌وجو به تصویر کشید.

بیشتر اوقات، عبارت جستجوی کاربران را به عنوان نقطه شروع کشف نیاز اطلاعاتی وی در نظر می‌گیرند، اما در بسیاری موارد، کاربران بدون توجه به روابط خاص میان اصطلاحات، آن‌ها را به صورت جداگانه در یک عبارت جستجو قرار می‌دهند. با استفاده از هستی‌شناسی، نظام می‌تواند اطلاعات دقیقی را که مورد نیاز کاربر است درک و با رفع ابهام از پرسش کاربر، آن را به صورت دقیق بیان کند. در این موقعیت این تعریف دقیق روابط معنایی و وجود قواعد است که به کمک نظام آمده و به درک نیاز اطلاعاتی کاربر و معنای واقعی عبارت جستجوی وی به نظام کمک نماید.

کاربران به ندرت به شکل ایده‌آل، نیازهای خود را با زبان‌های پرس‌وجو فرمول‌بندی می‌کنند زیرا آن‌ها دانش کمی از اطلاعات استراتژی سیستم بازیابی اطلاعات از منابع معنایی زمینه‌ای و از محتوای پایگاه داده را دارند. براین اساس، استراتژی‌های اصلاح و توسعه پرس‌وجو برای فراهم‌آوری فرمول‌بندی (نیمه) خودکار از پرس‌وجو توسعه می‌یابند. این فرمول‌بندی مجدد ممکن است پرس‌وجو را با اضافه کردن مفاهیم به آن یا با حذف مفاهیم «ضعیف» از آن یا با اصلاح وزن تعیین شده در اصطلاحات پرس‌وجو اصلاح کند.

گسترش پرس‌وجو می‌تواند به سه روش انجام شود. اولین روش گسترش پرس‌وجوی دستی است؛ تجربه کاربر در مورد این که کدام اصطلاح در پرس‌وجوی جدید بیاید تصمیم می‌گیرد و اصطلاحی را برای گسترش پرس‌وجوی اولیه انتخاب می‌کند. در دومین روش گسترش پرس‌وجو، پرس‌وجو به طور خودکار گسترش می‌یابد، وزن‌دهی برای تمامی اصطلاحات محاسبه می‌شود و اصطلاحاتی که بیشترین وزن را دارند برای اضافه شدن به پرس‌وجوی اولیه انتخاب می‌شوند. در گسترش پرس‌وجو بر اساس ارزیابی کاربر، به عنوان سومین روش گسترش پرس‌وجو، اصطلاحات توسط سیستم ایجاد می‌شوند و کاربر اصطلاح مناسب را برای گسترش پرس‌وجوی اولیه انتخاب می‌کند. محدودیت گسترش پرس‌وجو این است که ممکن است پرس‌وجو را به سوی دوری از هدف

کاربر ببرد به این دلیل که اصطلاح اضافه شده برای گسترش پرس‌وجو ممکن است منجر به بیان مفاهیم متفاوت می‌شود. اما گسترش پرس‌وجو بر اساس روابط معنایی می‌تواند به عنوان یکی از راه‌حل‌های این مشکل گسترش پرس‌وجو در نظر گرفته شود و این مسئله را تا حد زیادی رفع کند.

استخراج اصطلاحات مرتبط به شکل کاملاً خودکار، معمولاً با استفاده از مجموعه اسناد است و بر اساس تعداد تکرار و هم‌رخدادی. بنابراین چندان قابل اعتماد نیست زیرا آمار و فراوانی در تشخیص متون مرتبط یا غیرمرتبط ضعیف هستند.

به طور کلی روش‌های مختلفی برای استفاده از اصطلاحات مرتبط برای گسترش پرس‌وجو به کار می‌رود، بعضی از آن‌ها در اینجا اشاره شده است:

□ ایجاد لیست اصطلاحات مرتبط با خوشه‌بندی: که مستقیم به گسترش پرس‌وجو مرتبط نیست

بلکه فقط لیستی به کاربر ارائه می‌دهد. کاربر با استفاده از آن پس از انجام جستجو در

صورتی که به مدرک مرتبط با نیاز خود دسترسی نیافت، می‌تواند با این لیست پرس‌وجوی

خود را اصلاح نماید و بازایی را بهبود بخشد؛

□ بازخورد: اصطلاحات اصلی‌ترین سند^۱ از نظر کاربر به پرس‌وجو اضافه می‌شود. اصطلاح به

پرس‌وجو به طور کلی اضافه می‌شود نه به پرس‌وجوی خاص. شیوه کار به گونه‌ای است که

اگر اصطلاحاتی که از اصلی‌ترین سند می‌آیند با اصطلاح پرس‌وجو هم‌رخداد باشند، به

پرس‌وجو افزوده می‌شود. در صورتی که این کار بر اساس روابط معنایی انجام گیرد پاسخ

بهتری خواهد داشت. به این شکل که اصطلاح اصلی‌ترین سند با هستی‌شناسی مطابقت داده

می‌شود و بر اساس نوع روابطی که با اصطلاح سؤال دارد، به اصطلاح سؤال افزوده می‌شود؛

□ نوع روابط معنایی: محققان انواع مختلفی از روابط معنایی مؤثر بر بازایی اطلاعات را به

آزمایش گذاشتند. بعضی از آن‌ها نوع خاصی از روابط معنایی را مثل مترادف و شبه‌مترادف‌ها

و اصطلاحات اخص را برای بهبود بازیافت شناسایی کردند و بعضی از روابط معنایی را در

بهبود دقت مثل اصطلاحات مرتبط و اعم را. همان‌طور که از نتایج برمی‌آید مشخص است

که آن‌چه تا کنون بررسی شده روابط موجود در اصطلاح‌نامه و نه روابط تقویت شده در

هستی‌شناسی هاست.

مطالعاتی که تاکنون در این رابطه انجام شده است نشان می‌دهد که گسترش پرس‌وجو با روابط

¹ top ranked

معنایی (مخصوصاً روابط اعم و اخص و IS A) بهبود می‌یابد و به طوری که می‌توان با کاربرد این نوع روابط معنایی در گسترش پرس‌وجو، بازیابی بهتری انجام داد.

مطالعات پیشین نشان می‌دهد که گسترش پرس‌وجو به‌طور کلی با دو دسته از الگوریتم‌ها انجام می‌شود. دسته اول الگوریتم‌هایی است که مبتنی بر مجموعه هستند و فنونی مانند بازخورد ربط، تحلیل محلی خودکار و تحلیل جهانی خودکار را برمی‌گیرد و دسته دوم که مورد توجه این پژوهش است دسته‌ای است که بر اساس رابطه انجام می‌شود. بنابراین در اینجا نحوه استفاده از هستی‌شناسی در گسترش پرس‌وجو بر اساس رابطه به‌طور خلاصه بحث می‌شود.

در حالی که روابط در گسترش پرس‌وجوی مبتنی بر مجموعه از مجموعه‌ای از اصطلاحات در اسناد به دست می‌آیند، گسترش پرس‌وجوی مبتنی بر رابطه از روابطی خارج از منابع استفاده می‌کنند. این منابع را به دو حوزه تقسیم کرده‌اند: اصطلاح‌نامه‌ها و هستی‌شناسی‌ها. اصطلاح‌نامه‌ها مجموعه‌ای از اصطلاحات هستند که به این عنوان شناخته شده‌اند که با یکدیگر در ارتباط هستند. در ساده‌ترین آن‌ها، روابط مترادف وجود دارند. بعضی از اصطلاح‌نامه‌ها شامل اطلاعاتی در مورد انواع دیگر ارتباطات است. در حالی که نمی‌توان تمایز واضحی بین روابط در اصطلاح‌نامه و هستی‌شناسی را شناسایی کرد اما هستی‌شناسی‌ها روابط را به شکل رسمی‌تری مشخص می‌کنند و در بیشتر زمینه‌های بازیابی دانش استفاده می‌شوند. به همین ترتیب، محدودیت‌های بازیابی اطلاعات مبتنی بر هستی‌شناسی سخت‌تر هستند و نتایج دقیق‌تر از آن چیزی است که در گسترش پرس‌وجوی مبتنی بر اصطلاح‌نامه وجود دارد.

بسیاری از محققان معتقدند استفاده از روابط معنایی در گسترش پرس‌وجو موجب افزایش دقت و بازیافت می‌شود. روابط معنایی را می‌توان از یک ساختار دانشی مثل اصطلاح‌نامه، تاکسونومی، شبکه واژگانی یا هستی‌شناسی گرفت. ساختارهای دانشی مثل هستی‌شناسی یا اصطلاح‌نامه بین انواع خاصی از روابط تمایز قائل می‌شوند مثلاً روابط سلسله‌مراتبی، مترادف و اصطلاحات مرتبط. چنین ساختارهای دانشی معمولاً دستی یا نیمه‌خودکار ایجاد می‌شوند. لیست‌های روابط معنایی معمولاً از طریق تحلیل مجموعه اسناد و آمارهای هم‌رخدادی ایجاد می‌شوند (دو اصطلاح در صورتی مرتبط هستند که با هم رخ دهند).

گسترش پرس‌وجوی مبتنی بر رابطه، مانند گسترش پرس‌وجوی مبتنی بر مجموعه مفید است، این حوزه بسیار مورد بررسی قرار گرفته است. بعضی از حوزه‌ها بسیار به کار نیاز دارد. اولین نقطه انواع روابط مورد استفاده در گسترش پرس‌وجو است. مجموعه‌ای از روابط مفید برای گسترش پرس‌وجو در حال رشد است و طبقه‌بندی ساده‌ای از روابط وجود ندارد، باعث گسترش کامل پرس‌وجو می‌شود،

زیرا ذهن انسان پیچیده‌تر از این است. با این وجود، احتمالاً دسته‌بندی‌های پیچیده‌تری از آن چه در حال حاضر استفاده می‌شود، وجود دارد و این پیچیدگی ناشناخته است. در ارتباط با نوع رابطه، مطالعات سیستماتیک کمی وجود دارد. سیستم‌های گسترش پرس و جوی موجود از مجموعه‌های محافظه کارانه روابط برای گسترش استفاده می‌کنند و در مورد سیستم‌های مبتنی بر هستی‌شناسی، نیز زیاد است و تنها روابط محدود (مثل بر اساس اشتراک) پذیرفته شده‌اند. در زمینه گسترش پرس و جوی، پرس و جوی مبتنی بر رابطه به اندازه پرس و جوی مبتنی بر مجموعه موفق نبوده است.

یکی از فرایندهایی که در گسترش پرس و جوی انجام می‌شود، پیشنهاد پرس و جوی است. پیشنهاد پرس و جوی تکنیکی است که هدف آن بهبود اثربخشی جستجوی وب کاربر با پیشنهاد جایگزین برای پرس و جوی اولیه است. پیشنهاد برای پرس و جوی ابزار مفیدی برای کاربران برای بیان نیاز اطلاعاتی خود با فراهم‌آوری پرس و جوی جایگزین است. مهم‌ترین هدف این پیشنهاد پرس و جوی کمک به کاربر برای رسیدن به هدف جستجوی آن‌ها است. بیشتر موتورهای جستجو پیشنهادهای پرس و جوی را با کامل کردن خود کار مثل کامل کردن خود کار اصطلاحات تایپ شده کاربر فراهم می‌کنند. معمولاً پرس و جوهای جایگزین که به کاربران پیشنهاد می‌شوند از لاگ پرس و جویی است که شامل پرس و جوهای جایگزینی است که به کاربران پیشنهاد شده، اسنادی که ممکن نیست بازیابی شوند و همچنین بعضی از اسناد بی‌ارتباط که ممکن است بازیابی شود هست. این به این دلیل است که کاربران همیشه پرس و جوهای مبتنی بر کلیدواژه را به موتورهای جستجو وارد می‌کنند. بنابراین برای غلبه بر مشکل ابهام، پرس و جوهای مبتنی بر مفهوم باید به جای پرس و جوهای کلیدواژه‌ای برای برآوردن هدف کاربر پیشنهاد شود.

فصل سوم

**روش تبدیل اصطلاح نامه علوم زیستی به
هستی شناسی**

۳-۱. مقدمه

پروژه تبدیل اصطلاح‌نامه به هستی‌شناسی، با تبدیل اصطلاح‌نامه شیمی پژوهشگاه ایرانداک در شهریورماه سال ۱۳۹۹ آغاز گردید. هستی‌شناسی ایجاد شده باید توسط نرم‌افزار پروتژه (Protege) قابل بازیابی، ویرایش، و ذخیره باشد. این پروژه با مطالعه بر روی ساختار داخلی فایل‌های هستی‌شناسی ساخته شده توسط نرم‌افزار پروتژه آغاز شد. نحوه ساخت سوپر کلاس‌ها، زیر کلاس‌ها، شخصیت‌ها (Individual) و سایر موارد مورد بررسی و در پیاده‌سازی نهایی مورد استفاده قرار گرفتند.

در ادامه تبدیل سایر اصطلاح‌نامه‌ها به هستی‌شناسی، دومین موردی که به صورت رسمی تبدیل به هستی‌شناسی شد اصطلاح‌نامه علوم زیستی بود که در شهریورماه سال ۱۴۰۱ کار بر روی آن به اتمام رسید. در این پروژه علاوه بر مواردی که برای تبدیل اصطلاح‌نامه شیمی به هستی‌شناسی شیمی استفاده شد، بیش از ۱۸۰ هستی‌شناسی خارجی مورد مطالعه قرار گرفت تا بتوان حداکثر تعداد روابط معنایی دودویی بین مفاهیم را استخراج نمود. این روش خوشبختانه به نتایج خوبی ختم شد که در ادامه در بخش‌های بعدی این گزارش روند ساخت و تکمیل هستی‌شناسی و همچنین ویژگی‌های هستی‌شناسی علوم زیستی مورد تحلیل و بررسی قرار خواهد گرفت.

۳-۲. نحوه تبدیل اصطلاح‌نامه به هستی‌شناسی

برای تبدیل اصطلاح‌نامه به هستی‌شناسی از روشی که توسط (سورگل و همکاران ۲۰۰۴) ذکر شد، استفاده شده است. مراحل تبدیل شامل موارد زیر است:

- مفاهیم تبدیل به کلاس‌ها می‌شوند؛
- مفاهیم اعم تبدیل به کلاس بالاتر یا Super class می‌شوند؛
- مفاهیم اخص تبدیل به زیر کلاس یا Sub class می‌شوند؛
- اصطلاح‌های نامرجح تبدیل به Individual یا Instance می‌شوند؛
- اصطلاح‌های مرتبط و ترجمه تبدیل به ویژگی یا Property شدند.

برای ساخت هستی‌شناسی ابتدا کدنویسی‌هایی که در هستی‌شناسی‌های قابل تطبیق با نرم‌افزار پروتزه بودند، مورد مطالعه قرار گرفتند. به همین منظور از ابتدای کدهای نوشته شده در هستی‌شناسی علوم زیستی، شروع به توضیح داده می‌شود.

۱-۲-۳. کدهای مقدماتی هستی‌شناسی

در شروع ساخت هستی‌شناسی یک‌سری کدهای مقدماتی وجود دارد که به‌صورت زیر در سورس فایل هستی‌شناسی نمایش داده می‌شوند:

```
<?xml version="1.0"?>
<Ontology xmlns="http://www.w3.org/2002/07/owl#"

xml:base="http://www.semanticweb.org/irandoc/ontologies/2020/8/initial_ont.owl"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:xml="http://www.w3.org/XML/1998/namespace"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"

ontologyIRI="http://www.semanticweb.org/irandoc/ontologies/2020/8/initial_ont.owl"
l">
  <Prefix name="owl" IRI="http://www.w3.org/2002/07/owl#" />
  <Prefix name="rdf" IRI="http://www.w3.org/1999/02/22-rdf-syntax-ns#" />
  <Prefix name="xml" IRI="http://www.w3.org/XML/1998/namespace" />
  <Prefix name="xsd" IRI="http://www.w3.org/2001/XMLSchema#" />
  </"#Prefix name="rdfs" IRI="http://www.w3.org/2000/01/rdf-schema">
```

در این کدهای اولی ویژگی‌های کلی فایل هستی‌شناسی مشخص می‌گردد. به‌عنوان مثال xml:base در مورد موضوع هستی‌شناسی است و این موضوع از طریق آدرس اینترنتی مقابل آن بازیابی می‌شود. سایر موارد نیز هر یک ویژگی خاصی را در هستی‌شناسی مشخص می‌کند. البته باید توجه داشت که به‌دلیل این که هنوز هستی‌شناسی‌های طراحی شده بر روی وب قرار نگرفته‌اند، از این ویژگی‌ها نمی‌توان استفاده چندانی برد، و البته نداشتن این موارد بر روی کار ما تأثیری نخواهند داشت.

۱-۲-۳. ایجاد و تعریف Annotation Property

Annotation propertyها برای ایجاد ویژگی‌هایی از اصطلاح‌نامه در هستی‌شناسی هستند مانند ترجمه، اصطلاح مرتبط، وضعیت و یادداشت دامنه. برای ایجاد آن‌ها در فایل سورس هستی‌شناسی به‌صورت زیر عمل شده است:

```
<Declaration>
  </"AnnotationProperty IRI="#Translation>
</Declaration>
<Declaration>
```

```

</"AnnotationProperty IRI="#Related_Term>
<Declaration/>
<Declaration>
  </"AnnotationProperty IRI="#Status>
<Declaration/>
<Declaration>
  </"AnnotationProperty IRI="#Scope_Note>
<Declaration/>

```

این نکته باید مورد توجه قرار گیرد که برای استفاده از هر یک از این ویژگی، آن‌ها ابتدا باید مثل موارد بالا در فایل تعریف شوند، در غیر این صورت نرم‌افزار پروتژه هنگام باز کردن هستی‌شناسی با خطا مواجه خواهد شد.

۳-۲-۳. ایجاد و تعریف Object property

Object property ها روابط معنایی بین کلاس‌ها را در هستی‌شناسی مشخص می‌نمایند. برای استفاده از این موارد در هستی‌شناسی، همانند Annotation property باید ابتدا در فایل سورس هستی‌شناسی باید تعریف شوند. کدهای زیر تعریف برخی از این ویژگی‌ها در فایل هستی‌شناسی را نشان می‌دهد:

```

<Declaration>
  </"ObjectProperty IRI="#has_role>
<Declaration/>
<Declaration>
  </"ObjectProperty IRI="#part_of>
<Declaration/>
<Declaration>
  </"ObjectProperty IRI="#has_output>
<Declaration/>
<Declaration>
  </"ObjectProperty IRI="#has_primary_output>
<Declaration/>
<Declaration>
  </"ObjectProperty IRI="#has_input>
<Declaration/>

```

۳-۲-۴. شروع ساخت کلاس‌ها

برای ساختن کلاس‌ها در هستی‌شناسی ابتدا باید اعلان آن‌ها ساخته شوند. برای ساخت اعلان‌ها اسم اصلی کلاس نیز باید اعلام شود. در این پروژه به جای اسم اصلی کلاس، کدی که در فایل اصطلاح‌نامه در هنگام ایجاد آن برای هر مفهوم در نظر گرفته شده بود در نظر گرفته شده است. کد زیر اعلان یک کلاس را نشان می‌دهد:

```

<Declaration>
  </"Class IRI="#p21569>

```

<Declaration/>

در این قطعه کد، کلاسی با نام p21569 در فایل هستی‌شناسی اعلان شد. در ادامه ساخت تمامی ویژگی‌های این کلاس با نام p21569 به آن مرتبط می‌شود.

حال برای عنوان کلاس یا برچسب آن، کد زیر نوشته می‌شود:

<AnnotationAssertion>

</"AnnotationProperty abbreviatedIRI="rdfs:label>

<IRI>#p21565</IRI>

<Literal>کلورامین</Literal>

<AnnotationAssertion/>

همان‌طور که مشاهده می‌شود، برای آن کلاس برچسب «کلورامین» نسبت داده می‌شود.

برای نسبت دادن ترجمه نام این کلاس نیز کد زیر نوشته می‌شود. دقت شود که این کد هم شامل ویژگی Translation است که در بخش Annotation property تعریف شد، و هم شامل کد اصلی کلاس یعنی p21565:

<AnnotationAssertion>

<"AnnotationProperty IRI="#Translation/>

<IRI>#p21565</IRI>

<Literal>chloramine</Literal>

<AnnotationAssertion/>

برای سایر Annotation property ها نیز روش مشابهی استفاده می‌شود.

برای نمایش زیر کلاس‌ها به صورت زیر عمل می‌شود:

<SubClassOf>

</"Class IRI="#p21565>

</"Class IRI="#p21553>

<SubClassOf/>

این قطعه کد نشان می‌دهد که کلاسی با نام p21565 دارای زیر کلاسی با نام p21553 می‌باشد.

در این روش ابتدا نام سوپر کلاس یا کلاس والد آورده می‌شود و سپس نام زیر کلاس.

۳-۳. یافتن روابط دودویی در هستی‌شناسی

در فاز دوم پروژه، سعی بر استخراج روابط دودویی تعریف شده در هستی‌شناسی علم زیستی هستیم که قبلاً و در خارج از ایران توسعه یافته‌اند. برای استخراج روابط دودویی ابتدا جستجو برای هستی‌شناسی‌های موجود در دنیا انجام شد. پس از تلاش فراوان به لیستی از آنتولوژی‌های دست‌یافتیم که تقریباً شامل تمامی آنتولوژی‌های موجود در حوزه علوم زیستی می‌باشد. این مجموعه دارای ۱۸۳ هستی‌شناسی با بیش از هفت میلیون مفهوم می‌باشد. برخی از هستی‌شناسی‌های مورد مطالعه و حاضر

در این لیست به شرح زیر می‌باشد:

- Antibiotic Resistance Ontology
- Apollo Structured Vocabulary (Apollo-SV)
- Compositional Dietary Nutrition Ontology
- Uber-anatomy ontology
- Sickle Cell Disease Ontology

لیست کامل آنتولوژی‌های مورد استفاده در صفحه اینترنتی زیر موجود می‌باشد:

- <https://www.ebi.ac.uk/ols/ontologies>

۳-۱. نحوه جستجو برای روابط دودویی

برای یافتن روابط دودویی موجود در آنتولوژی ساخته شده از اصطلاح‌نامه، ابتدا تمامی آنتولوژی‌های در یافتی از وبسایت ebi.ac.uk مورد مطالعه قرار گرفت. به همین منظور تمامی کلاس‌های آنتولوژی‌ها و تمامی روابط معنایی بین آن‌ها استخراج و در فایل‌های مجزایی ذخیره شدند. پس از ذخیره سازی کلاس‌ها و روابط معنایی آن‌ها، به آنتولوژی علوم زیستی که از اصطلاح‌نامه ایرانداک ساخته شد مراجعه می‌کنیم. سپس شروع به جستجوی کلاس‌های استخراجی از آنتولوژی‌های خارجی در آنتولوژی ساخته می‌نماییم.

در حقیقت در این مرحله باید به دنبال روابط معنایی گشت. توجه شود که یک رابطه معنایی دودویی دو کلاس را به هم مرتبط می‌کند. پس در این مرحله ابتدا دنبال کلاس‌هایی می‌گردیم که در یک طرفه این روابط معنایی باشند. پس از پیدا کردن این کلاس‌ها، به طرف دیگر روابط معنایی رجوع کرده و آن‌ها را نیز با آنتولوژی ساخته شده تطبیق می‌دهیم. اگر این کلاس‌ها در این آنتولوژی پیدا شوند، یک رابطه معنایی به آنتولوژی ما افزوده می‌شود، و در غیر این صورت به دلیل این که ما فقط کلاس یک طرف از رابطه معنایی را داریم، آن را نمی‌توانیم به آنتولوژی اضافه کنیم. مثال زیر می‌تواند این موضوع را به وضوح نشان دهد.

کلمه Glycine که در آنتولوژی MAXO وجود دارد، دارای روابط معنایی با سایر کلاس‌ها به شرح زیر است (روابط معنایی با رنگ قرمز مشخص شده‌اند):

- glycine **Sub_class** serine family amino acid
- glycine **has role** neurotransmitter
- glycine **has role** micronutrient
- glycine **has role** nutraceutical
- glycine **has role** hepatoprotective agent
- glycine **has role** ec 2.1.2.1 (glycine hydroxymethyltransferase) inhibitor
- glycine **has role** nmda receptor agonist
- glycine **has role** fundamental metabolite
- glycine **is conjugate** acid of glycinate

ایجاد هستی‌شناسی از اصطلاح‌نامه علوم زیستی ایرانداک □ ۶۹

- glycine **is conjugate** base of glycinium
- glycine **is tautomer** of glycine zwitterion

همان‌طور که مشاهده می‌شود کلمه glycine با ۱۱ کلاس دیگر رابطه معنایی دارد. اما ما از این ۱۱ رابطه فقط یکی از آن را می‌توانیم استفاده کنیم و ۱۰ مورد دیگر عملاً قابل استفاده نخواهند بود، زیرا طرف دیگر این روابط در داخل آنتولوژی ما وجود ندارند. تنها رابطه مورد استفاده به شرح زیر است:

- glycine **has role** neurotransmitter

در مورد اخیر هر دو کلاس در دو طرف رابطه دودویی در آنتولوژی ما وجود دارد، یعنی کلمات neurotransmitter و glycine.

۳-۲. روابط دودویی تعریف شده در آنتولوژی‌های مورد مطالعه

تمامی روابط دودویی تعریف شده در آنتولوژی‌های مورد مطالعه بررسی شدند و لیست استخراجی از این روابط به شرح زیر می‌باشد:

- has_role
- part_of
- has_output
- has_primary_output
- has_input
- characteristic_of
- has_part
- develops_from
- capable_of
- directly_develops_from
- results_in_growth_of
- contributes_to_morphology_of
- produces
- surrounded_by
- connected_to
- composed_primarily_of
- produced_by
- continuous_with
- channel_for
- extends_fibers_into
- supplies
- channels_from
- distalmost_part_of
- proximalmost_part_of
- connects

ایجاد هستی‌شناسی از اصطلاح‌نامه علوم زیستی ایرانداک □ ۷۰

- contained_in
- derives_from/develops_from
- disease_has_location
- has_functional_parent
- has_parent_hydrate
- adjacent_to

لازم به ذکر است که تمامی موارد ذکر شده دارای کلاس و زیرکلاس در آنتولوژی ساخته شده علوم زیستی ایرانداک می‌باشند.

۳-۴. نتایج و تحلیل نهایی

۳-۴-۱. نتایج

پس از بررسی روابط معنایی طبق روشی که پیشتر بیان شد، به لیستی از روابط معنایی دودویی دست یافتیم که کلاس‌های هر دو طرف آن‌ها در اصطلاح‌نامه علوم زیستی ایرانداک یا آنتولوژی ساخته شده‌مان وجود دارند. این روابط در پیوست به همراه نام اختصاری آنتولوژی خارجی که آن روابط در آن‌ها یافت شده‌اند، آورده شده است. این روابط شامل ۸۲۰ مفهوم از مجموع ۵۰۰۰ مفهوم اصطلاح‌نامه علوم زیستی می‌باشد. این بدان معنی است که ما توانستیم برای حدود یک ششم از کل مفاهیم اصطلاح‌نامه روابط معنایی دودویی صحیح و استاندارد پیدا کنیم. همچنین تمامی روابط دودویی یافت شده که فقط یک طرف آن‌ها در اصطلاح‌نامه یافت شده‌اند نیز ۳۴۴۱ رابطه معنایی می‌باشد، که به دلیل زیاد بودن حجم آن‌ها در این گزارش نیامده است.

۳-۴-۲. تحلیل نهایی

با توجه به یافتن حدود یک ششم عدد از کل تعداد مفاهیم موجود در اصطلاح‌نامه در طرفین روابط معنایی دودویی، می‌توان بیان داشت که برای تعداد مناسبی از مفاهیم روابط دودویی استاندارد تو سعه داده شد. با توجه به حجم یافته شده، تعداد زیادی نفر ساعت نیروی متخصص در حوزه علوم زیستی صرفه‌جویی گردید.

اما دلیل این که تعداد باقی‌مانده از مفاهیم درون آنتولوژی‌های خارجی یافته نشدند، می‌تواند این مورد باشد که ترجمه مفاهیم در اصطلاح‌نامه علوم زیستی ایرانداک نرمالیزه نشده‌اند. نرمالیزه شدن بدان معنی است که برای هر مفهوم در همه آنتولوژی‌ها و اصطلاح‌نامه‌ها یک اصطلاح یکتا به کار برده شود. به مثال زیر توجه بفرمایید:

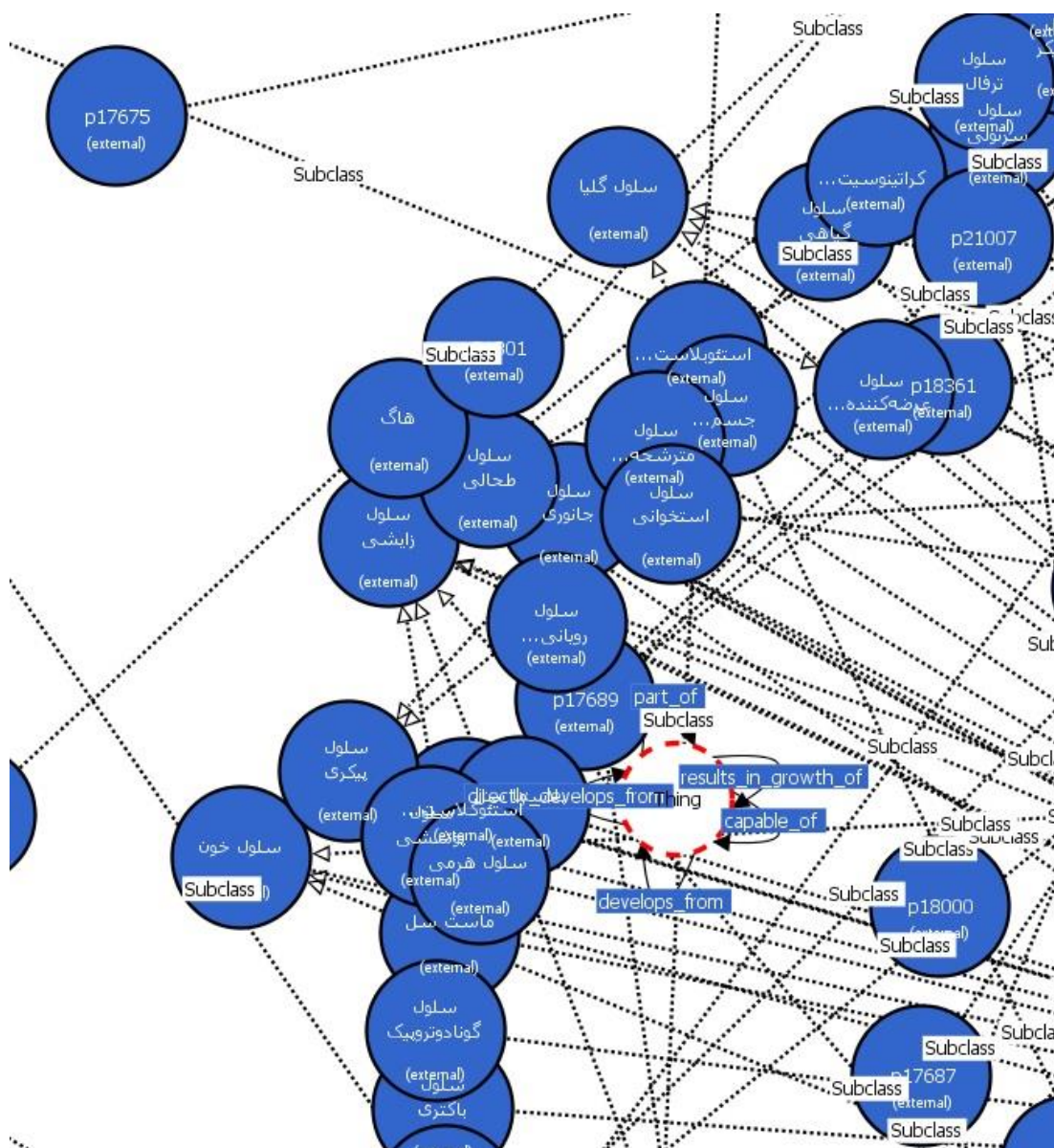
- Found term: **Cyanotic**
- Normalized term: **Cyanosis**

مثال دوم:

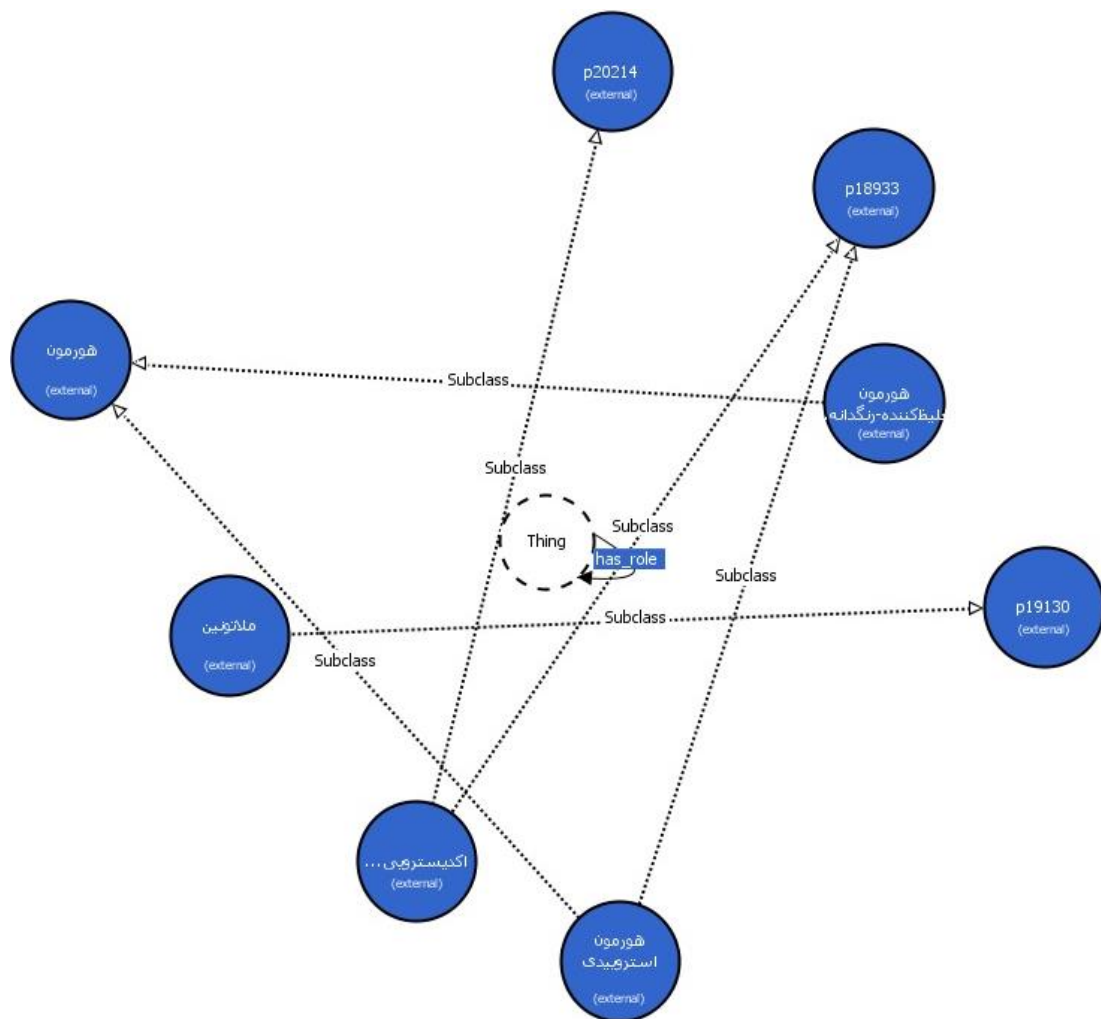
- Found term: **Left atrium fibrillated**
- Normalized term: **Left atrial fibrillation**

مثال‌های بالا به خوبی تفاوت مفاهیم یافته شده و مفاهیم نرمالیزه شده را نشان می‌دهند. دقت بفرمایید این مفاهیم از درون اصطلاح‌نامه یافته نشده‌اند و فقط برای درک مفهوم کلمه نرمالیزه بیان شدند. با توجه به این مثال‌ها باید مجدداً ترجمه مفاهیم درون اصطلاح‌نامه را بررسی کرد تا از نرمالیزه بودن آن‌ها اطمینان حاصل نمود. برای این منظور هم از روش‌های خودکار و بر مبنای هوش مصنوعی می‌توان بهره گرفت (پیوست ۱ و ۲).

۳-۵. پیوست‌ها



شکل ۳-۱. برشی از نقشه مفهومی سلول و زیرکلاس‌ها



شکل ۳-۲. برشی از نقشه مفهومی هورمون‌ها

فصل چهارم

تحليل روابط معنایی مفاهيم هستی‌شناسی

۴-۱. تحلیل روابط معنایی مفاهیم

1	strychnine	has role	cholinergic antagonist	ecto_classes_relation.txt
2	atropine	has role	muscarinic antagonist	dron_classes_relation.txt
3	benzodiazepine	has role	gaba modulator	ecto_classes_relation.txt
4	pupation	part of	pupal development	go_classes_relation.txt
5	ecdysteroid	Sub_class	steroid	cl_classes_relation.txt
6	ecdysteroid	has role	hormone	cl_classes_relation.txt
7	juvenile hormone	has role	hormone	cl_classes_relation.txt
8	protein biosynthesis	has output	protein	txpo_classes_relation.txt
9	gluconeogenesis	has primary output	glucose	cl_classes_relation.txt
10	hematophagy	has input	blood	ecocore_classes_relation.txt
11	protein biosynthesis	has output	protein	txpo_classes_relation.txt
12	gluconeogenesis	has primary output	glucose	cl_classes_relation.txt
13	rotenone	has role	toxin	phipo_classes_relation.txt
14	gibberellin	Sub_class	carboxylic acid	ecto_classes_relation.txt
15	gibberellin	has role	plant hormone	ecto_classes_relation.txt
16	chromosome morphology	characteristic of	chromosome	oba_classes_relation.txt
17	plastid	part of	cytoplasm	oba_classes_relation.txt
18	sarcoplasmic reticulum	Sub_class	endoplasmic reticulum	zp_classes_relation.txt
19	vacuole	part of	cytoplasm	maxo_classes_relation.txt
20	nasal mucosa	part of	nose	bto_classes_relation.txt
21	gastric mucosa	Sub_class	mucosa	bto_classes_relation.txt
22	gastric mucosa	part of	abomasum	bto_classes_relation.txt
23	mucosa	has part	epithelium	ohpi_classes_relation.txt
24	mucosa	has part	connective tissue	ohpi_classes_relation.txt
25	synaptic membrane	part of	synapse	maxo_classes_relation.txt
26	cytosol	part of	cytoplasm	wbphenotype_classes_relation.txt

27	axoneme	part of cytoskeleton	cl_classes_relation.txt
28	axoneme	has part microtubule	cl_classes_relation.txt
29	nucleosome	part of chromatin	oba_classes_relation.txt
30	chromatin	part of chromosome	maxo_classes_relation.txt
31	odontoclast	Sub_class osteoclast	cl_classes_relation.txt
32	osteoclast	develops from monocyte	cl_classes_relation.txt
33	osteoclast	capable of bone resorption	cl_classes_relation.txt
34	neutrophil	Sub_class granulocyte	cl_classes_relation.txt
35	phagocyte	capable of phagocytosis	maxo_classes_relation.txt
36	trophoblast	Sub_class embryonic structure	maxo_classes_relation.txt
37	trophoblast	part of blastocyst	maxo_classes_relation.txt
38	luteal cell	part of corpus luteum	cl_classes_relation.txt
39	adipocyte	part of brown adipose tissue	bto_classes_relation.txt
40	adipocyte	part of white adipose tissue	bto_classes_relation.txt
41	platelet	Sub_class blood cell	cl_classes_relation.txt
42	platelet	develops from megakaryocyte	cl_classes_relation.txt
43	platelet	capable of blood coagulation	cl_classes_relation.txt
44	eosinophil	Sub_class granulocyte	cl_classes_relation.txt
45	basophil	Sub_class granulocyte	cl_classes_relation.txt
46	neutrophil	Sub_class granulocyte	cl_classes_relation.txt
47	granulocyte	Sub_class blood cell	cl_classes_relation.txt
48	granulocyte	has part secretory granule	cl_classes_relation.txt
49	spermatozoid	develops from spermatid	clyh_classes_relation.txt
50	spermatocyte	develops from spermatogonium	maxo_classes_relation.txt
51	vitelline membrane	part of zygote	xenopus_anatomy_classes_relation.txt
52	ovum	Sub_class germ cell	bto_classes_relation.txt
53	gamete	Sub_class germ cell	maxo_classes_relation.txt
54	splenocyte	Sub_class leukocyte	maxo_classes_relation.txt
55	splenocyte	part of spleen	maxo_classes_relation.txt
56	myocyte	Sub_class cell	xenopus_anatomy_classes_relation.txt
57	myocyte	part of muscle	xenopus_anatomy_classes_relation.txt
58	hepatocyte	Sub_class epithelial cell	maxo_classes_relation.txt
59	hepatocyte	Sub_class somatic cell	maxo_classes_relation.txt
60	hepatocyte	part of liver	maxo_classes_relation.txt
61	kidney cell	Sub_class animal cell	maxo_classes_relation.txt
62	kidney cell	Sub_class somatic cell	maxo_classes_relation.txt

63	kidney cell	part of kidney		maxo_classes_relation.txt
64	oligodendrocyte	part of central nervous system		cl_classes_relation.txt
65	schwann cell	part of peripheral nervous system		maxo_classes_relation.txt
66	mast cell	has part	secretory granule	cl_classes_relation.txt
67	melanocyte	Sub_class	somatic cell	maxo_classes_relation.txt
68	cell	directly develops from	cell	ohpi_classes_relation.txt
69	protein biosynthesis	has output	protein	txpo_classes_relation.txt
70	gluconeogenesis	has primary output	glucose	cl_classes_relation.txt
71	phospholipase	Sub_class	lipase	txpo_classes_relation.txt
72	phospholipase	has role	phospholipase	txpo_classes_relation.txt
73	blood coagulation	Sub_class	hemostasis	cl_classes_relation.txt
74	bile secretion	has output	bile	txpo_classes_relation.txt
75	diacylglycerol	has role	second messenger	txpo_classes_relation.txt
76	cell growth	results in growth of	cell	cl_classes_relation.txt
77	pollen germination	part of	pollination	go_classes_relation.txt
78	chemotropism	Sub_class	tropism	zp_classes_relation.txt
79	nasal mucosa	part of	nose	bto_classes_relation.txt
80	nose	Sub_class	olfactory organ	ohpi_classes_relation.txt
81	nose	contributes to morphology of	respiratory system	ohpi_classes_relation.txt
82	ocellus	Sub_class	sense organ	lepao_classes_relation.txt
83	eye lens	part of	eye	spd_classes_relation.txt
84	eye	Sub_class	sense organ	ohpi_classes_relation.txt
85	ear	Sub_class	sense organ	ohpi_classes_relation.txt
86	sense organ	part of	sensory system	ohpi_classes_relation.txt
87	gill	part of	respiratory system	envo_classes_relation.txt
88	lung	contributes to morphology of	respiratory system	ohpi_classes_relation.txt
89	salivary gland	Sub_class	exocrine gland	ma_classes_relation.txt
90	salivary gland	part of	mouth	ma_classes_relation.txt
91	sebaceous gland	produces	sebum	maxo_classes_relation.txt
92	sweat gland	produces	sweat	maxo_classes_relation.txt
93	parathyroid gland	Sub_class	endocrine gland	maxo_classes_relation.txt
94	adrenal cortex	Sub_class	cortex	maxo_classes_relation.txt
95	adrenal cortex	part of	adrenal gland	maxo_classes_relation.txt
96	adrenal cortex	contributes to morphology of	adrenal gland	maxo_classes_relation.txt
97	adrenal medulla	part of	adrenal gland	maxo_classes_relation.txt

- 98 adrenal medulla surrounded by adrenal cortex
maxo_classes_relation.txt
- 99 adrenal medulla contributes to morphology of adrenal gland
maxo_classes_relation.txt
- 100 seminiferous tubule part of testis ma_classes_relation.txt
- 101 luteal cell part of corpus luteum cl_classes_relation.txt
- 102 corpus luteum part of endocrine system ma_classes_relation.txt
- 103 corpus luteum part of ovary ma_classes_relation.txt
- 104 follicular fluid part of ovarian follicle cl_classes_relation.txt
- 105 follicular fluid has part glycoprotein cl_classes_relation.txt
- 106 ovarian follicle part of ovary cl_classes_relation.txt
- 107 ovarian follicle has part oocyte cl_classes_relation.txt
- 108 ovarian follicle has part follicular fluid cl_classes_relation.txt
- 109 endocrine gland part of endocrine system
ohpi_classes_relation.txt
- 110 heart ventricle part of heart ma_classes_relation.txt
- 111 heart valve part of heart ma_classes_relation.txt
- 112 heart atrium part of heart ma_classes_relation.txt
- 113 heart contributes to morphology of cardiovascular system
ohpi_classes_relation.txt
- 114 nephron contributes to morphology of kidney
maxo_classes_relation.txt
- 115 pylorus contributes to morphology of stomach
maxo_classes_relation.txt
- 116 gastric mucosa Sub_class mucosa bto_classes_relation.txt
- 117 gastric mucosa part of abomasum bto_classes_relation.txt
- 118 gastric fundus part of stomach bto_classes_relation.txt
- 119 rumen epithelium part of rumen oba_classes_relation.txt
- 120 forestomach part of stomach ma_classes_relation.txt
- 121 abomasum part of ruminant stomach
uberon_classes_relation.txt
- 122 stomach contributes to morphology of digestive system
ohpi_classes_relation.txt
- 123 brain stem part of brain bto_classes_relation.txt
- 124 brain part of central nervous system ohpi_classes_relation.txt
- 125 brain contributes to morphology of nervous system
ohpi_classes_relation.txt
- 126 rumen epithelium part of rumen oba_classes_relation.txt
- 127 mucosa has part epithelium ohpi_classes_relation.txt
- 128 mucosa has part connective tissue ohpi_classes_relation.txt
- 129 sternum connected to rib maxo_classes_relation.txt
- 130 periosteum part of bone ma_classes_relation.txt
- 131 patella Sub_class sesamoid bone maxo_classes_relation.txt
- 132 patella part of knee maxo_classes_relation.txt
- 133 patella connected to femur maxo_classes_relation.txt
- 134 fat body Sub_class adipose tissue
xenopus_anatomy_classes_relation.txt
- 135 tendon part of musculoskeletal system ohpi_classes_relation.txt

136	ligament	has part	connective tissue	maxo_classes_relation.txt
137	endometrium	contributes to morphology of	uterus	ohpi_classes_relation.txt
138	myometrium	part of	uterus	ma_classes_relation.txt
139	sarcomere	part of	myofibril	maxo_classes_relation.txt
140	sarcoplasmic reticulum	Sub_class	endoplasmic reticulum	zp_classes_relation.txt
141	pleura	contributes to morphology of	respiratory system	maxo_classes_relation.txt
142	phagocyte	capable of	phagocytosis	maxo_classes_relation.txt
143	phagocyte	capable of	phagocytosis	maxo_classes_relation.txt
144	reticuloendothelial system	part of	immune system	uberon_classes_relation.txt
145	reticuloendothelial system	composed primarily of	phagocyte	uberon_classes_relation.txt
146	mast cell	has part	secretory granule	cl_classes_relation.txt
147	dermis	composed primarily of	connective tissue	ohpi_classes_relation.txt
148	sebaceous gland	produces	sebum	maxo_classes_relation.txt
149	sweat gland	produces	sweat	maxo_classes_relation.txt
150	gill	part of	respiratory system	envo_classes_relation.txt
151	nose	Sub_class	olfactory organ	ohpi_classes_relation.txt
152	nose	contributes to morphology of	respiratory system	ohpi_classes_relation.txt
153	pleura	contributes to morphology of	respiratory system	maxo_classes_relation.txt
154	larynx	contributes to morphology of	respiratory system	ohpi_classes_relation.txt
155	lung	contributes to morphology of	respiratory system	ohpi_classes_relation.txt
156	paranasal sinus	part of	respiratory system	ma_classes_relation.txt
157	air sac	part of	respiratory system	uberon_classes_relation.txt
158	trachea	contributes to morphology of	respiratory system	ohpi_classes_relation.txt
159	oviductal ampulla	part of	oviduct	bto_classes_relation.txt
160	oviduct	part of	female reproductive system	ohpi_classes_relation.txt
161	vulva	part of	female reproductive system	ma_classes_relation.txt
162	cervical mucus	Sub_class	mucus	ohpi_classes_relation.txt
163	cervical mucus	produced by	cervix	ohpi_classes_relation.txt
164	cervix	part of	uterus	ohpi_classes_relation.txt
165	cervix	continuous with	vagina	ohpi_classes_relation.txt
166	cervix	contributes to morphology of	uterus	ohpi_classes_relation.txt

167	female reproductive system	Sub_class	reproductive system	ohpi_classes_relation.txt
168	bulbourethral gland	Sub_class	exocrine gland	ma_classes_relation.txt
169	seminal vesicle part of testis			maxo_classes_relation.txt
170	prepuce has part	mucosa		maxo_classes_relation.txt
171	male reproductive system	Sub_class	reproductive system	maxo_classes_relation.txt
172	cloaca	part of digestive tract		ohpi_classes_relation.txt
173	reproductive system	has part	gonad	ohpi_classes_relation.txt
174	endocrine gland	part of endocrine system		ohpi_classes_relation.txt
175	endocrine system	composed primarily of	endocrine gland	ohpi_classes_relation.txt
176	ureter	channel for	urine	ohpi_classes_relation.txt
177	nephron	contributes to morphology of	kidney	maxo_classes_relation.txt
178	bladder	part of urinary system		xenopus_anatomy_classes_relation.txt
179	urethra	channel for	urine	ohpi_classes_relation.txt
180	cloaca	part of digestive tract		ohpi_classes_relation.txt
181	sense organ	part of sensory system		ohpi_classes_relation.txt
182	sensory system	has part	part of nervous system	ohpi_classes_relation.txt
183	parasympathetic nervous system		part of autonomic nervous system	maxo_classes_relation.txt
184	parasympathetic nervous system		contributes to morphology of autonomic nervous system	maxo_classes_relation.txt
185	sympathetic nervous system	part of	autonomic nervous system	ohpi_classes_relation.txt
186	sympathetic nervous system		contributes to morphology of autonomic nervous system	ohpi_classes_relation.txt
187	autonomic nervous system	part of	peripheral nervous system	ohpi_classes_relation.txt
188	autonomic nervous system		contributes to morphology of nervous system	ohpi_classes_relation.txt
189	peripheral nerve	part of	peripheral nervous system	ma_classes_relation.txt
190	peripheral nervous system	part of	nervous system	ohpi_classes_relation.txt
191	spinal cord	part of	central nervous system	ohpi_classes_relation.txt
192	brain	part of	central nervous system	ohpi_classes_relation.txt
193	brain		contributes to morphology of nervous system	ohpi_classes_relation.txt
194	central nervous system	part of	nervous system	ohpi_classes_relation.txt
195	nerve fiber	has part	axon	cl_classes_relation.txt
196	nerve fiber	extends_fibers_into	neuron	cl_classes_relation.txt

197	ganglion	part of nervous system	ohpi_classes_relation.txt
198	tendon	part of musculoskeletal system	ohpi_classes_relation.txt
199	flight muscle	Sub_class muscle	bto_classes_relation.txt
200	blood	has part blood plasma	ohpi_classes_relation.txt
201	coronary vessel	part of heart	ohpi_classes_relation.txt
202	coronary vessel	contributes to morphology of heart	ohpi_classes_relation.txt
203	mesenteric artery	supplies intestine	ohpi_classes_relation.txt
204	pulmonary artery	Sub_class artery	ohpi_classes_relation.txt
205	pulmonary artery	supplies lung	ohpi_classes_relation.txt
206	hepatic artery	Sub_class artery	maxo_classes_relation.txt
207	hepatic artery	supplies liver	maxo_classes_relation.txt
208	artery	channels_from heart	ohpi_classes_relation.txt
209	coronary vessel	part of heart	ohpi_classes_relation.txt
210	coronary vessel	contributes to morphology of heart	ohpi_classes_relation.txt
211	blood vessel	channel for blood	ohpi_classes_relation.txt
212	heart	contributes to morphology of cardiovascular system	ohpi_classes_relation.txt
213	cardiovascular system	part of circulatory system	ohpi_classes_relation.txt
214	cardiovascular system	has part heart	ohpi_classes_relation.txt
215	cardiovascular system	has part blood vessel	ohpi_classes_relation.txt
216	phagocyte	capable of phagocytosis	maxo_classes_relation.txt
217	reticuloendothelial system	part of immune system	uberon_classes_relation.txt
218	reticuloendothelial system	composed primarily of phagocyte	uberon_classes_relation.txt
219	lymph	develops from interstitial fluid	envo_classes_relation.txt
220	bile duct	channel for bile	maxo_classes_relation.txt
221	anus	part of digestive tract	ohpi_classes_relation.txt
222	anus	distalmost_part_of digestive tract	ohpi_classes_relation.txt
223	rectum	part of large intestine	ohpi_classes_relation.txt
224	rectum	distalmost_part_of large intestine	ohpi_classes_relation.txt
225	rectum	contributes to morphology of large intestine	ohpi_classes_relation.txt
226	cecum	part of large intestine	ma_classes_relation.txt
227	colon	part of large intestine	ohpi_classes_relation.txt
228	colon	contributes to morphology of large intestine	ohpi_classes_relation.txt
229	large intestine	part of intestine	ohpi_classes_relation.txt
230	large intestine	contributes to morphology of intestine	ohpi_classes_relation.txt
231	duodenum	part of small intestine	maxo_classes_relation.txt

- 232 duodenum proximalmost part of small intestine
maxo_classes_relation.txt
- 233 duodenum continuous with pylorus
maxo_classes_relation.txt
- 234 duodenum contributes to morphology of small intestine
maxo_classes_relation.txt
- 235 intestine part of intestine ohpi_classes_relation.txt
- 236 stomach contributes to morphology of digestive system
ohpi_classes_relation.txt
- 237 tooth enamel part of tooth bto_classes_relation.txt
- 238 tooth part of jaw ma_classes_relation.txt
- 239 tooth part of mouth ma_classes_relation.txt
- 240 palate part of mouth ma_classes_relation.txt
- 241 lip part of mouth ohpi_classes_relation.txt
- 242 lip contributes to morphology of mouth ohpi_classes_relation.txt
- 243 mouth has part part of respiratory system
ohpi_classes_relation.txt
- 244 mouth proximalmost part of digestive tract ohpi_classes_relation.txt
- 245 pancreatic juice produced by pancreas
maxo_classes_relation.txt
- 246 salivary gland Sub_class exocrine gland ma_classes_relation.txt
- 247 salivary gland part of mouth ma_classes_relation.txt
- 248 cloaca part of digestive tract ohpi_classes_relation.txt
- 249 pancreatic juice produced by pancreas
maxo_classes_relation.txt
- 250 pancreatic duct connected to intestine
maxo_classes_relation.txt
- 251 digestive tract part of digestive system ohpi_classes_relation.txt
- 252 digestive tract connects mouth ohpi_classes_relation.txt
- 253 trophoblast Sub_class embryonic structure
maxo_classes_relation.txt
- 254 trophoblast part of blastocyst maxo_classes_relation.txt
- 255 blastoderm Sub_class embryonic structure
zfa_classes_relation.txt
- 256 trophoblast Sub_class embryonic structure
maxo_classes_relation.txt
- 257 trophoblast part of blastocyst maxo_classes_relation.txt
- 258 blastocyst Sub_class embryonic structure
maxo_classes_relation.txt
- 259 blastocyst has part trophoblast
maxo_classes_relation.txt
- 260 blastocyst develops from morula maxo_classes_relation.txt
- 261 blastomere Sub_class cell clyh_classes_relation.txt
- 262 umbilical cord connects placenta
maxo_classes_relation.txt
- 263 chorioallantoic membrane develops from allantois
uberone_classes_relation.txt
- 264 allantois develops from cloaca maxo_classes_relation.txt

265	pancreatic juice	produced by	pancreas	maxo_classes_relation.txt
266	urine	Sub_class	excreta	ohpi_classes_relation.txt
267	urine	produced by	kidney	ohpi_classes_relation.txt
268	bodily secretion	produced_by	exocrine gland	ovae_merged_classes_relation.txt
269	blood serum	part of	blood plasma	maxo_classes_relation.txt
270	blood plasma	part of	blood	ohpi_classes_relation.txt
271	blood	has part	blood plasma	ohpi_classes_relation.txt
272	lymph	develops from	interstitial fluid	envo_classes_relation.txt
273	allantoic fluid	contained in	allantois	bto_classes_relation.txt
274	oviductal fluid	contained in	oviduct	bto_classes_relation.txt
275	follicular fluid	part of	ovarian follicle	cl_classes_relation.txt
276	follicular fluid	has part	glycoprotein	cl_classes_relation.txt
277	cervical mucus	Sub_class	mucus	ohpi_classes_relation.txt
278	cervical mucus	produced by	cervix	ohpi_classes_relation.txt
279	mucus	produced by	mucosa	ohpi_classes_relation.txt
280	nectary	part of	flower	bto_classes_relation.txt
281	sieve tube	part of	phloem	bto_classes_relation.txt
282	sapwood	Sub_class	secondary xylem	pso_classes_relation.txt
283	sclerenchyma	part of	shoot	bto_classes_relation.txt
284	periderm	part of	integument	ohpi_classes_relation.txt
285	endodermis	part of	leaf	bto_classes_relation.txt
286	endodermis	part of	root	bto_classes_relation.txt
287	pulvinus	part of	petiole	bto_classes_relation.txt
288	nectar	derives from/develops from	nectary	bto_classes_relation.txt
289	cutin	part of	plant cuticle	envo_classes_relation.txt
290	sieve tube	part of	phloem	bto_classes_relation.txt
291	sapwood	Sub_class	secondary xylem	pso_classes_relation.txt
292	vascular bundle	Sub_class	vascular tissue	bto_classes_relation.txt
293	vascular bundle	part of	leaf	bto_classes_relation.txt
294	vascular bundle	part of	stem	bto_classes_relation.txt
295	root	develops from	root primordium	envo_classes_relation.txt
296	stem nodule	part of	stem	bto_classes_relation.txt
297	internode	part of	stem	bto_classes_relation.txt
298	endosperm	part of	seed	bto_classes_relation.txt
299	plumule	part of	hypocotyl	bto_classes_relation.txt
300	epicotyl	part of	plumule	bto_classes_relation.txt
301	coleoptile	part of	plumule	bto_classes_relation.txt
302	sapwood	Sub_class	secondary xylem	pso_classes_relation.txt
303	corolla	part of	perianth	pso_classes_relation.txt
304	perianth	part of	flower	pso_classes_relation.txt
305	carpel	part of	gynoecium	pso_classes_relation.txt
306	gynoecium	part of	flower	maxo_classes_relation.txt
307	anther	part of	stamen	ppo_classes_relation.txt

- 308 stamen part of androecium ppo_classes_relation.txt
- 309 pollen derives from/develops from microspore
bto_classes_relation.txt
- 310 androecium part of flower ppo_classes_relation.txt
- 311 nectary part of flower bto_classes_relation.txt
- 312 inflorescence has part flower ppo_classes_relation.txt
- 313 endocarp part of pericarp mondo_classes_relation.txt
- 314 mesocarp part of pericarp mondo_classes_relation.txt
- 315 pericarp part of fruit mondo_classes_relation.txt
- 316 fruit develops from gynoecium maxo_classes_relation.txt
- 317 photoreceptor Sub_class sense organ dpo_classes_relation.txt
- 318 myocardial infarction disease has location heart
mondo_classes_relation.txt
- 319 pyruvic acid has role cofactor
maxo_classes_relation.txt
- 320 pyruvic acid has functional parent propionic acid
maxo_classes_relation.txt
- 321 glyoxylic acid Sub_class aldehydic acid mco_classes_relation.txt
- 322 malonic acid Sub_class lipid zp_classes_relation.txt
- 323 malic acid has functional parent succinic acid
mco_classes_relation.txt
- 324 tyrosine Sub_class aromatic amino acid
maxo_classes_relation.txt
- 325 tyrosine has functional parent propionic acid
maxo_classes_relation.txt
- 326 phenylalanine Sub_class aromatic amino acid
maxo_classes_relation.txt
- 327 aromatic amino acid Sub_class amino acid
maxo_classes_relation.txt
- 328 phenylalanine Sub_class aromatic amino acid
maxo_classes_relation.txt
- 329 methionine Sub_class sulfur-containing amino acid
maxo_classes_relation.txt
- 330 methionine has functional parent butyric acid
maxo_classes_relation.txt
- 331 tyrosine Sub_class aromatic amino acid
maxo_classes_relation.txt
- 332 tyrosine has functional parent propionic acid
maxo_classes_relation.txt
- 333 cysteine Sub_class sulfur-containing amino acid
maxo_classes_relation.txt
- 334 glycine has role neurotransmitter
maxo_classes_relation.txt
- 335 tyrosine Sub_class aromatic amino acid
maxo_classes_relation.txt
- 336 tyrosine has functional parent propionic acid
maxo_classes_relation.txt

337	phenylalanine	Sub_class	aromatic amino acid	
			maxo_classes_relation.txt	
338	tyrosine	Sub_class	aromatic amino acid	
			maxo_classes_relation.txt	
339	tyrosine	has functional parent	propionic acid	
			maxo_classes_relation.txt	
340	cysteine	Sub_class	sulfur-containing amino acid	
			maxo_classes_relation.txt	
341	phenylalanine	Sub_class	aromatic amino acid	
			maxo_classes_relation.txt	
342	glycine	has role	neurotransmitter	
			maxo_classes_relation.txt	
343	methionine	Sub_class	sulfur-containing amino acid	
			maxo_classes_relation.txt	
344	methionine	has functional parent	butyric acid	
			maxo_classes_relation.txt	
345	cysteine	Sub_class	sulfur-containing amino acid	
			maxo_classes_relation.txt	
346	melatonin	has role	hormone	cl_classes_relation.txt
347	methionine	Sub_class	sulfur-containing amino acid	
			maxo_classes_relation.txt	
348	methionine	has functional parent	butyric acid	
			maxo_classes_relation.txt	
349	beta-alanine	has role	neurotransmitter	
			mco_classes_relation.txt	
350	dopamine	Sub_class	catecholamine	cl_classes_relation.txt
351	nitric oxide	has role	neurotransmitter	
			cl_classes_relation.txt	
352	histamine	has role	neurotransmitter	
			cl_classes_relation.txt	
353	glycine	has role	neurotransmitter	
			maxo_classes_relation.txt	
354	peroxidase	Sub_class	enzyme	txpo_classes_relation.txt
355	peroxidase	has role	antioxidant	txpo_classes_relation.txt
356	lipoprotein	Sub_class	conjugated protein	
			cl_classes_relation.txt	
357	streptomycin	has role	protein synthesis inhibitor	
			phipo_classes_relation.txt	
358	xylose	Sub_class	aldopentose	mco_classes_relation.txt
359	galactose	Sub_class	aldohexose	
			maxo_classes_relation.txt	
360	glucose	Sub_class	aldohexose	maxo_classes_relation.txt
361	mannose	Sub_class	aldohexose	mco_classes_relation.txt
362	fructose	Sub_class	ketohexose	maxo_classes_relation.txt
363	propionic acid	Sub_class	saturated fatty acid	
			maxo_classes_relation.txt	
364	oleic acid	has role	antioxidant	cido_classes_relation.txt

365	unsaturated fatty acid	Sub_class	fatty acid	
				maxo_classes_relation.txt
366	oleic acid	has role	antioxidant	cido_classes_relation.txt
367	prostaglandin	Sub_class	prostanoid	zp_classes_relation.txt
368	thromboxane	Sub_class	prostanoid	oba_classes_relation.txt
369	phosphatidic acid	Sub_class	glycerophospholipid	
				zp_classes_relation.txt
370	phospholipid	Sub_class	lipid	
				wbphenotype_classes_relation.txt
371	lipoprotein	Sub_class	conjugated protein	
				cl_classes_relation.txt
372	pyridoxine	has role	cofactor	
				maxo_classes_relation.txt
373	riboflavin	has role	antioxidant	
				maxo_classes_relation.txt
374	riboflavin	has role	cofactor	
				maxo_classes_relation.txt
375	benzyl alcohol	has role	antioxidant	cido_classes_relation.txt
376	acetaldehyde	Sub_class	aldehyde	fypo_classes_relation.txt
377	cerulenin	has role	antimetabolite	fypo_classes_relation.txt
378	cerulenin	has role	antifungal agent	
				fypo_classes_relation.txt
379	cerulenin	has role	fatty acid synthesis inhibitor	
				fypo_classes_relation.txt
380	formamide	has functional parent	formic acid	
				fypo_classes_relation.txt
381	hydroxylamine	has parent hydride	ammonia	
				fypo_classes_relation.txt
382	methadone	has role	serotonin uptake inhibitor	
				ecto_classes_relation.txt
383	aldehydic acid	Sub_class	aldehyde	mco_classes_relation.txt
384	goat	Sub_class	mammal	mondo_classes_relation.txt
385	goat	Sub_class	mammal	mondo_classes_relation.txt
386	myelin	part of	nervous system	hp_classes_relation.txt
387	steroid hormone	Sub_class	steroid	
				maxo_classes_relation.txt
388	steroid hormone	has role	hormone	
				maxo_classes_relation.txt
389	peptide hormone	Sub_class	peptide	cl_classes_relation.txt
390	peptide hormone	has role	hormone	
				cl_classes_relation.txt
391	calcitonin	Sub_class	polypeptide	cl_classes_relation.txt
392	calcitonin	Sub_class	peptide hormone	
				cl_classes_relation.txt
393	thymus hormone	has role	hormone	
				txpo_classes_relation.txt
394	estradiol	has role	estrogen	cl_classes_relation.txt
395	ecdysteroid	Sub_class	steroid	cl_classes_relation.txt

396	ecdysteroid	has role	hormone	cl_classes_relation.txt
397	juvenile hormone	has role	hormone	cl_classes_relation.txt
398	cortisone	Sub_class	glucocorticoid	ecto_classes_relation.txt
399	corticosterone	Sub_class	glucocorticoid	mondo_classes_relation.txt
400	corticosterone	Sub_class	glucocorticoid	mondo_classes_relation.txt
401	pancreatic hormone	has role	hormone	txpo_classes_relation.txt
402	corticotropin-releasing hormone	Sub_class	polypeptide	oba_classes_relation.txt
403	corticotropin-releasing hormone	Sub_class	peptide hormone	oba_classes_relation.txt
404	corticotropin-releasing hormone	has role	neurotransmitter	oba_classes_relation.txt
405	cosyntropin	Sub_class	polypeptide	dron_classes_relation.txt
406	corticotropin	Sub_class	polypeptide	cl_classes_relation.txt
407	corticotropin	Sub_class	peptide hormone	cl_classes_relation.txt
408	pituitary hormone	has role	hormone	txpo_classes_relation.txt
409	blood brain barrier	part of	nervous system	cl_classes_relation.txt
410	blood brain barrier	adjacent to	brain	cl_classes_relation.txt

فصل پنجم

جمع‌بندی و نتیجه‌گیری

۵-۱. جمع‌بندی

بررسی تاریخی و کارکردی اصطلاح‌نامه‌ها در طول سال‌ها نشان می‌دهد که با وجود تغییر در محیط‌های اطلاعاتی، اهداف و ساختار آن‌ها همواره ثابت بوده است. به‌روزرسانی اصطلاح‌نامه‌های چاپی با تأخیر صورت می‌گیرد و اصطلاح‌نامه‌های الکترونیکی کمی در محیط‌های بازیابی اطلاعات ایران به کار می‌روند. تغییر محیط اطلاعاتی و ماهیت منابع اطلاعاتی مشهود است که می‌تواند بر ضرورت بازنگری در ساختار اصطلاح‌نامه‌ها تأکید داشته باشد. کاربران در محیط‌های اطلاعاتی جدید، به بازیابی اطلاعات با کمک ابزارهای پیشرفته جستجوی اطلاعات فراتر از جستجوی متن آزاد می‌پردازند که از ویژگی‌های این ابزارهای پردازش است. بنابراین محیط اطلاعاتی جدید، اصطلاح‌نامه‌هایی با قابلیت‌های بیشتر از آن‌چه اصطلاح‌نامه‌های کنونی با ساختارهای محدود فراهم می‌آورند، می‌طلبد. چالش‌های اصطلاح‌نامه را در بازیابی مدارک مرتبط شامل دشواری استخراج معانی جدید و قراردادن آن در ساختار معنایی موجود، ناتوانی در ارائه راهکارهای مناسب برای بازیابی اطلاعات غیرمتنی، توجه نکردن کافی اصطلاح‌نامه به سطح کاربر و اطلاعات، عدم روزآمدسازی اصطلاح‌های جدید علمی، دشواری ایجاد یکدستی در واژگان اصطلاح‌نامه‌های نظام‌های اطلاعاتی مختلف و مشکلات حل نشده استفاده از اصطلاح‌نامه‌های خودکار و نیز می‌توان به این چالش‌ها، محدود بودن ساختار و روابط معنایی موجود در آن به روابط سلسله‌مراتبی، هم‌بسته و ترجیحی را افزود. این چالش‌ها، نیاز به برداشتن گامی فراتر از اصطلاح‌نامه را مشهود می‌سازد. با توجه به عدم تغییر در ساختار اصطلاح‌نامه‌ها و همچنین نیاز به ابزار معنایی در محیط جدید اطلاعاتی، هستی‌شناسی‌ها پا به عرصه وجود گذاشتند و این قابلیت را فراهم آوردند تا روابط مفهومی دقیق‌تر و صریح‌تر از روابط موجود در اصطلاح‌نامه‌ها (سلسله‌مراتبی، هم‌بسته و ترجیحی) را بیان کنند.

اصطلاح آنتولوژی ریشه‌ای یونانی دارد. آنتولوژی دو واژه *onto* یعنی هستی و *logia* به معنای مطالعه را دربردارد و در کل به معنای هستی‌شناسی است. هستی‌شناسی شاخه‌ای از متافیزیک است که به مطالعه و شناخت ذات موجودات می‌پردازد. دغدغه هستی‌شناسی، شناسایی انواع پدیده‌هایی است

که در عالم واقع وجود دارند و سپس چگونگی توصیف آن‌هاست. در واقع هستی‌شناسی شاخه‌ای از علم فلسفه است که به مطالعه ماهیت چیستی و ساختار واقعیت می‌پردازد. هستی‌شناسی در فلسفه به دنبال پاسخ‌گویی به این سؤالات است که اشیاء چیستند؟ ماهیت درون چیزها که با تغییر ویژگی‌ها (تغییر رنگ، تغییر اندازه و ...) ثابت می‌ماند، چیست؟ آیا مفاهیم (درخت، کتاب، میز و غیره) خارج از ذهن قرار دارند؟ موجودیت‌های جهان را چگونه می‌توان طبقه‌بندی کرد؟

بحث هستی‌شناسی به عنوان زمینه‌ای در فلسفه، سال‌هاست که وجود داشته است. متخصصین علوم رایانه، هستی‌شناسی را نیز مانند بسیاری از اصطلاحات این رشته از دیگر علوم قرض گرفته و سپس معنایی فنی به آن داده‌اند.

پس از ورود هستی‌شناسی به علوم رایانه، تعاریف متفاوتی از آن مطرح شد که در اینجا به بعضی از آن‌ها اشاره می‌شود.

«هستی‌شناسی، به توصیف اصطلاحات پایه‌ای، ارتباطات و واژگان یک حوزه موضوعی و قوانین اتصال اصطلاحات و ارتباطی که تعمیم لغات را شرح می‌دهد، می‌پردازد». این تعریف توصیفی، بیانگر روش ایجاد هستی‌شناسی بوده و رهنمودهای مبهمی را ارائه می‌کند. این تعریف، اصطلاحات پایه‌ای و ارتباطات بین آن‌ها را فراهم می‌کند. اما هستی‌شناسی تنها شامل اصطلاحاتی نیست که صریحاً در آن تعریف شده، بلکه شامل دانش استنباط شده از آن نیز می‌باشد.

هستی‌شناسی‌ها در واقع بر اساس مفاهیم و ارتباطی که بین آن‌ها وجود دارد، سعی در استنباط و خلق دانش جدید از حوزه‌های موضوعی متفاوت دارند. بنابراین، هستی‌شناسی‌ها را نمی‌توان تنها لغات و روابط بین آن‌ها دانست.

در تعریف هستی‌شناسی در علوم رایانه شاید بهتر باشد به تعریف گروبر^۱ اشاره کرد که هستی‌شناسی را «مشخصات واضح و رسمی یک مفهوم» می‌نامد. رسمی بودن این مفاهیم امکان ماشینی شدن ارائه مفاهیم را می‌دهند. هستی‌شناسی در وب معنایی اصطلاحات و ارتباط بین آن‌ها را در دامنه مورد استفاده بیان می‌کند. به عبارت دیگر، هستی‌شناسی شامل مفاهیم، ویژگی‌ها و ارتباط بین آن‌هاست.

بنابراین، ویژگی اصلی و متمایزکننده هستی‌شناسی‌ها از سایر واژگان را می‌توان توجه به قوانین و محدودیت‌هایی دانست که برای روابط بین مفاهیم تعریف می‌شود.

بورست^۲ (۱۹۹۷) تعریف گروبر را به صورت زیر اندکی اصلاح کرد: «هستی‌شناسی به منزله مشخصه سازی رسمی از یک ادراک مشترک است». بورست، با افزودن ادراک مشترک به نکته‌ای

^۱ Gruber

^۲ Borst

خاص اشاره کرده است که در آن مشترک بودن مفهوم میان عده‌ای از افراد را مورد نظر قرار داده است. به واقع، برای توصیف حوزه موضوعی خاص با مفاهیم خاص نیاز است تا ادراک مشترکی از آن مفاهیم بین متخصصان یا استفاده‌کنندگان اصطلاحات آن حوزه به‌وجود آید.

استودر^۱ و همکارانش (۱۹۹۸) تعریف گروبر و بورست را با یکدیگر ادغام و تعریف جدیدی به شرح زیر ارائه کردند:

«هستی‌شناسی، مشخصه‌سازی رسمی و آشکار از ادراکی مشترک است. در این تعریف، ادراک به مدل انتزاعی از برخی پدیده‌های موجود در جهان با معرفی مفاهیم مربوط به آن پدیده‌ها اشاره دارد. آشکار، بدین معناست که انواع مفاهیم و محدودیت‌های حاکم بر استفاده از آن‌ها به وضوح توصیف و شرح داده شده است، رسمی، به این واقعیت بازمی‌گردد که هستی‌شناسی باید به وسیله ماشین قابل استفاده باشد. مشترک، همان‌طور که پیش از این اشاره شده، این عقیده را منعکس می‌کند که هستی‌شناسی دانشی را کسب می‌کند که مورد توافق و رضایت گروهی از افراد است نه مختص به چند فرد خاص». استودر در تکمیل تعاریف قبلی ارائه شده توسط سایر متخصصان، ابعاد کامل‌تری از هستی‌شناسی‌ها را بیان کرده و سعی در توسعه تعریف با ویژگی‌های خاص آن داشته است. مشخصه‌سازی، رسمی بودن، آشکار بودن و ادراک مشترک ویژگی‌هایی است که در تعریف استودر به آن اشاره شده است.

حوزه دیگری که به کاربرد هستی‌شناسی توجه نموده و حوزه‌ای وابسته به علوم رایانه در نظر گرفته می‌شود، حوزه هوش مصنوعی^۲ است. در این حوزه، وسیع‌ترین کاربرد لغت هستی‌شناسی برای بازنمون‌های رسمی آن چیزی است که برای ما انسان‌ها به منزله عقل سلیم تعریف می‌شود. در این مفهوم مقولات شامل چیزهایی مانند: فضا، زمان و ساختار اشیاء فیزیکی است. چنین بازنمون‌های رسمی در خلق عوامل هوشمند که بتواند وظایف خاصی را انجام دهد، مفید است. یک چنین هستی‌شناسی، یک هستی‌شناسی غیرزبانی است. کاربرد هستی‌شناسی در بازنمون مقولات (که محدود به یک نوع مقوله خاص نیست) آن را از کاربرد هستی‌شناسی در سایر حوزه‌ها متمایز می‌سازد.

در هر یک از فرایندهای مربوط به بازنمون مقولات، می‌توان مراحل این چنین را تصور نمود:

□ در ابتدا یک هستی‌شناسی به‌منظور بازنمایی دنیای واقعی برای سیستم طراحی می‌شود و بر

اساس آن موجودیت‌های مختلف، روابط ما بین هر یک و همین‌طور ویژگی‌های آن‌ها، برای

سیستم مدل‌سازی شده و هر موجودیت در یک سلسله‌مراتب از طبقات قرار می‌گیرد؛

^۱ Studer

^۲ Artificial intelligence

- در مرحله بعد اقدام به ساخت یک موتور استنتاج برای تعامل با هستی‌شناسی مذکور می‌شود؛
- در مرحله نهایی زبانی برای کار با هستی‌شناسی وضع و پیاده‌سازی می‌شود. این زبان، رابط خارجی هستی‌شناسی با دیگر سیستم‌هاست.

این فرایندها می‌تواند در طراحی ربات‌ها و ماشین‌های جستجوی مختلف به کار گرفته شود. هستی‌شناسی را ابزار سازماندهی منابع وب برای بازیابی تمام‌متن می‌دانند. چیزی که فراتر از رده‌بندی‌ها در وب کاربرد دارد. در وب، هستی‌شناسی، فهم مشترک از یک دامنه است. چنین فهم مشترکی برای غلبه بر تفاوت در اصطلاحات ضروری است. وجود اصطلاحات یکسان با معانی متفاوت در برنامه‌ها نیز مشکل دیگری است. چنین اختلافی با نگاشت اصطلاحات متناظر در یک هستی‌شناسی مشترک یا با تعریف نگاشت مستقیمی بین هستی‌شناسی‌ها قابل حل است. به این دلیل که هستی‌شناسی‌ها، به صورت گسترده برای اهداف مختلف (پردازش زبان طبیعی، مدیریت دانش، تجارت الکترونیک، یکپارچه‌سازی هوشمند اطلاعات، وب معنایی و غیره) در جامعه‌های مختلف (مهندسی دانش، مهندسی پایگاه داده و نرم‌افزار) به کار می‌روند. تعریف جدیدی از اصطلاح هستی‌شناسی برای استفاده در دیگر رشته‌ها ارائه کردند. تعریف آن‌ها به این شکل است:

یک هستی‌شناسی، شکل‌های گوناگونی را دربردارد، اما ضرورتاً شامل واژگان اصطلاحات و مشخص‌سازی معنای آن‌ها است. هستی‌شناسی، تعاریف و نشانه‌های چگونگی ارتباط مفاهیمی که است ساختاری از دامنه را دربردارد و دربردارنده تفاسیر ممکن اصطلاحات است.

مسئله اصلی که این دو محقق در تعریف خود مورد توجه قرار دادند، توجه به این واقعیت است که هستی‌شناسی سعی در مشخص‌سازی معنا دارد. این مشخص‌سازی با ارتباط بین مفاهیم و قواعدی که این مفاهیم را محدود می‌کند انجام می‌شود و باعث روشن شدن مفاهیم و جایگاه آن‌ها در حوزه موضوعی و در جهان دانش می‌شود.

تعاریف مختلف دیدگاه‌های مکمل و متفاوتی از یک واقعیت را فراهم می‌کنند. بعضی از این تعاریف، تعریف خود را بر اساس فرایند دنبال شده برای ایجاد یک هستی‌شناسی و استفاده از آن در برنامه‌ها ارائه می‌کنند، در حالی که تعاریف دیگر متأثر از فرایند توسعه آن است. نتیجه اصلی از این تعاریف، این است که می‌توان گفت، هدف هستی‌شناسی‌ها، گرفتن دانش توافقی در یک روش کلی است و این که ممکن است در برنامه‌های نرم‌افزاری یا گروه‌های افراد به اشتراک گذاشته شوند و مجدداً استفاده شوند. معمولاً هستی‌شناسی‌ها، توسط گروه‌های مختلفی از افراد در جاهای مختلف به صورت مشترک ایجاد می‌شوند. در واقع تعریف اصلی بدون توجه به کاربردهای خاص آن در

حوزه‌های مختلف، تعیین دامنه یک حوزه موضوعی، مفاهیم مربوط به یک حوزه و روابط بین این مفاهیم است.

در حوزه اطلاع‌رسانی نیز همچون سایر حوزه‌ها تعاریف متعددی از آن ارائه شده است؛ به طوری که به‌ندرت می‌توان تعریف یکسانی از آن به‌دست آورد. «سوا» (۱۹۹۹) در مورد هستی‌شناسی می‌گوید: «هستی‌شناسی پدیده‌ها یا عناصر موجود در یک حوزه خاص را دسته‌بندی می‌کند، بر اساس این بررسی، فهرستی از مفاهیم را ارائه می‌دهد که به تفصیل، انواع موضوعات و روابط منطقی میان آن‌ها را بیان می‌کند و از این رو هستی‌شناسی نامیده می‌شود» تمرکز این تعریف بر دسته‌بندی و وجود روابط منطقی است که می‌توان گفت به کاربرد هستی‌شناسی در علم اطلاع‌رسانی توجه دارد. به‌طور کلی می‌توان هستی‌شناسی را نوعی نقشه‌معنایی دانست که بر اساس سلسله‌مراتب و ژگانی تدوین شده و اغلب مفاهیمی را دربرمی‌گیرد که برای شناخت انسان از دنیای پیرامون، بنیادی و اساسی هستند، دربرگیرنده ویژگی‌های اصلی هستی‌شناسی است.

سورگل^۱، هستی‌شناسی را شکل گسترش‌یافته رده‌بندی، و هدف آن را در کمک به کاربر در بازیابی متون بر روی وب می‌داند. بنابراین، سایر حوزه‌ها مانند هوش مصنوعی، مهندسی نرم‌افزار و نظایر آن برای پاسخ به نیاز به رده‌بندی مفاهیم بر روی وب، به هستی‌شناسی‌ها روی آورده‌اند. سورگل با توجه به این نکته هستی‌شناسی را یک ابزار سازماندهی منابع وب که به بازیابی متون با زبان طبیعی کمک می‌کند، توصیف کرده است. اما وی نیز عقیده دارد که در واقع مشکل ناشی از سازماندهی اطلاعات نیست، بلکه این مجموعه اصطلاحات هستند که مشکل ساز شده‌اند و در نتیجه، زمینه ظهور هستی‌شناسی‌ها به‌صورت مجموعه لغات یا اصطلاح‌نامه‌های تکمیل‌شده یا گسترش‌یافته را فراهم کرده‌اند. این توصیف سورگل از هستی‌شناسی‌ها، بر اهمیت گردآوری مفاهیم با روابط صحیح و دقیق در اصطلاح‌نامه‌ها برای بازیابی منابع بر روی وب تأکید دارد. اگر با دیدگاهی که سورگل ارائه می‌دهد به هستی‌شناسی‌ها نگاه کنیم، می‌توان هستی‌شناسی را اصطلاح‌نامه‌ای غنی شده برشمرده که تعاریف و ارتباطات معنایی را دربردارد و نقش آن کاهش ابهام مفهومی و اصطلاح‌شناختی در حوزه‌ای خاص و تعریف شده است. بنابراین، در علم اطلاع‌رسانی نیز می‌توان برای بازنمون مدارک از هستی‌شناسی و این ساختار غنی‌شده استفاده نمود.

در نهایت، در ساده‌ترین تعریف، مفهوم هستی‌شناسی را می‌توان مانند اصطلاح‌نامه‌ها، مجموع لغات کنترل شده دانست؛ با این تفاوت که مفاهیم یک حوزه خاص با روابط معنایی خاص با توجه به حوزه موضوعی و بعضاً متفاوت از حوزه‌های موضوعی دیگر به شکل خاصی طبقه‌بندی می‌شوند که

¹ Soergel

هدف آن‌ها نمایش مفاهیم در قالب زبان طبیعی است. بدین سبب برخی آن را نوع تکامل یافته اصطلاح‌نامه و برخی دیگر آن را ابزاری کاملاً جدید می‌دانند. منطق به کار رفته در ساخت هستی‌شناسی‌ها منطبق ساختن کاربر با نظام سازماندهی دانش و نه انطباق نظام کاربر است. این نظام ساختار معنایی دقیق و تعریف شده بر اساس واقعیت‌های موجود یک حوزه به بیان روابط دقیق میان اصطلاحات در قالبی واقعی می‌پردازد و بدین ترتیب به کاربر در فهم و شناخت بیشتر یک حوزه کمک می‌کند. هستی‌شناسی‌ها همچنین روابط موجود در اصطلاح‌نامه‌ها را حمایت می‌کنند. این ابزار با فراهم‌آوری روابط میان اصطلاحات مانند اصطلاح‌نامه اما در قالبی واقعی و دقیق‌تر به گونه‌ای که معانی را نیز ارائه می‌کند به کاربر در این شناخت یاری می‌رساند. این روابط که به‌طور واضح نام‌گذاری می‌شوند از حوزه‌ای به حوزه دیگر متفاوت است برای مثال نوع روابط مبهم عام^۱ خاص^۲ و نیز مرتبط^۳ در اصطلاح‌نامه‌ها به روابط مفهومی «دارای نمونه»^۴، «یک ... است»^۵ و نظایر این‌ها تبدیل می‌شود. در واقع، تفاوت هستی‌شناسی‌ها با اصطلاح‌نامه‌ها و واژگان کنترل شده، نمایش مفاهیم در قالب زبان طبیعی است. روابط معنایی دقیق‌تر و واقعی موجود در هستی‌شناسی آن را به زبان طبیعی نزدیک‌تر می‌کند.

بدین ترتیب مندرجات هستی‌شناسی‌ها ترکیبی از لغات کنترل شده و محدود و قابل گسترش، رده‌ها و روابط بین اصطلاحات، روابط فرعی سلسله‌مراتبی صریح و اصول منطقی برای ایجاد روابط توسعه ماشین است.

هستی‌شناسی‌ها در برخورد با مفاهیم، به شکل متفاوتی از سایر واژگان کنترل شده (مثل اصطلاح‌نامه و فرهنگ) رفتار می‌کنند. مفاهیم و هستی‌شناسی‌های دربردارنده مفاهیم در ایجاد تصور ذهنی نسبت به مفهوم نقش مهمی دارند.

به‌طور کلی، مؤلفه‌های عمومی انواع مختلف هستی‌شناسی‌ها را می‌توان به‌صورت زیر تعریف کرد:

- کلاس: کلاس‌ها برای نمایش مفاهیم به کار می‌روند. برای مثال کلاس خودرو که می‌تواند شامل زیرکلاس‌هایی مانند خودروی سبک و سنگین باشد؛
- رابطه: رابطه بیانگر ارتباط مفاهیم با یکدیگر در یک قلمرو است؛
- صفات: ویژگی‌ها و مشخصاتی که اشیاء یا کلاس‌ها می‌توانند داشته باشند؛

¹ BT (Boarder Term)

² NT (Narrower Term)

³ RT (Related Term)

⁴ Has instance

⁵ Is a

- نمونه‌ها: عناصر و نمونه‌های هر کلاس هستند؛
 - توابع^۱: توابع نمونه‌هایی از روابط هستند. برای مثال تابع پرداخت، قیمت یک اتاق را پس از اعمال تخفیف روی آن محاسبه می‌کند؛
 - اصول موضوعه^۲: اصول موضوعه برای مدلسازی جملاتی به کار می‌روند که همواره صحیح هستند. اصول موضوعه زمانی برای بازنمایی دانش به کار گرفته می‌شوند که مؤلفه‌های دیگر قادر به انجام آن نباشند. اصول موضوعه برای استنتاج دانش جدید و همچنین سازگاری هستی‌شناسی استفاده می‌شود.
- همان‌طور که مؤلفه‌های هستی‌شناسی نشان می‌دهند، تفاوت ساختار هستی‌شناسی و اصطلاح‌نامه به صورت مشخص، به تصویر کشیده شده است. وجود کلاس‌ها و زیرکلاس‌ها، وجود روابطی که محدود به روابط اصطلاح‌نامه‌ای نیست و وجود صفات و نمونه‌ها، ساختار معنایی متفاوتی با اصطلاح‌نامه‌ها ایجاد می‌کند.

¹ functions

² axiom

۵-۲. نتیجه‌گیری

طرح تبدیل اصطلاح‌نامه‌ها به هستی‌شناسی، با تبدیل اصطلاح‌نامه شیمی پژوهشگاه علوم و فناوری اطلاعات (ایرانداک) در شهریورماه ۱۳۹۹ آغاز گردید. هستی‌شناسی ایجادشده توسط نرم‌افزار پروتزه قابل بازیابی، ویرایش و ذخیره می‌باشد. این پروژه با مطالعه بر روی ساختار داخلی فایل‌های هستی‌شناسی ساخته شده توسط نرم‌افزار پروتزه آغاز شد و در پیاده‌سازی نهایی مورد استفاده قرار گرفت.

در ادامه تبدیل اصطلاح‌نامه‌ها به هستی‌شناسی دومین موردی که به‌صورت رسمی تبدیل به هستی‌شناسی شد، اصطلاح‌نامه علوم زیستی بود که در شهریورماه ۱۴۰۱ کار بر روی آن به اتمام رسید. در این پروژه علاوه بر مواردی که برای تبدیل اصطلاح‌نامه شیمی به هستی‌شناسی استفاده شد، بیش از ۱۸۰ هستی‌شناسی خارجی مورد مطالعه قرار گرفت تا بتوان حداکثر تعداد روابط معنایی دودویی بین مفاهیم را استخراج نمود. برای تبدیل اصطلاح‌نامه به هستی‌شناسی از روشی که توسط سورگل و همکاران (۲۰۰۴) ذکر کرده، استفاده شده است و مراحل زیر را شامل می‌شود:

۱. مفاهیم تبدیل به کلاس شدند؛
۲. مفاهیم اعم تبدیل به کلاس بالاتر یا سوپرکلاس شدند؛
۳. مفاهیم اخص تبدیل به زیرکلاس یا ساب کلاس شدند؛
۴. اصطلاح‌های نامرجح تبدیل به اینستانس یا اصطلاح منفرد شدند؛
۵. اصطلاح‌های مرتبط و ترجمه تبدیل به ویژگی‌ها شدند.

برای ساخت هستی‌شناسی ابتدا کدنویسی‌هایی که در هستی‌شناسی‌ها قابل تطبیق با نرم‌افزار پروتزه بودند، مورد مطالعه قرار گرفتند. به‌همین منظور از ابتدا کدهای نوشته‌شده در هستی‌شناسی علوم زیستی شروع به گسترش داده می‌شوند. در این کدهای اولیه ویژگی‌های کلی فایل هستی‌شناسی مشخص می‌گردد. در فاز دوم پروژه سعی بر استخراج روابط دودویی تعریف‌شده در هستی‌شناسی علوم زیستی می‌شود که قبلاً در کشورهای دیگر توسعه یافته‌اند. برای استخراج روابط دودویی غنی‌سازی هستی‌شناسی به مجموعه ۱۸۳ هستی‌شناسی با بیش از هفت میلیون مفهوم می‌رسد. پس از بررسی روابط معنایی، به فهرستی از روابط معنایی دودویی دست یافتیم که کلاس‌های هر دو طرف آن‌ها در اصطلاح‌نامه علوم زیستی ایرانداک یا آنتولوژی ساخته شده وجود دارند. این روابط در پیوست گزارش به همراه نام اختصاری آنتولوژی خارجی که آن روابط در آن‌ها یافت شده‌اند، آورده شده است. این روابط شامل ۸۲۰ مفهوم از مجموع ۵۰۰۰ مفهوم اصطلاح‌نامه علوم زیستی می‌باشد. بنابراین برای تعداد مناسبی از مفاهیم روابط دودویی استاندارد تو سعه داده شد. دلیل این که تعداد باقیمانده

از مفاهیم درون آنتولوژی‌های خارجی یافت نشدند، می‌تواند این مورد باشد که ترجمه مفاهیم در اصطلاح‌نامه علوم زیستی یکسان‌سازی نشده‌اند، بدین معنی که برای هر مفهوم در همه آنتولوژی‌ها و اصطلاح‌نامه‌ها یک اصطلاح یکتا به کار برده شوند.

بنابراین بایستی مجدداً ترجمه مفاهیم درون اصطلاح‌نامه را بررسی کرد تا از یکدست بودن آنها اطمینان حاصل شود. برای این منظور هم از روش‌های خودکار و بر مبنای هوش مصنوعی می‌توان بهره گرفت و غنی‌سازی روابط معنایی آنتولوژی را توسعه داد.

فهرست منابع و مآخذ

آسوشه، عباس و همکاران. ۱۳۹۰. هستان نگاری، روش شناسی‌ها، ابزارها و زبان‌های توسعه، انتشارات دانشگاه تربیت مدرس

اکبری، اسماعیل. اصطلاح‌نامه جامع، علوم زیستی. ۱۳۸۴. مرکز اطلاعات و مدارک علمی ایران. تهران.

پالمر، فرنک ر. ۱۳۸۰. نگاهی تازه به معنی‌شناسی. ترجمه‌ی کورش صفوی. تهران: نشر مرکز، کتاب ماد.

ترابی، لاله. ۱۴۰۰. طراحی الگوی هستان‌شناسی پزشکی هسته‌ای. پایان‌نامه دکتري. دانشگاه آزاد اسلامی.

حری، عباس. ۱۳۹۱. اصطلاح‌نامه و هستی‌نگاشت. کلیات کتاب ماه، سال ۱۶، شماره ۲، بهمن ۱۳۹۱.

حسینی بهشتی، ملوک‌السادات. ۱۳۹۲. ساختواژه، اصطلاح‌شناسی و مهندسی دانش. پژوهشگاه علوم و فناوری اطلاعات

حسینی بهشتی، ملوک‌السادات. ۱۳۹۱. ایجاد هستی‌شناسی واژگانی از اصطلاح‌نامه‌های علوم پایه. همایش زبان‌شناسی. بهمن ۱۳۹۱.

حسینی بهشتی، ملوک‌السادات. ۱۳۷۸. اصطلاح‌نامه جامع (علوم پایه)، مجموعه مقالات نخستین هم‌اندیشی مسائل واژه‌گزینی و اصطلاح‌شناسی، فرهنگستان زبان و ادب فارسی، مرکز نشر دانشگاهی.

حسینی بهشتی، ملوک‌السادات و همکاران. ۱۳۹۳. طراحی و پیاده‌سازی هستی‌شناسی علوم پایه بر اساس مفاهیم و روابط موجود در اصطلاح‌نامه‌های مرتبط، پژوهشنامه پردازش و مدیریت اطلاعات.

حسینی بهشتی، ملوک‌السادات و همکاران. ۱۳۹۸. توسعه و روزآمدسازی اصطلاح‌نامه‌های ایرانداک، تهران. پژوهشگاه علوم و فناوری اطلاعات. تهران.

حسینی بهشتی، ملوک‌السادات و همکاران. ۱۳۹۳. غنی‌سازی و پالایش روابط معنایی در هستان‌شناسی علوم پایه با تمرکز بر رشته شیمی، مقاله ارائه شده در سومین همایش ملی زبان‌شناسی رایانشی، دانشگاه صنعتی شریف، تهران.

حسینی بهشتی، ملوک‌السادات. ۱۳۹۱. تبدیل اصطلاح‌نامه به هستان‌شناسی. کلیات کتاب ماه. سال ۱۶، شماره ۲، بهمن ۱۳۹۱.

داورپناه، محمدرضا. ۱۳۸۴. ضرورت‌های نوین بازنگری در ذخیره و بازیابی اطلاعات. کتابداری و اطلاع‌رسانی، ۳ (۸): ص ۶۷-۸۸.

رجبی، تقی. ۱۳۸۲. اصطلاح‌نامه جامع: شیمی، مرکز اطلاعات و مدارک علمی ایران. تهران.

ایجاد هستی‌شناسی از اصطلاح‌نامه علوم زیستی ایرانداک □ ۱۰۰

رجبی، تقی. ۱۳۸۲. اصطلاح‌نامه شیمی، مقاله ارائه شده در دومین هم‌اندیشی مسائل واژه‌گزینی و اصطلاح‌شناسی، فرهنگستان زبان و ادب فارسی. تهران.

رجبی، تقی. ۱۳۹۵. طراحی ساختار درختی و تدوین اصطلاح‌نامه شیمی از واژگان مصوب فرهنگستان، پژوهشگاه علوم و فناوری اطلاعات. تهران

رجبی، تقی و همکاران. ۱۳۹۸. روزآمدسازی و توسعه اصطلاح‌نامه‌های علمی و فنی ایرانداک. فصلنامه مدیریت اطلاعات، ۵ (۱): ۹۹-۱۱۸.

زندى روان، نرگس. ۱۳۸۵. تجزیه قراردادی مفهوم زیرساخت پردازش زبان طبیعی در رایانه. در اولین همایش سراسری دانشجویی کتابداری و اطلاع‌رسانی پزشکی ایران. دانشگاه علوم پزشکی اصفهان.

شریف، عاطفه. ۱۳۸۵. شناختی از روابط معنایی در هستی‌شناسی وب. اطلاع‌شناسی، ۴ (۱ و ۲): ۸۴-۶۵.

صفری، مهدی. ۱۳۸۳. مدلسازی مفهومی در بازنمون رسمی دانش: شناختی از هستی‌شناسی در هوش مصنوعی و نظام‌های اطلاعاتی. اطلاع‌شناسی، ۴ (۲): ۷۵-۱۰۴.

صفوی، کورش. ۱۳۸۳. درآمدی بر معنی‌شناسی. تهران: سوره مهر (حوزه‌ی هنری سازمان تبلیغات اسلامی).

صنعت‌جو، اعظم و اکرم فتحیان. ۱۳۹۱. روش‌شناسی طراحی، ساخت و پیاده‌سازی هستی‌شناسی: رویکردها، زبان‌ها و ابزارها (مطالعه موردی طراحی هستی‌شناسی ASFAONT در حوزه کتابداری و اطلاع‌رسانی). فصلنامه کتابداری و اطلاع‌رسانی، ۱۵ (شماره ۱ (پیاپی ۵۷)): ۱۱۳-۱۴۲.

فلبر، هلموت. ۱۳۸۱. مبانی اصطلاح‌شناسی. ترجمه محسن عزیزی. تهران: مرکز اطلاعات و مدارک علمی ایران.

گراسمن، دیوید و افیر فریدر. ۱۳۸۴. بازیابی اطلاعات: الگوریتم‌ها و روش‌های اکتشافی. ترجمه جعفر مهرداد، سارا کلینی. مشهد: انتشارات کتابخانه رایانه‌ای؛ شیراز: کتابخانه منطقه‌ای علوم و تکنولوژی شیراز.

لاینز، جان. ۱۳۸۳. مقدمه‌ای بر معناشناسی زبان شناختی. ترجمه حسین واله. تهران: گام نو.

محمدی، مسعود. ۱۴۰۰. طراحی چارچوب مفهومی نظام بازیابی اطلاعات در حوزه گوارش و کبد. پایان‌نامه دکتری. دانشکده پیراپزشکی / دانشگاه نهران.

مهرداد، جعفر و محمدرضا فلاحتی فومنی. ۱۳۸۴. معناشناسی و بازیابی اطلاعات: هفت گفتار. مشهد: انتشارات کتابخانه رایانه‌ای؛ شیراز: کتابخانه منطقه‌ای علوم و تکنولوژی شیراز.

نشاط، نرگس. ۱۳۸۵. نقش زبان در بازنمایی معنا. اطلاع‌شناسی، ۴ (۱ و ۲): ص ۱۱۹-۱۳۰.

هماوندی، هما. ۱۳۹۸. طراحی چارچوب هستی‌شناسی رفتار اطلاعاتی کاربران. پایان‌نامه دکتری. دانشکده مدیریت / دانشگاه تهران.

Akbari, F. (2020) Iran's Language Planning Confronting English Abbreviations, Springer Briefs in Linguistics, Switzerland.

Almeida, M.B. (2013) Revisiting Ontologies: A Necessary Clarification, Journal of the American Society for Information Science and Technology, 64 (8):1682-1693.

- Amirhosseini, M.; Salim, J. (2019) A Synthesis Survey of Ontology Evaluation Tools, Applications and Methods to Propose a Novel Branch in Evaluating the Structure of Ontologies: Graph-Independent Approach, *International Journal of Computer (IJC)* Vol. 33, No.1, pp 46-68.
- Amirhosseini, M.; Salim, J. (2019) Structural Analysis of Semantic Relations Regarding Integration and Association of Semantic Network in Vocbench as an Agricultural Ontology, *International Journal of Engineering Technologies and Management Research*, Vol. 6, (Iss. 5).
- Andrew, A. M. (2004). "A Review of Knowledge Representation, Reasoning and Declarative Problem Solving" *Cybernetics*, 33 (1): 147-150. Retrieved Sep, 5, 2007, from [http://www.cybsoc.org/knowrep \(K\). HTML](http://www.cybsoc.org/knowrep (K). HTML).
- Asadi, s., & Jamali M., H. R. (2004). "Shifts in Search Engine Development: A Review of Past, Present and Future Trends in Research on Search Engines". *Webology*, 1 (2). Retrieved feb, 18, 2005, From <http://www.webology.ir/2004/v1n2/a6.html>.
- Asanee, Kawtrakul, et al (2005). "Automatic Term Relationship Cleaning and Refinement for AGROVOC (online). available: <ftp://ftp.fao.org/docrep/fao/008/af240e/af240e00.pdf>
- Berners-Lee, T. et al (2001) "The Semantic Web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities". Retrieved Oct, 5, 2006, from <http://www.w3c.org/2001/sw.html>
- Bitters, Barry, (2005). "A Geographical Ontology of Objects in the Visible Domain". A dissertation for a degree of doctor of philosophy. The Florida State University College of Social Science.
- Brewster, c.l, O'Hara, K. (2007). "Knowledge Representation with Ontologies: Present challenges-future Possibilities: *Human-computer Studies* 65: 563-568. Retrieved May, 5, 2007, from Elsevier Database.
- Broughton, Vanda (2006). "Essential Thesaurus Construction. London: Facet Publishing.
- Budin, G. and E. Oeser (1997). "Beitraege zur Terminologie and Wissenstechnik". Vienna, TermNet.
- Cimatti, F. (2018). *A Biosemiotic Ontology*. Springer Nature Swizerland.
- Corcho, O., Fernandez-Lopez, M., & Gomez-Perez, A. (2007). Ontological engineering: what are ontologies and how can we build them?. In *Semantic web services: Theory, tools and applications* (pp. 44-70). IGI Global.
- Daconta, M.C., et al (2003). "The Semantic web: a guide to future of XML web services and Knowledge management". Indianapolis: Wiley Publishing, INC.
- Davies, J., et al (2007). "Semantic Web Technologies: Trends and Research in Ontology-based Systems". John Wiley & Sons.
- Dean, M. et al (2004) "OWL Web Ontology Language Reference" Retrieved June, 5, 2006, from <http://www.w3.org/TR/owl-ref/>
- Dimitriadis, A.; Saulwick, A.; Windhouwer, M. (2005) "Semantic Relations in Ontology Mediated Linguistic Data Integration". Retrieved May, 5, 2007, From <http://home.medewerker.uva.nl/a.g.saulwick/bestanden/SemanticRelations-EMELD-2005.pdf>
- Ejei, F.; M.S., Hosseini Beheshti; T. Rajabi; Z. Ejei (2017) Enriching Semantic Relations of Basic Sciences Ontology. *Knowledge Organization*. 44 (5) 312-325.
- Fernández-López, M., Gómez-Pérez, A. and Juristo, N. (1997, March). METHONTOLOGY: From Ontological Art Towards Ontological Engineering. In *AAAI-97 Spring Symposium Series*, Stanford University, EEUU .33-40.
- Fernandez-Lopez, M.; Gomez-Perez, A.; (1999) Building a Chemical Ontology Using Methontology and the Ontology Design Environment.
- Gomez-Perez et al (2004) *Ontological Engineering*. Springer Verlag.

- Gomez-Perez, A., Fernández-López, M., & Corcho, O. (2006). *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer Science & Business Media.
- Goodarzi, N., Norouzi, Y., Hosseini Beheshti, MS. (2022) Folksonomy-based Ontology: A link between intelligence and meaning in the field of psychology and educational sciences, *Sciences and Techniques of Information Management*.
- Grimm, S.; Abecker, A.; Voelker, j.; Studer, R. (2011) *Ontologies and the Semantic Web*, Handbook of Semantic Web Technologies, Springer Verlag
- Gruber, T. (1995). "Towards Principles for the Design of Ontologies used for Knowledge Sharing". *International Journal of Human Computer Studies*, 43: 907-928.
- Gruber, T.R. (1993). "A Translation Approach to Portable Ontology Specifications". *Knowledge Acquisition*. No. 5, P.: 199-220.
- Guarino, N. (1998, June). Formal ontology and information systems. In *Proceedings of FOIS (Vol. 98 (1998), 81-97)*. Retrieved from <http://www.loa.istc.cnr.it/old/Papers/FOIS98.pdf> [Accessed: 13.5.2016].
- Guarino, N. and Giaretta, P. (1995). *Ontologies and knowledge bases towards a terminological clarification. Towards very large knowledge bases: knowledge building & knowledge sharing*, 25-32. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.320.8006&rep=rep1&type=pdf> [Accessed: 20.10.2016].
- Guarion, Nicola. (1999). "The Role of Conditions in Ontology Design". *Proceeding of the IJCAI-99 Workshop on Ontologies and Problem Solving Method (KRR5)*. P.: 1-7. Available at: <http://ladesb.pd.cnr.it/infor/ontology/ontology.html> (Date: 2006, 25, Dec)
- Hassanzadeh, M. (2022) *Metaverse, Meta university and the Future of Higher Education*. *Sciences and Techniques of Information Management*, 8(2), 7-22.
- Hess, C.; Schleider, c. (2006) "Ontology-based Verification of Core Model Conformity in Conceptual Modeling". *Computers, environment and Urban Systems*, 30: 543-561. Retrieved May, 5, 2007, from Elsevier Database.
- Heylighen, f. (2001) "Bootstrapping Knowledge representation from Entailment Meshes via Semantic Nets to Learning Web" *Cybernetics*. 30 (5/6): 691-725. Retrieved Des, 20, 2006, From Emerald Database.
- Heylighen, f; & Bollen, J. (1996). "The World-Wide Web as a Super- Brain: from metaphor to model" *Cybernetics and Systems*. Retrieved Des, 15, 2006, From <http://pespmc1.wub.ac.be/Papers/wwwsuperBRAIN.html>
- Hjorland B. (2004) "Theory of Knowledge Organization and the Feasibility of Universal Solutions" 8 International ISKO Conference, London, 16th 2004, retrieved Apr, 4, 2007, From: <http://dlist.sir.arizona.edu/389/01/hjorlandisko8.pdf>.
- Hjorland, B. (2015) *Theories are Knowledge Organizing Systems (KOS)*, *Knowledge Org.* 42(2015) No.2, 113-128.
- Hjorland, B. (2018) *Reviews of Concepts in Knowledge Organization*, *Knowledge Organization*, 45 (2018) No.2.
- Huang, Jin-Xia et al (2005) *Enriching Core Ontology with Domain Thesaurus through Concept and Relation Classification*
- ISOCO (2006) "HOW TO UNDERSTAND THE SEMANTIC WEB". Retrieved July, 20, 2006, From http://www.isoco.com/in/innovation_ws_how.html

- Kalibatiene, D., & Vasilecas, O. (2011, October). Survey on ontology languages. In International Conference on Business Informatics Research (pp. 124-141). Springer, Berlin, Heidelberg.
- Kang, Sin-Jae; dong-II and Lee, Jong-Hyeok. (2001). Semi automatic Ontology Construction by Using Thesaurus, computational dictionaries and Large corpora. ACL 2001 Workshop on Human Language Technology and Knowledge. Available at:
<http://truth.Yust.edu/~dongil/research/ontology.pdf> (Date: 2006, 25, Dec).
- Khosravi, F.&A. Vazifedoost (2008)Creating a Persian Ontology through Thesaurus Reengineering for Organizing the Digital Library of the National Library of Iran. In: Faslname Ketab, 70, 19-36.
- Mancini, c.; Shum, s.J. B. (2006) Modeling discourse in contested domains: a semiotic and cognitive framework. Intelligent Journal of Human-Computer Studies. 64: 1154-1171. Retrieved Des, 20, 2006, from Elsevier Database.
- McGuinness, D.L. & Harmelen, f.V. (2004) "OWL web Ontology Language Overview". Retrieved Jane, 5, 2006, From <http://www.w3.org/TR/owl-features/>
- Miller, G.; Beck with, C.; Fellbaum, r.et al. (1990) Introduction to Word Net: An Online Lexical Database International Journal of Lexicography 3, 4: 235-44.
- Mondal, P. (2020). Language, Biology and Cognition. A Critical Perspective. Springer Nature Switzerland.
- Nichoson, D, Dunsire, g, and Neill, S. (2002). HILT: Moving towards interoperability in subject terminology. Journal of Internet Cataloging. Vol. 5, No. 4, P.: 97-111.
- Noy, N. F. & McGuinness, D. L. (2001). Ontology Development 101: A Guide to Creating Your First Ontology. Retrieved from
http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html[Accessed: 12.5.2016].
- Noy, N.F.; Mac Guinness(2000) Ontology Development101:A Guide to Creating Your First Ontology, Stanford University
- Obitko, Marek; Snasel, Vaclav; and Smid, Jan. (2004). Ontology Design with formal Concept Analysis. CLA. P.: 111-1190 Available at:
<http://Ftp. Informatik.rwth-aachen.de/publications/CEUR-WS/Vol-110/paper12.pdf> (Date: 2006, 25, Dec)
- Ogden, C. K., & Richards, I. A. (1923). The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism (Vol. 29). K. Paul, Trench, Trubner & Company, Limited.
- Pangaro, P. (2006) "Cybernetics: a Definition". Retrieved Jane, 5, 2007, from
<http://pangaro.com/published/sybermacmilan.html>
- Preece, H. (2003) "OWL: The Web Ontology Language". Retrieved Jun, 20, 2006, From
<http://www.csd.abdn,as.uk/~apreece/talks/OWLtheWebOntologyLanguage/OWLThewebOntologyLanguage.pdf>
- Prevot, L. et al (2010) Ontology and the Lexicon. A Multidisciplinary Perspective. In: Ontology and Lexicon. Cambridge Uni. Press.
- Pulido, J.R.G. et al (2006). "Ontology languages for the semantic web: a never completely updated review" Knowledge-Based Systems, 19: 479-497. Retrieved may, 5, 2007, from Elsevier Database.
- Raskin, R. G.; Pan, M.J. (2004) "Knowledge Representation in the Semantic Web for Earth and Environmental Terminology (SWEET)". Computers & Geosciences, 31: 1119-1125. Retrieved Jun, 15, 2007, From Elsevier Database

- Roussey, C., Pinet, F., Kang, M. A., & Corcho, O. (2011). An introduction to ontologies and ontology engineering. In *Ontologies in Urban Development Projects* (pp. 9-38). Springer London.
- Sattar Chaudary, Abdus, Pei Jiun, Tan. (2005). Enhancing access to digital information resources on heritage: a case of development of taxonomy at the Integrated Museum and Archives Systems in Singapore. *Journal of Documentation*. Vol. 61, No 6, P.: 751-7760.
- Sawsaa, A. 2013. A Generic Model of Ontology to Visualize Information Science Domain (OIS). PhD thesis, University of Huddersfield.
- Schwarz, K. (2005). Domain model enhanced search: A comparison of taxonomy, thesaurus and ontology. Unpublished Master's Thesis, University of Utrecht, MA. Retrieved Sep, 4, 2007, From www.cwi.nl/~media/publications/masterthesis_kat_domainmodel_2005.pdf
- Shamsfard, M. et al (2006) An Ontology Based System for Ranking Documents. In: *International Journal of Computer Science*, Vol. 1, N. 3.
- Shamsfard, M. et al (2010) Ontology Population and Semantic Annotation of Persian Texts. In: *Computer Science and Engineering (CSI)*.
- Shiri, A. & C., Revie (2000) Thesauri on the Web: Current Developments and Trends. *Online Information Review* 24 (4): 273-280.
- Shiri, A. et al (2011) An Evaluation of Thesaurus-Enhanced Visual Interfaces for Multilingual Digital Libraries. *TPDL* 2011: 239-243.
- Shiri, A. et al (2011) User Evaluation of Searching: A Visual Interface for Bilingual Digital Libraries. *The Electronic Library* 29 (1) 71-89.
- Smiraglia, R.P. (2014) *The Elements of Knowledge Organization*, Springer International Publishing Switzerland
- Soergel, D. et al (2004). Reengineering Thesauri for new Application: the AGROVOC Example" *Journal of Digital Information*, 4 (4). Retrieved Des, 15, 2006, From <http://Jodi.ecs.soton.ac.uk/Articles/v04/i04/soergel/>
- Soergel, Dagobert, Boris Lauser, 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. (1993). Review of the book *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project*, by D.B. Lenat and R.V. Guha. *Artificial Intelligent*. 61(1), 95-104. Retrieved from <http://www.sciencedirect.com/science/article/pii/000437029390096T> [Accessed: 5.9.2016].
- Sowa, J. F. (2000). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co. Pacific Grove, CA, Retrieved from <http://www.jfsowa.com/ontology/index.htm> [Accessed: 5.9.2016].
- Sowa, J. F. (2010) The Role of Logic and Ontology in Language and Reasoning. Chapter 11 of *Theory and Applications of Ontology. Philosophical Perspectives*. Edit. by R. Poli & J. Seibt. Berlin: Springer, 2010, pp. 231-263.
- Steenmsen, M. (2003) Digital CULT's Expert 13 Tanole with the Semantic Web. Towards a Semantic Web for Heritage resources, *digi CULT thematic Issue* 3, P.:14-20. Available at: <http://www.digicult. Iffo/pages/Themiss.php> (Date: 2006, 25, Dec)
- Studer, R.; Benjamins, V.R.; Fensel, D.; (1998) *Knowledge Engineering: Principles and Methods*, In: *Data & Knowledge Engineering*, 25 (1998) 161-197.
- Svenonius, e. (2004) "The Epistemological Foundation of Knowledge Representation". *Library Trend*, 52 (3): 571-587. Retrieved Des, 15, 2006, From Elsevier Database.

- Torabi, L. , Mihosseini, Z., Abazari, Z., Hosseini Beheshti, MS. (2022) Evaluating the Efficiency of Nuclear Medicine Ontology in Knowledge Representation and Concepts Retrieval by Ontometric approach, Sciences and Techniques of Information Management.
- Torabi, Laleh, Mirhosseini, Zohre, Hosseini Beheshti, Molouk Sadat (2021) Design and Construction of Nuclear Medicine Ontology, Health Information Management, Vol.18/2, 81-88.
- Uschold, M., & Gruninger, M. (1996). Ontologies: Principles, methods and applications. The knowledge engineering review, 11(02), 93-136.
- Van Heijst, G., Schreiber, A. T., and Wielinga, B. J. 1996. Using Explicit Ontologies in KBS Development. International Journal of Human and Computer Studies.
- Villazon-Terrazas et al (2010) A Pattern-Based Method for Reengineering Non-Ontological Resources into Ontologies. In: International Journal on Semantic Web and Information Systems, 6 (4), 27-63.
- Weinbrenner, S. & J. Engler (2011) SCY Ontology and Metadata Scheme, DIV.2, SCY Consortium.
- Weng, S. et al (2006). "Ontology Construction for Information Classification". Expert Systems with Application, 31: 1-12. Retrieved May, 5, 2007, From Elsevier Database.
- Wielinga, B.J. and et all. (2001). From thesaurus to Ontology. In International conference On Knowledge. Canada: ACM Special Interest Group on Artificial Intelligence. P.: 194-201. Available at: <http://www.cs.vu.nl/~guus/papers/Wielinga.1a.pdf> (Date: 2006, 25, Dec)
- Yao, Haining; Orme, Anthony Mark; and Lutzkorn, Letha. (2005). Cohension Metric for Ontology Design and Application. Journal of computer Science. Col. 1, P.: 107-113. Available at: <Http://www.scipub.org/fulltext/jcs/jcs11107-113.pdf> (Date: 2006, 25, Dec)

Abstract

Ontology is the science of recognizing and categorizing the concepts that already exist or might exist in different areas. Generally, the science of ontology stems from philosophy and tries to recognize the concepts, and in the continuation of Aristotelian view, tries to categorize and classify existing concepts.

In Ontology, it is endeavored that the concepts of a specialized field are categorized and the concepts' classifications, the semantic relations between them, their traits and functions, the circumstances that can happen for a concept, the current rules for semantic patterns, the limitations of concepts and the conditions for analogy between concepts are determined. Therefore, it is made possible to create the ontology needed in information retrieval systems and a more accurate and faster search by looking at the needed concepts. To create an ontology, a developed lexicon or a thesaurus is used. This lexicon not only includes the meaning of the phrases, but also specifies the phrases that are somehow connected to the intended concept logically, and therefore extracts all the intended concepts in their own specific discussion, including objects or people and traits from a lexicon or a thesaurus, and creates the ontology. Eventually, by creating ontology and coding the logical patterns, a shared understandable language for humans and computers is achieved.

The goal for creating the biology ontology and using it in the information retrieval system is to give structure to information and eventually representing information with the help of the computer using the technology and standards which are being coded and completed by the worldwide web consortium. In other words, in semantic web, the computers participate in the process of analyzing and extracting the intended information by the user through identifying the existing documents and function more intelligently in this area. In semantic web, the search engines identify the semantic relations of concepts and separating the data more than before and use precise and powerful tools such as ontology to achieve that.

Continuing the conversion of thesauri to ontologies, we develop concepts and expand their semantic relations through the development of the biology ontology using IRANDoc's biology thesaurus and conforming it with the General Ontology for Biology which is accessible on the Bioportal Ontologies website. In order to do that, the protégé software is used, which is accessible and usable for Persian language, and a biology ontology with 5000 entries is developed.

Keywords: Thesaurus, Ontologies, Biology, Information Retrieval System, Semantic Web.



**Iranian Research Institute
for Information Science and Technology
(IranDoc)**

Research Project

Building Ontology from IRANDOC Bioscience Thesaurus

By:

Molouk Sadat Hosseini Beheshti

Co-Operators:

Esmail Akbari

Omid Ghasvand

February 2022