

ارزیابی هستی‌شناسی‌ها: بررسی معیارها، رویکردها و سطوح

اکرم فتحیان دستگردی* | دانشجوی دکترای رشته کتابداری و اطلاع‌رسانی،
دانشگاه فردوسی مشهد

دریافت: ۱۳۸۸/۰۵/۲۸ | پذیرش: ۱۳۸۸/۰۷/۲۵

فصلنامه علمی پژوهشی
پژوهشگاه علوم و فناوری اطلاعات ایران
شاپا(چاپی) ۸۲۲۳-۲۲۵۱
شاپا(الکترونیکی) ۸۲۳۱-۲۲۵۱
نمایه در SCOPUS، LISA و ISC
<http://jlist.irandoc.ac.ir>
ویژه‌نامه ذخیره، بازیابی و مدیریت اطلاعات
ص-ص ۳-۲۵ زمستان ۱۳۹۰

نوع مقاله: مروری

*fathian2000@gmail.com

چکیده: هستی‌شناسی، دانش ساختاریافته در مورد یک حوزه خاص است که از طریق ارائه مفاهیم و روابط میان آنها در آن حوزه شکل می‌گیرد. امروزه، به دلیل طراحی هستی‌شناسی در حوزه‌های موضوعی مختلف، ارائه معیارهایی برای گزینش مناسب‌ترین هستی‌شناسی، ضروری است. مقاله حاضر در نظر دارد معیارها و رویکردها و سطوح مختلف ارزیابی هستی‌شناسی را بررسی کند. بدین منظور، ابتدا مراحل مختلف ارزیابی هستی‌شناسی‌ها بررسی شد. سپس، مهمترین رویکردها در ارزیابی هستی‌شناسی‌ها، شامل رویکرد استاندارد طلایی، رویکرد وظیفه‌محور، رویکرد استخراج داده، و رویکرد معیارمحور، مطرح گردید. معیارهای ارزیابی هستی‌شناسی شامل معیارهای مطرح‌شده توسط صاحب‌نظران این حوزه و نیز معیارهای مؤسسه تحقیقات هستی‌شناسی امریکا، در بخش دیگری از مقاله ارائه شد. سطوح مختلف ارزیابی هستی‌شناسی شامل سطح واژگانی و داده‌ای، سطح ساختار سلسله‌مراتبی و رده‌بندی، سطح روابط معنایی، سطح بافت و کاربرد هستی‌شناسی، سطح نحوی هستی‌شناسی، و سطح معماری و ساختار هستی‌شناسی نیز در بخش پایانی مقاله بررسی گردید.

کلیدواژه‌ها: هستی‌شناسی، معیارهای ارزیابی هستی‌شناسی، رویکردهای ارزیابی هستی‌شناسی، سطوح ارزیابی هستی‌شناسی، مراحل ارزیابی هستی‌شناسی

۱. مقدمه

امروزه، تمرکز نظام‌های اطلاعاتی جدید از پردازش داده به پردازش معنی تغییر یافته است (Brank, Grobelnic, and Mladenic 2005). به همین دلیل، در نظام‌های کنونی، مبنای پردازش، به جای داده، یک مفهوم معنایی است که تفسیر خاص خود را دارد و درون متن در کنار سایر مفاهیم قرار می‌گیرد و با آنها دارای ارتباط است.

هستی‌شناسی نیز به عنوان یک ابزار معنایی، عبارت است از دانش ساختاریافته در مورد حوزه‌ای خاص که از طریق ارائه مفاهیم و روابط میان آنها در آن حوزه شکل می‌گیرد (Brank, Grobelnic, and Mladenic 2005). هستی‌شناسی، توصیفی واضح و رسمی برای تعریف یک پایگاه دانش شامل مفاهیم (یا رده‌ها)^۱ در حوزه موضوعی، نقش‌ها^۲ (یا ویژگی‌ها)^۳ بین نمونه‌ها^۴، مفاهیم، محدودیت‌ها^۵ و قید و شرط‌های مربوط به نقش‌ها، همراه با یک مجموعه از عناصر و اعضا^۶ یا نمونه‌ها) است که پایگاه دانش را تعریف می‌کند. مفاهیم برای بیشتر هستی‌شناسی‌ها، اهمیت ویژه‌ای دارند و نقش‌ها، ویژگی‌های مفاهیم و اعضا را توصیف می‌کنند. هستی‌شناسی‌ها می‌توانند به ایجاد فهم مشترک معنایی از اطلاعات کمک کنند و فرضیات را به طور روشن و واضح توصیف نمایند (Lu 2006).

امروزه، رشته‌های مختلفی با هستی‌شناسی در ارتباط هستند. گروه‌های مختلف ممکن است دیدگاه‌های مختلفی در مورد هستی‌شناسی داشته باشند. در واقع، جامعه مرتبط با هستی‌شناسی یکسان و هماهنگ نیست، اما تاحدودی همه گروه‌ها با وجود دیدگاه‌های مختلف در مورد هستی‌شناسی، بر استفاده از فناوری‌های معنایی در این زمینه، اتفاق نظر دارند. برخی از گروه‌های مرتبط با هستی‌شناسی عبارتند از:

- منطق‌دان‌ها، هستی‌شناسان رسمی، معنی‌شناسان رسمی، و برخی دانشمندان علوم رایانه؛
- کتابداران و دانشمندان اطلاع‌رسانی؛
- مهندسان نرم‌افزار شیء‌گرا؛
- مهندسان دانش هوش مصنوعی کلاسیک؛
- نظریه پردازان و متخصصان پایگاه‌های اطلاعاتی؛
- جامعه وب جهانی؛
- تحلیل‌گران دولتی و تجاری؛ و
- متخصصان حوزه‌های موضوعی (Obrst, Hughes, and Ray 2006).

1. Class	2. role	3. property/ slot
4. instance	5. restriction	6. individual

امروزه، هستی‌شناسی‌های بسیاری در حوزه‌های مختلف موضوعی مانند ژنتیک، زیست‌شناسی، گیاه‌شناسی، و زبان‌شناسی، به‌منظور حمایت از خدمات وب معنایی، به‌وجود آمده‌اند. این هستی‌شناسی‌ها در محیط وب از وب‌سایت‌های بزرگی که به طبقه‌بندی حوزه‌های مختلف دانش می‌پردازند (مانند یاهو^۱) تا وب‌سایت‌هایی که به طبقه‌بندی و ارائه مشخصات محصولات برای فروش پرداخته‌اند (مانند سایت آمازون^۲)، متنوع هستند (Lu 2006). ابزارهای نمایه‌سازی وب مانند یاهو، ساختار سلسله مراتبی عظیمی از رده‌ها و موضوعات فراهم می‌کنند که می‌تواند جنبه‌های مختلف تلاش‌های بشری را نمایش دهد. لایبرو و فینین نیز پس از انجام پژوهشی بر روی یاهو به این نتیجه رسیدند که طبقه‌ها و رده‌های یاهو^۳ و سایر ابزارهای نمایه‌سازی وب، می‌توانند برای توصیف معنایی محتوای مدارک استفاده شوند و به‌همین دلیل مانند هستی‌شناسی قادرند روشی استاندارد و جامع را برای توصیف طبیعت اشیا و فعالیت‌ها و مدارک موجود در جهان واقعی ارائه کنند (Labrou and Finin 1999).

۲. پیشینه بحث

در ایران، در رشته کتابداری و اطلاع‌رسانی تاکنون پژوهشی در زمینه ارزیابی هستی‌شناسی صورت نگرفته است. اما، در خارج از ایران، پژوهش‌های بسیاری در این زمینه انجام شده و معیارها و رویکردها و ابزارهای مختلفی نیز ارائه شده است. در این بخش، به برخی از این پژوهش‌ها پرداخته می‌شود.

گومز - پرز در پژوهشی، مجموعه‌ای از ایده‌های مطرح‌شده در ارزیابی هستی‌شناسی‌ها را ارائه نمود. این ایده‌ها از سویی، برای طراحان هستی‌شناسی در آزمایشگاه‌ها به‌عنوان مبنایی برای انجام ارزیابی‌های فنی سودمند هستند و از سوی دیگر، توسط کاربران نهایی در شرکت‌ها به‌منظور انتخاب بهترین هستی‌شناسی برای استفاده در نظام‌ها، به‌کار می‌روند. در این مقاله، ابتدا برای سؤالاتی در مورد ارزیابی هستی‌شناسی‌ها، پاسخ‌هایی ارائه شده است. این سؤالات به‌اختصار عبارتند از: چرا؟ چه چیزی؟ چه زمانی؟ چگونه؟ کجا؟ و چه کسی ارزیابی را انجام می‌دهد؟ در ادامه، معیارهایی برای ارزیابی هستی‌شناسی‌ها ارائه شده و درنهایت، این معیارها برای ارزیابی هستی‌شناسی "داده‌های کتابشناختی"^۴ مورد استفاده قرار گرفته‌اند (Gómez-Pérez 1995).

میدچ و ستاب نیز در پژوهش خود با در نظر گرفتن دو سطح واژگانی و معنایی برای هستی‌شناسی‌ها، معیارهایی را برای ارزیابی شباهت میان آنها ارائه نمودند. این معیارها، به‌منظور گسترش و یکپارچه‌سازی هستی‌شناسی‌ها، ایجاد سازگاری و هماهنگی میان آنها، و نیز مقایسه

1. Yahoo

2. Amazon.com

3. Yahoo categories

4. Bibliographic-Data

آنها با سایر هستی‌شناسی‌ها مورد استفاده قرار می‌گیرد. آنها معیارهای ارائه‌شده را در یک ارزیابی تجربی در سه مرحله بر روی یک هستی‌شناسی به کار بردند تا کیفیت معیارها را بر روی داده‌های واقعی بررسی کنند. این معیارها در عملکردهای مختلف به کار می‌رود (Maedche and Staab 2002).

لوزانو- تلو و گومز- پرز در پژوهشی به بررسی یکی از شیوه‌های انتخاب هستی‌شناسی به نام ONTOMETRIC پرداختند. این روش به منظور انتخاب مناسب‌ترین هستی‌شناسی برای یک طرح خاص به کار می‌رود. در این پژوهش، از این شیوه برای ارزیابی دو هستی‌شناسی Cisholm's و BWW-ontology استفاده شد تا بهترین هستی‌شناسی به منظور تحلیل نظام‌های تجاری انتخاب شود (Lozano- Tello and Gómez-Pérez 2004).

بروتون- جونز و دیگران در پژوهش خود براساس نظریه نمادشناسی^۱ در رشته زبان‌شناسی، معیارهایی را برای ارزیابی کیفیت هستی‌شناسی ارائه کردند. این نظریه توسط ستامپر و دیگران^۲ در شش سطح به بررسی ویژگی‌های نمادها و نشانه‌ها می‌پردازد. در این پژوهش، فقط از چهار سطح نحوی، معنایی، عملی، و اجتماعی این نظریه استفاده شده است. برای ارزیابی کیفیت هستی‌شناسی‌ها در هر یک از سطوح، معیارهایی پیشنهاد شده است. برای سطح نحوی، معیارهای صحت نحوی و غنا و وسعت هستی‌شناسی؛ برای کیفیت معنایی، معیارهای قابلیت شرح و تفسیر، ثبات و سازگاری، و وضوح و صراحت؛ برای جنبه عمل‌گرا، معیارهای جامعیت و دقت و ربط؛ و برای جنبه اجتماعی، معیارهای استناد و پیشینه هستی‌شناسی ارائه شده است. سپس، این معیارها در نمونه اولیه ابزاری به نام Ontology Auditor به کار گرفته شده و عملکرد آن در ارزیابی تعدادی از هستی‌شناسی‌های مبتنی بر زبان DAML^۳ بررسی شده است (Burton-Jones et al. 2005).

علاوه بر این، پژوهش‌هایی نیز به بررسی ابزارهای ارزیابی هستی‌شناسی پرداخته‌اند. برای نمونه مرشد و سینق نیز در پژوهش خود معیارهایی را برای ارزیابی و رتبه‌بندی ابزارهای طراحی هستی‌شناسی ارائه نمودند. برخی از معیارهای مطرح‌شده در این پژوهش عبارتند از: امکان ویرایش و تغییر مداوم در مفاهیم و روابط، اعتبار و پایایی، میزان پوشش مفاهیم و روابط، میزان دسترس‌پذیری ویرایشگر از طریق وب، کاربرد آسان ویرایشگر، و امکان نمایش گرافیکی هستی‌شناسی. در بخش پایانی مقاله نیز، سه ویرایشگر طراحی هستی‌شناسی (Protégé, OntoEdit Free, Metis) براساس این روش رتبه‌بندی مورد ارزیابی قرار گرفته‌اند (Murshed and Singh 2005).

1. Semiotic theory

2. Stamper's et al

3. DARPA Agent Markup language (DAML)

لو در پایان‌نامه خود به بررسی یکی از ابزارهای ارزیابی هستی‌شناسی به نام ¹ ontoKBEval پرداخت. این ابزار با استفاده از استدلال‌گر RacerPro، هستی‌شناسی‌هایی را که در قالب زبان ² OWL طراحی شده‌اند، براساس منطق توصیف ³ ارزیابی می‌کند (Lu 2006). مقاله حاضر با توجه به پژوهش‌های انجام‌شده در این زمینه و معیارهایی که توسط صاحب‌نظران این حوزه ارائه شده است، برخی از رویکردها و معیارهای ارزیابی هستی‌شناسی را مورد بررسی قرار می‌دهد.

۳. هستی‌شناسی در رشته کتابداری و اطلاع‌رسانی

در حرفه کتابداری و اطلاع‌رسانی، یکی از مهمترین اهداف، سازماندهی علوم به گونه‌ای است که بیشترین واقعیت‌نمایی را از حوزه موضوعی مربوط داشته باشد و درنهایت، سریع‌ترین و مناسب‌ترین بازیابی را در هنگام جستجو متناسب با نیاز کاربر فراهم کند. ساخت ابزارهای معنایی مانند هستی‌شناسی‌ها نیز به منظور تحقق این هدف انجام شده است. از سویی، ظهور فناوری‌های اطلاعاتی و ارتباطی و دسترسی به محیط وب برای تبادل و سازماندهی اطلاعات، سبب تغییر در محیط‌های اطلاعاتی و در نتیجه، تغییر ابزارهای بازنمون دانش شده است. بنابراین، لازم است کتابداران به منظور سازماندهی دانش در عصر جدید، از این ابزارها و فنون جدید در راستای تحقق اهداف خود استفاده نمایند. هستی‌شناسی، ابزاری نوین است که در عصر حاضر با استفاده از فناوری‌های نوین، امکان طراحی و استفاده از آن با تمام قابلیت‌ها امکان‌پذیر شده است.

رشته کتابداری و اطلاع‌رسانی، به دلیل نیاز این رشته به ابزارهایی به منظور ارائه مفاهیم یک حوزه و روابط معنایی دقیق میان آنها، از هستی‌شناسی استفاده می‌کند. هستی‌شناسی، یادآور ابزارهای مورد استفاده در حرفه کتابداری و اطلاع‌رسانی برای طبقه‌بندی مفاهیم است. طرح‌های رده‌بندی، اصطلاحنامه، یا دیگر انواع واژگان کنترل‌شده نظام‌هایی بوده‌اند که از دیرباز در کتابداری و اطلاع‌رسانی استفاده می‌شده‌اند. اما، ظهور هستی‌شناسی‌ها در حرفه کتابداری ناشی از نیاز به یک ابزار معنایی است که روابط مفهومی به مراتب صریح‌تر و تعریف‌شده‌تر از روابط مفهومی اصطلاحنامه‌ها فراهم آورد (Soergel et al. 2004 در صنعت جو ۱۳۸۴).

از آنجا که رشته کتابداری و اطلاع‌رسانی در مقایسه با سایر رشته‌هایی که هستی‌شناسی را مورد استفاده قرار داده‌اند، با کاربران بسیاری سر و کار دارد، باید طراحی این ابزار براساس نیازهای کاربران باشد. به همین دلیل، ارزیابی هستی‌شناسی‌ها اهمیت ویژه‌ای برای کتابداران و اطلاع‌رسانان دارد.

1. Ontology Knowledge Base Evaluation Tool (ontoKBEval)
2. Ontology Web Language (OWL)
3. Description Logic (DL)

۴. ضرورت ارزیابی هستی‌شناسی‌ها

رشد و توسعه فعالیت‌ها و عملکرد مدیریت دانش، وب معنایی و بسیاری از خدمات مربوط به آن، به هستی‌شناسی‌ها، به‌عنوان ستون فقرات این فعالیت‌ها وابسته است (Fensel et al. 2004 cited in Brewster et al. 2003). در محیط وب معنایی، هستی‌شناسی با عنوان مفهوم‌سازی بخش محدودی از واقعیت‌های موجود در حوزه‌ای معین از دانش معرفی می‌شود که در میان جامعه کاربران آن حوزه، مشترک است. این مفهوم‌سازی به زبان رسمی و قابل پردازش برای ماشین و با هدف اشتراک اطلاعات معنایی در نظام‌های خودکار، تعریف شده است. در محیط وب، هستی‌شناسی، ابزار فشرده و منظمی برای تعریف معانی منابع وب است (ژاکوب ۲۰۰۳). هستی‌شناسی‌ها در مدیریت دانش نیز نقش مهمی را ایفا می‌کنند. هستی‌شناسی‌ها، قابلیت یکپارچه‌سازی معنایی و بازنمایی انواع منابع دانش سازمانی را دارند و از بنیانی‌ترین زیرساخت‌ها در رویکردهای پیشرفته خودکارسازی مدیریت دانش هستند. هستی‌شناسی‌ها نه تنها تسهیل‌کننده اشتراک و استفاده مجدد دانش توسط عاملان هوشمند و رایانه‌ها هستند، میان افراد نیز چنین نقشی را ایفا می‌کنند، به‌طوری‌که ارتباطات و امکان استفاده مجدد از دانش را میان موجودیت‌ها - افرادی - که در زمینه موضوعی، دانش‌شان را به اشتراک می‌گذارند، فراهم می‌کند (شریف ۱۳۸۷). بنابراین، ارزیابی هستی‌شناسی‌ها، امری مهم و ضروری برای رشد و توسعه وب معنایی و مدیریت دانش است.

با توجه به کاربرد هستی‌شناسی‌ها در حوزه‌های مختلف، لازم است تا کاربران بتوانند بهترین هستی‌شناسی را که متناسب با نیاز دانشی آنهاست، انتخاب کنند. طراحان هستی‌شناسی باید معیارهایی را در طراحی و ساخت هستی‌شناسی مورد توجه قرار دهند. آنها باید نتایج به‌دست آمده از کاربرد هستی‌شناسی را براساس این معیارها ارزیابی کنند و در مرحله ویرایش و اصلاح هستی‌شناسی نیز این معیارها را به کار گیرند.

از سوی دیگر، آنچه سبب اهمیت رویکردی خاص می‌گردد، قابلیت ارزیابی و مقایسه آن با سایر پدیده‌های مشابه در یک حوزه است. همین مسأله در مورد وب معنایی و هستی‌شناسی‌ها نیز مطرح است. هستی‌شناسی، ساختاری مبتنی بر داده برای مفهومی کردن دانش است، اما ممکن است هستی‌شناسی‌های مختلفی برای مفهومی کردن حوزه‌ای یکسان از دانش وجود داشته باشند. بنابراین، نیاز است که بتوان هستی‌شناسی‌هایی را که بیش از همه، معیارهای تعیین‌شده را در خود دارند، تعیین کرد (Brank, Grobelnic, and Mladenec 2005).

استفاده از معیارها در ارزیابی هستی‌شناسی‌ها می‌تواند توانمندی‌ها و ضعف‌های مربوط به مدل‌های مختلف هستی‌شناسی را آشکار سازد. بهترین هستی‌شناسی آن است که قادر باشد

اهداف و رسالت‌هایی را که به آن منظور طراحی شده است، برآورده سازد. همچنین، باید به بهترین نحو، در رفع نیازهای دانشی کاربران خود سودمند و مؤثر باشد. تاکنون در ایران مبحث ارزیابی هستی‌شناسی در رشته کتابداری و اطلاع‌رسانی بررسی نشده است. به همین دلیل، مقاله حاضر این موضوع را مورد توجه قرار می‌دهد.

۵. دو رویکرد کمی و کیفی در ارزیابی هستی‌شناسی

نوع کیفی ارزیابی هستی‌شناسی‌ها، به‌طور عمده بر قضاوت‌های کاربران یا متخصصان آن متکی است که بر این اساس، طراحان هستی‌شناسی، کاربران نظام یا متخصصان حوزه، باید هستی‌شناسی را ارزیابی کنند (Porzel and Malaka 2004). یکی از مسائلی که این رویکرد با آن روبرو است، انتخاب مجموعه‌ای از معیارها و اصول برای ارزیابی هستی‌شناسی است.

در ارزیابی کیفی، کل هستی‌شناسی یا بخش‌هایی از آن در دسترس کاربران قرار می‌گیرد و از آنها درخواست می‌شود تا آن را ارزیابی کنند. یکی از مشکلات این رویکرد، انتخاب کاربران مناسب و معیارهای مناسب برای ارزیابی است. آیا متخصصان حوزه باید به‌عنوان کاربران در نظر گرفته شوند، یا مهندسان دانش یا حتی کاربران نهایی؟ آیا کاربران باید هستی‌شناسی را با معیارهای تخصصی مانند جامعیت، صحت و ... ارزیابی کنند و منظور ما از این واژه‌ها چیست؟ علاوه بر این بیشتر کاربران نمی‌توانند درستی و صحت یک هستی‌شناسی را به‌لحاظ منطقی ارزیابی کنند (Brewster et al. 2004). رویکرد معیارمحور که در بخش‌های بعدی مقاله به توضیح آن پرداخته می‌شود، از انواع ارزیابی کیفی است.

در رویکرد کمی، از معیارهای قابل سنجش و فرمول‌های آماری برای ارزیابی هستی‌شناسی استفاده می‌شود. برای مثال، بروستر و دیگران در سال ۲۰۰۴ در پژوهش خود تلاش کردند با استفاده از معیارهایی مانند جامعیت^۱ و مانعیت^۲ (دقت و بازیافت) بر روی مجموعه‌ای از متون که بازنمونی از حوزه موضوعی مورد نظر بودند، میزان تناسب آن هستی‌شناسی با حوزه موضوعی مربوط را بررسی کنند (Brewster et al. 2004). گانگمی و دیگران^۳ نیز در پژوهش خود معیارهایی را که بیشتر بر ارزیابی ساختار، عملکرد و کاربردپذیری هستی‌شناسی متمرکز بود، ارائه نمودند. برخی از این معیارها عبارتند از: تعداد رده‌ها و زیررده‌ها، میانگین نمونه‌های موجود در هر رده، و تراکم رده‌ها و زیررده‌ها (Yu, Thom, and Tam 2007). رویکرد وظیفه‌محور که در بخش‌های بعدی مقاله توضیح داده می‌شود، از انواع ارزیابی کمی است.

1. Recall

2. Precision

3. Gangemi et al

۶. مراحل مختلف ارزیابی هستی‌شناسی

امروزه، روش‌های مختلف برای ارزیابی هستی‌شناسی‌ها وجود دارد. برخی روش‌ها، هستی‌شناسی‌ها را در مرحله قبل از مدل‌سازی به‌منظور جلوگیری از ایجاد ناسازگاری‌ها و ناهماهنگی‌ها در طراحی آنها مورد ارزیابی قرار می‌دهند. برخی روش‌ها، هستی‌شناسی‌ها را در مرحله مدل‌سازی به‌منظور کنترل هماهنگی و بررسی خطاهای منطقی موجود در هستی‌شناسی، ارزیابی می‌کنند. برخی نیز، هستی‌شناسی‌ها را در مرحله پس از طراحی و انتشار، کنترل می‌کنند. روش آخر شامل مخازن هستی‌شناسی‌ها نیز می‌شود؛ روش اول و دوم نیز با طرح‌های موجود در زمینه هستی‌شناسی ارتباط دارد.

۶-۱. ارزیابی پیش از مدل‌سازی هستی‌شناسی

در این روش ارزیابی، منابعی که در مراحل اولیه به‌منظور طراحی و ساخت هستی‌شناسی، در اختیار مهندس هستی‌شناسی قرار می‌گیرد، ارزیابی می‌شود. این منابع ممکن است از طریق مهندسی مجدد صفحات وب و کشف مجموعه‌های متون در حوزه‌های موضوعی به‌دست آمده باشد. پیش از آنکه مهندس هستی‌شناسی از منابع گردآوری‌شده استفاده کند، باید کیفیت این منابع مورد بررسی قرار گیرد.

۶-۲. ارزیابی در روند مدل‌سازی هستی‌شناسی

این روش زمانی استفاده می‌شود که مهندس هستی‌شناسی تصمیم می‌گیرد در فواصل زمانی منظم، به کنترل کیفیت هستی‌شناسی بپردازد. در این مرحله، ممکن است هستی‌شناسی‌های مشابه موجود و در دسترس، به‌عنوان منبعی برای مقایسه، مورد استفاده قرار گیرند. کنترل کیفیت هستی‌شناسی، شامل کنترل هماهنگی و سازگاری هستی‌شناسی و کنترل خطاهای منطقی و خطاهای مربوط به زبان هستی‌شناسی نیز هست. در این مرحله، از مجموعه پرسش‌هایی برای سنجش قابلیت و کارآمدی هستی‌شناسی نیز استفاده می‌شود که می‌تواند روشی برای ارزیابی محتوای هستی‌شناسی باشد (این مسأله بررسی می‌شود که آیا مدل هستی‌شناسی ارائه‌شده برای آن حوزه موضوعی، مفاهیم ضروری را برای پاسخ به این پرسش‌ها در بردارد یا خیر؟).

۶-۳. ارزیابی پس از انتشار و ارائه هستی‌شناسی

ارزیابی هستی‌شناسی پس از انتشار، به‌طور عمده توسط کاربران انجام می‌شود. این ارزیابی

شامل بررسی عملکرد هستی‌شناسی طراحی شده و نیز مقایسه هستی‌شناسی‌های مختلف در آن حوزه موضوعی است. میزان استفاده کاربران از مفاهیم یک هستی‌شناسی، می‌تواند بیانگر این باشد که تا چه حد هستی‌شناسی، به بازنمون دانش حوزه پرداخته است. اما، هدف عمده این نوع ارزیابی، تصمیم‌گیری در مورد انتخاب هستی‌شناسی است و شامل مقایسه و انطباق ویژگی‌های هستی‌شناسی با یک مجموعه معیارهای کمی و کیفی است (Hartmann et al. 2005).

۷. رویکردهای مختلف ارزیابی هستی‌شناسی

براساس اینکه چه نوع هستی‌شناسی و با چه هدفی مورد ارزیابی قرار می‌گیرد، می‌توان رویکردهای مختلفی برای ارزیابی هستی‌شناسی‌ها در نظر گرفت. اما به هر حال، رویکردهای ارزیابی هستی‌شناسی به‌طور عمده در یکی از موارد زیر می‌گنجد:

- رویکرد استاندارد طلایی^۱: آن دسته از ارزیابی‌ها که به مقایسه هستی‌شناسی با یک استاندارد اصلی می‌پردازند که این استاندارد ممکن است خود یک هستی‌شناسی دیگر باشد (مانند پژوهشی که میدچ و ستاب^۲ سال ۲۰۰۲ انجام دادند).
- رویکرد کاربرد/وظیفه‌محور^۳: آن دسته از ارزیابی‌ها که به بررسی هستی‌شناسی در بافت وظیفه و کاربردی خاص می‌پردازند و نتایج به‌دست‌آمده را مورد ارزیابی قرار می‌دهند (مانند پژوهشی که پورزل و مالاکا^۴ سال ۲۰۰۴ انجام دادند).
- رویکرد استخراج داده^۵ (براساس متون و مدارک حوزه‌ای خاص): آن دسته از ارزیابی‌ها که به مقایسه هستی‌شناسی با یک منبع داده (مانند مجموعه‌ای از مدارک) در آن حوزه خاص می‌پردازند تا بررسی کنند که هستی‌شناسی مورد نظر تا چه حد مدارک آن حوزه را تحت پوشش قرار داده است (مانند پژوهشی که بروستر و دیگران^۶ در سال ۲۰۰۴ انجام دادند).
- رویکرد معیارمحور^۷ (ارزیابی توسط انسان^۸): آن دسته از ارزیابی‌ها که توسط عوامل انسانی انجام می‌شوند و آنها تلاش می‌کنند تا تعیین کنند هستی‌شناسی تا چه میزان، مجموعه‌ای از معیارها و استانداردها و عناصر مهم را تحت پوشش قرار داده است (مانند پژوهشی که لوزانوتلو و گومز-پرز^۹ سال ۲۰۰۴ انجام دادند) (Brank, Grobelnic, and Mladenic 2005).

1. Golden standard Approach

3. Task/Application- based Approach

5. Data- driven Approach

7. Criteria-based Approach

9. Lozano-Tello and Gomez-Perez

2. Maedche and Staab

4. Porzel and Malaka

6. Brewster et al.

8. Assessment by humans

۱-۷. رویکرد استاندارد طلایی

این رویکرد، یک هستی‌شناسی را با هستی‌شناسی دیگری که می‌تواند به‌عنوان هستی‌شناسی معیار در نظر گرفته شود، مقایسه می‌کند. به‌طور معمول، این نوع ارزیابی برای تعیین اینکه آیا فرآیند تولید یک هستی‌شناسی (به‌صورت نیمه‌خودکار یا براساس یک الگوریتم یادگیری) به‌صورت مناسب و کارآمد انجام شده است یا خیر، انجام می‌گیرد. میدچ و ستاب، در پژوهش خود، نمونه‌ای از ارزیابی هستی‌شناسی را براساس استاندارد طلایی مطرح کرده‌اند. آنها روش‌های تجربی و آزمایشی را برای ارزیابی شباهت‌های بین هستی‌شناسی‌ها هم از لحاظ واژگانی و هم از لحاظ مفهومی ارائه کرده‌اند. معیارها براساس همپوشانی در روابط تعیین شده‌اند (معیارهای ارزیابی دقت و صحت یادگیری روابط عمومی^۱). این معیارها دقت و صحت روابطی را که از طریق نظام یادگیری هستی‌شناسی پیشنهادی آنها ایجاد شده است در مقایسه با یک هستی‌شناسی موجود، تعیین می‌کنند (Yu, Thom, and Tam 2007).

۲-۷. رویکرد کاربرد/وظیفه‌محور

این رویکرد، رویکردی کمی در ارزیابی هستی‌شناسی‌ها به‌شمار می‌آید. در این رویکرد، چندین هستی‌شناسی مربوط به یک حوزه موضوعی در نظر گرفته می‌شود و عملکرد آنها در انجام وظایف و فعالیت‌های یکسان، مورد ارزیابی قرار می‌گیرد تا مناسب‌ترین هستی‌شناسی برای عملکرد مربوط انتخاب گردد. یکی از مراحل مهم پژوهش‌هایی که از این رویکرد در ارزیابی هستی‌شناسی‌ها استفاده کرده‌اند، تعیین مجموعه‌ای از فعالیت‌ها و عملکردهای یکسان است که امکان می‌دهد تا هستی‌شناسی‌های مختلف از نظر نحوه انجام آن عملکردها، مورد ارزیابی قرار گیرند.

این رویکرد، هستی‌شناسی را براساس قابلیت‌ها و توانایی‌های آن در تکمیل وظایف و عملکردها ارزیابی می‌کند. با استفاده از این رویکرد می‌توان به‌روشی کمی در مورد تناسب هستی‌شناسی برای انجام فعالیت یا وظیفه‌ای خاص، قضاوت کرد و این کار از طریق بررسی و ارزیابی عملکرد آن هستی‌شناسی در بافت آن فعالیت صورت می‌گیرد. اشکال و ضعف این رویکرد این است که ارزیابی فعالیت یا وظیفه‌ای خاص ممکن است قابل تعمیم به وظیفه و فعالیت دیگر نباشد. به‌همین دلیل، لازم است ارزیابی‌ها برای هر وظیفه و فعالیت به‌صورت مستقل انجام شوند (Yu, Thom, and Tam 2007).

یک هستی‌شناسی ممکن است در انواع فعالیت‌ها و وظایف مورد استفاده قرار گیرد. نتیجه

آن فعالیت یا عملکرد خاص ممکن است تا حدی به هستی‌شناسی انتخاب‌شده در آن فعالیت، وابسته باشد. در نتیجه می‌توان گفت یک هستی‌شناسی خوب و باکیفیت آن است که برای عملکرد و فعالیتی خاص، نتایج و پاسخ‌های مناسب و مؤثری ارائه کند. بنابراین، هستی‌شناسی‌ها ممکن است از طریق استفاده و به‌کارگیری آنها در عملکردی خاص و ارزیابی نتایج آن عملکرد، مورد ارزیابی قرار گیرند (Brank, Grobelnic, and Mladenec 2005).

برای ارزیابی و اندازه‌گیری میزان تأثیر هستی‌شناسی در عملکرد مربوط (که نتیجه آن به مدل هستی‌شناسی بستگی دارد)، باید فعالیت‌ها و عملکردها در طول ارزیابی، به‌صورت مداوم و ثابت انجام شوند و نحوه ارزیابی باید به‌گونه‌ای انتخاب شود که نتیجه انجام آن وظیفه، تا حد ممکن به هستی‌شناسی انتخاب و استفاده‌شده، وابسته باشد.

در این رویکرد، هستی‌شناسی، دست‌کم در سه سطح ارزیابی می‌شود:

۱. دامنه‌ها (رده‌ها و مفاهیم هستی‌شناسی)؛
۲. تناسب روابط مربوط به طبقه‌بندی مفاهیم (روابط سلسله‌مراتبی is-a)؛ و
۳. تناسب روابط معنایی میان مفاهیم (Porzel and Malaka 2004).

چهار عنصر مهم در این رویکرد عبارتند از:

۱. وظیفه^۱: لازم است وظیفه‌ای که در این رویکرد انتخاب می‌شود به اندازه کافی جامع و گسترده باشد تا شرایط مناسب را برای بررسی و ارزیابی آن هستی‌شناسی مشخص، فراهم کند، به‌خصوص اگر هدف ارزیابی این است که روابط معنایی را نیز شامل شود. در این صورت لازم است وظایفی در نظر گرفته شود که نتیجه حاصل از انجام آنها، به‌طور عمده، به نحوه مدل‌سازی این روابط معنایی در هستی‌شناسی، وابسته باشد.
۲. یک (یا تعداد بیشتری) هستی‌شناسی: دست‌کم یک هستی‌شناسی برای اجرای این رویکرد ارزیابی، نیاز است. می‌توان فقط یک هستی‌شناسی را براساس عملکرد آن در انجام یک وظیفه مشخص، ارزیابی کرد. همچنین، می‌توان هستی‌شناسی‌های مختلف را از نظر انجام آن وظیفه، با یکدیگر مقایسه نمود.
۳. کاربرد^۲: عملکرد، الگوریتم و فرآیندی خاص است که هستی‌شناسی برای انجام وظیفه‌ای مشخص، از آن استفاده می‌کند. بزرگ‌ترین مشکل در این رویکرد، تعیین و اجرای الگوریتم‌ها و حفظ آنها در طول ارزیابی است.
۴. استاندارد طلایی: برای ارزیابی عملکرد هر الگوریتمی که عناصر اصلی مانند نقش‌ها و تگ‌های کلمات (اسم، صفت، مفعول و ...)، معانی و مفاهیم کلمات، و تگ‌های روابط

1. task

2. application

موجود در هستی‌شناسی را ایجاد می‌کند، مجموعه‌ای معین از پاسخ‌ها نیاز است. در این رویکرد، این مجموعه تفسیرشده و کامل از پاسخ‌ها، "استاندارد طلایی" نامیده می‌شود. رویکرد وظیفه‌محور در ارزیابی هستی‌شناسی‌ها، دارای چندین اشکال است. برخی از این اشکالات عبارتند از:

۱. زمانی که هستی‌شناسی به‌روش خاصی برای عملکردی خاص استفاده می‌شود، می‌توان مناسب یا نامناسب بودن آن را تعیین نمود، اما این مشاهده و بررسی را به‌سختی می‌توان تعمیم داد.
۲. یک هستی‌شناسی ممکن است فقط بخش کوچکی از یک وظیفه یا فعالیت باشد، بنابراین تأثیر آن هستی‌شناسی بر نتیجه آن فعالیت ممکن است غیرمستقیم و به‌نسبت کم باشد.
۳. مقایسه هستی‌شناسی‌های مختلف فقط زمانی امکان دارد که همه آنها بتوانند در همان فعالیت خاص به کار گرفته شوند (Brank, Grobelnic, and Mladenic 2005).

۳-۷. رویکرد استخراج داده (براساس متون و مدارک حوزه‌ای خاص)

هستی‌شناسی ممکن است از طریق مقایسه آن با داده‌های حقیقی (به‌طور معمول، یک مجموعه مدارک متنی) موجود در حوزه تحت پوشش هستی‌شناسی، ارزیابی گردد (Brank, Grobelnic, and Mladenic 2005).

در این رویکرد، از طریق یک روش احتمالاتی، مشخص می‌شود که تا چه حد هستی‌شناسی برای بازنمون دانش حوزه مربوط، مناسب است.

بروستر و دیگران در پژوهش خود، از رویکرد استخراج داده استفاده نمودند. به کارگیری این رویکرد، آنها به معیارهایی برای تناسب هستی‌شناسی^۱ با مجموعه‌ای از متون مربوط به حوزه موضوعی مورد نظر دست یافتند. معیار آنها، همپوشانی واژگانی بین مفاهیم یک هستی‌شناسی معین و واژه‌های استخراج‌شده (توسط الگوریتم خوشه‌بندی معنایی پنهان) و گسترش یافته توسط نرم‌افزار WordNet بود. برای ارزیابی تناسب طبقه‌بندی مفاهیم، نویسندگان روش تنیس متریک^۲ را که توسط استونسون^۳ پیشنهاد شده بود، به کار بردند. روش پژوهش به این صورت بود که آنها در ابتدا، حوزه هنر و هنرمندان و پنج هستی‌شناسی را که در زمینه هنر بود، در نظر گرفتند. سپس، ۴۱ متن را در مورد تعدادی از هنرمندان از اینترنت انتخاب کردند. در مرحله بعد، به‌جای مقایسه هر هستی‌شناسی با هستی‌شناسی دیگر، هستی‌شناسی‌ها را با مجموعه متون استخراج‌شده در حوزه هنر مقایسه نمودند. به این منظور، تعدادی از واژه‌های مجموعه متون مورد نظر، به‌صورت خودکار استخراج گردید؛ تعداد واژه‌هایی که بین هر

1. Ontological fit

2. Tennis Metric

3. Stevenson

هستی‌شناسی و آن مجموعه متون همپوشانی داشتند، محاسبه شد و واژه‌هایی که در هر یک از این دو (هستی‌شناسی و مجموعه متون) وجود نداشتند، حذف گردید. از طریق مقایسه همپوشانی‌ها، میزان تناسب هر هستی‌شناسی با مجموعه متون مورد نظر، بررسی شد و هستی‌شناسی مناسب (هستی‌شناسی ARTEQUAKT) انتخاب گردید (Brewster et al. 2004).

روش دیگر برای این مقایسه، استفاده از نمایش فضای برداری از واژه‌هایی است که هم در هستی‌شناسی تحت ارزیابی قرار می‌گیرند و هم در مجموعه متون، موجود هستند. استفاده از این روش، معیاری کلی برای سنجش میزان تناسب هستی‌شناسی‌ها و مجموعه متون مورد نظر، ایجاد می‌کند (Brewster et al. 2004).

در این رویکرد، برای ایجاد تناسب میان هستی‌شناسی و مجموعه متون، سه مرحله وجود دارد و هر مرحله می‌تواند به روش‌های مختلف انجام شود:

۱. تعیین کلیدواژه‌ها: در این مرحله به‌طور عمده، شناسایی خودکار واژه‌ها صورت می‌گیرد و همه روش‌های موجود می‌توانند به این منظور به کار گرفته شوند (Magnard, 2004 and Ananiadou 2000 cited in Brewster et al. 2004). در پژوهش بروستر و دیگران، روش تحلیل معنایی پنهان^۱ و روش خوشه‌بندی استفاده شده است (Brewster et al. 2004).

۲. گسترش پرسش: یک مفهوم در هستی‌شناسی، بازنمون مجموعه‌ای از واژه‌های مختلف به روش‌های مختلف است. به همین دلیل، لازم است که چندین شکل از واژه‌های مفهومی، از طریق گسترش پرسش ایجاد شود. در پژوهش اشاره‌شده، از WordNet به منظور گسترش پرسش در هر خوشه استفاده شده است.

۳. نگاهت هستی‌شناسی: سرانجام مجموعه واژه‌هایی که در خوشه مشخص شده‌اند، باید در طراحی هستی‌شناسی مورد استفاده قرار گیرند.

استفاده از رویکرد تحلیل خوشه‌ای، معیاری برای تناسب ساختاری هستی‌شناسی ایجاد می‌کند. می‌توان دو هستی‌شناسی را با مجموعه‌های مفهومی یکسان، اما نحوه سازماندهی متفاوت، تصور کرد. در این صورت، مفاهیم در این دو هستی‌شناسی دارای فاصله متفاوت از یکدیگر هستند. با استفاده از روش Tennis Measure (طراحی‌شده توسط استونسون^۲) می‌توان میزان نزدیکی عناصر در خوشه‌ای از هستی‌شناسی را نسبت به عناصر موجود در خوشه‌های متفاوت، تعیین کرد. نتیجه محاسبات، به مدل احتمالاتی که برای استخراج خوشه‌ها استفاده شده است، بستگی دارد (Brewster et al. 2004).

1. Latent Semantic Analysis

2. Stevenson

۷-۴. رویکرد معیارمحور

در این رویکرد، هستی‌شناسی‌ها به صورت کیفی، براساس معیارهای تعیین شده توسط متخصصان این حوزه، به منظور تصمیم‌گیری برای انتخاب مناسب‌ترین هستی‌شناسی، مورد ارزیابی قرار می‌گیرند. هر هستی‌شناسی از نظر دارا بودن معیارها، ارزیابی می‌شود و به‌ازای هر معیار، نمره و رتبه‌ای را به خود اختصاص می‌دهد و جمع نمرات برای هر هستی‌شناسی محاسبه می‌گردد. مسأله اصلی در این رویکرد این است که چگونه می‌توان هستی‌شناسی‌ها را با توجه به معیارهای خاص، ارزیابی نمود (Brank, Grobelnic, and Mladenic 2005).

در حالی که برخی از این اصول و معیارها از لحاظ نظری با ارزش هستند، بسیار دشوار است که آزمون‌های خودکار، برای مقایسه و ارزیابی دو یا تعداد بیشتری از هستی‌شناسی‌ها، براساس استفاده مداوم آنها از معیارهای اصلی، طراحی نمود. زیرا این معیارها و اصول برای انجام ارزیابی، به یک معنی‌شناسی ظاهری وابسته هستند که در حال حاضر فقط بشر قادر به پیش‌بینی و تهیه آنهاست. به همین دلیل، فقط بشر قادر است که این نوع ارزیابی را به صورت کامل انجام دهد (Brewster et al. 2004).

در ادامه، به برخی از معیارهای ارزیابی هستی‌شناسی که توسط صاحب‌نظران این حوزه و نیز مؤسسه تحقیقات هستی‌شناسی امریکا^۱ مطرح شده است، اشاره می‌شود.

۷-۴-۱. معیارهای ارزیابی هستی‌شناسی براساس نظرات صاحب‌نظران این حوزه

- بورتون جونز و دیگران، ۱۰ معیار را برای ارزیابی هستی‌شناسی‌ها پیشنهاد کرده‌اند:
- قانونی و مجاز بودن هستی‌شناسی^۲ (تعداد خطاهای نحوی)؛
- غنی بودن و وسعت هستی‌شناسی^۳: چه تعداد از ویژگی‌های نحوی که در زبان رسمی در دسترس هستند، توسط هستی‌شناسی استفاده شده‌اند؟
- قابلیت شرح و تفسیر^۴: آیا واژه‌های استفاده شده در هستی‌شناسی در WordNet هم ظاهر شده‌اند؟
- ثبات و سازگاری^۵: چه تعداد از مفاهیم هستی‌شناسی سبب ایجاد ناهماهنگی و بی‌ثباتی در آن شده‌اند؟
- وضوح و صراحت^۶: آیا واژه‌های استفاده شده در هستی‌شناسی دارای معانی متعددی در WordNet هستند؟

1. US National Center for Ontological Research (NCOR <http://ncor.us>)
3. richness 4. interpretability 5. consistency

2. Lawfulness
6. clarity

- جامعیت و وسعت^۱: تعداد مفاهیم هستی‌شناسی نسبت به میانگین تمام مخزن هستی‌شناسی‌ها؛
 - دقت^۲: درصد شرح و توضیحات اشتباه در هستی‌شناسی؛
 - ربط^۳: چه تعداد از توضیحاتی که ویژگی‌های نحوی را شامل می‌شوند، به‌عنوان موارد مفید و قابل پذیرش توسط کاربر، ارائه شده‌اند؛
 - استناد^۴: تا چه حد سایر هستی‌شناسی‌ها از مفاهیم موجود در این هستی‌شناسی استفاده کرده‌اند؟ و
 - سابقه و پیشینه^۵: در مقایسه با سایر هستی‌شناسی‌ها در مخزن هستی‌شناسی، چه تعداد دسترسی به این هستی‌شناسی ایجاد شده است (Burton-Jones et al. 2004)؟
- علاوه بر بورتون، افراد دیگری نیز معیارهایی را پیشنهاد کرده‌اند. گروبر معیارهای زیر را ارائه نموده است:
- وضوح و صراحت؛
 - همبستگی و ارتباط میان مفاهیم^۶؛
 - توسعه‌پذیری^۷؛
 - کمترین خطا در هستی‌شناسی^۸؛ و
 - کمترین تعصب، تبعیض، و سوگیری در کدگذاری^۹ (Gruber 1993 cited in Yu, Thom, and Tam 2007).
- گرونینگر و فوکس نیز کارآمدی و صلاحیت^{۱۰} هستی‌شناسی را به‌عنوان معیار ارزیابی هستی‌شناسی بیان کرده‌اند (Gruninger and Fox 1995 cited in Yu, Thom, and Tam 2007).
- گومز - پرز نیز معیارهای زیر را پیشنهاد کرده است:
- سازگاری، ثبات، هماهنگی، و انسجام؛
 - جامعیت^{۱۱}؛
 - اختصار^{۱۲}؛
 - قابلیت توسعه و گسترش‌پذیری^{۱۳}؛ و
 - حساسیت و انعطاف‌پذیری^{۱۴} (Gómez-Pérez 1995).

1. comprehensiveness
4. authority
7. extendibility
9. minimal encoding bias
12. conciseness

2. accuracy
5. history
8. minimal Ontological commitment
10. Competency
13. expandability

3. Relevance
6. Coherence
11. Completeness
14. sensitiveness

گورینو و ولتی، صحت و درستی^۱ موجودیت‌ها و روابط هستی‌شناسی را به‌عنوان معیار ارزیابی آن، مطرح کرده‌اند. آنان صحت و درستی چهار عنصر مهم در هستی‌شناسی را مورد ارزیابی قرار داده‌اند. این چهار عنصر عبارتند از:

- هستی و ماهیت^۲؛
 - ثبات و پیوستگی^۳؛
 - هویت و عینیت^۴؛ و
 - یکپارچگی و وحدت^۵ (Guarino and Welty 2002 cited in Yu, Thom, and Tam 2007).
- لوزانو- تلو و گومز- پرز نیز مجموعه‌ای مفصل از ۱۱۷ معیار را که در سه سطح سازماندهی شده‌اند، پیشنهاد نموده‌اند (Lozano-Tello and Gómez-Pérez 2004).
- براساس تحلیل‌ها برخی معیارها که توسط پژوهشگران مختلف پیشنهاد شده‌اند، جنبه‌های مشابهی از ارزیابی هستی‌شناسی‌ها را مشخص کرده‌اند و با یکدیگر همپوشانی دارند. این معیارها به‌طور کلی عبارتند از:

- وضوح و صراحت؛
- ثبات و سازگاری و دوام؛
- اختصار؛
- گسترش‌پذیری؛
- صحت و درستی؛
- جامعیت؛
- کمترین خطای هستی‌شناسی؛ و
- کمترین تعصب و سوگیری در کدگذاری.

اکنون به تحلیل برخی از این معیارها پرداخته می‌شود:

برخی از این معیارها می‌توانند توسط ابزارهای هستی‌شناسی، ارزیابی شوند. استدلال‌گرهایی مانند FaCT و Racer ابزارهایی را فراهم می‌کنند که قادرند خطاهای موجود در هستی‌شناسی مانند واژه‌های حشو و زائد، تناقض و ناهماهنگی بین تعاریف و مفاهیم و تعاریف گمشده را کنترل و بررسی کنند. علاوه بر این، دونگ و دیگران فنون و ابزارهای موجود در زمینه مهندسی نرم‌افزار را برای کنترل خطاهای هستی‌شناسی‌های حوزه نظامی به‌کار برده‌اند (Dong et al. 2004 cited in Yu, Thom, and Tam 2007). برخی معیارها مانند وضوح و گسترش‌پذیری هستی‌شناسی را به‌سختی می‌توان ارزیابی کرد، زیرا ابزاری برای اندازه‌گیری و

1. Correctness
4. identity

2. essence
5. unity

3. Rigidity

تعیین آنها وجود ندارد. جامعیت هستی‌شناسی را نیز نمی‌توان به‌سادگی آزمود. سایر معیارها می‌توانند چالش‌های بیشتری را در ارزیابی به‌وجود آورند، زیرا ممکن است به‌صورت کمی، قابل اندازه‌گیری نباشند. برای سنجش این معیارها، مشاهده و بررسی مستقیم هستی‌شناسی نیاز است. برای مثال، به‌منظور تعیین "صحت و درستی" هستی‌شناسی، لازم است یک متخصص موضوعی یا یک طراح هستی‌شناسی، به‌صورت شخصی و مستقیم صحت تعاریف را با مراجعه به دنیای واقعی مشخص کند. البته این راه‌حل، ممکن است برای یک هستی‌شناسی بزرگ یا مخزن هستی‌شناسی به‌راحتی امکان‌پذیر نباشد (Yu, Thom, and Tam 2007).

۲-۴-۷. معیارهای ارزیابی هستی‌شناسی از نظر مرکز ملی تحقیقات هستی‌شناسی امریکا

این مرکز در اکتبر ۲۰۰۵ افتتاح شد. هدف اولیه مرکز این است که به‌منظور افزایش کیفیت هستی‌شناسی‌ها، مبحث مهندسی هستی‌شناسی را به رشته‌ای علمی و مهندسی، تبدیل کند.

برای رسیدن به این هدف، مرکز اهداف جانبی زیر را در نظر دارد:

- پیشرفت و توسعه کیفیت زبان‌های بازنمون معنایی دانش، محتوای هستی‌شناسی و روش‌های منطقی و استدلالی؛
- ایجاد روش‌ها و راهکارهایی برای کمک به تعریف و ارزیابی و اطمینان از کیفیت دانش؛
- تسهیل آموزش گروه‌های مرتبط با طراحی هستی‌شناسی و ارتقای مهارت‌ها برای تولید و رشد هستی‌شناسی؛ و
- ایجاد بهترین استانداردهای مرتبط با هستی‌شناسی و تسهیل ارتباط بین مؤسسات استاندارد.

این مرکز سعی دارد که برخی سؤالات در مورد کیفیت هستی‌شناسی‌ها را پاسخ دهد. این سؤالات عبارتند از:

۱. کدام زبان‌ها برای بازنمون دانش، مناسب هستند؟ (هستی‌شناسی‌ها، پایگاه‌های دانش، قوانین، مدارک و اسناد، و پرسش‌ها) و ویژگی‌های مطلوب این زبان‌ها چیست؟
۲. کدام زبان‌ها برای ارائه منطقی و استدلال حاکم بر دانش، مناسب هستند؟
۳. اصول روش‌شناسی و معیارهای طراحی برای رشد و ایجاد محتوای دانش، چه باید باشد؟
۴. آیا پرسش‌های سنجش قابلیت و توانایی هستی‌شناسی، موارد کاربرد هستی‌شناسی و طرح‌های تولید هستی‌شناسی، می‌توانند برای ارزیابی هستی‌شناسی به‌کار روند؟

۵. چگونه از کیفیت محتوای دانش، اطمینان حاصل کنیم؟ آیا باید ویژگی‌های رسمی محتوی (مانند صحت، جامعیت، هماهنگی، ثبات و دوام، دقت، مانعیت، و تابعیت) را به‌عنوان روشی اصولی برای ارزیابی کیفیت محتوی، بپذیریم؟
۶. چگونه محتوای دانش هستی‌شناسی‌های مختلف در حوزه‌ای خاص، با یکدیگر پیوند می‌خورد و یکپارچه می‌گردد؟ آیا یک هستی‌شناسی در حوزه‌ای خاص، با هستی‌شناسی‌های بالایی و میانی و مرجع در آن حوزه ارتباط دارد؟ مرکز ملی تحقیقات هستی‌شناسی امریکا سه روش را برای ارزیابی هستی‌شناسی پیشنهاد می‌کند:

۷-۴-۱. ایجاد رقابت در ارزیابی هستی‌شناسی

این مسأله شامل ایجاد رقابتی در ارزیابی هستی‌شناسی در راستای اهداف کنفرانس ارزیابی بازیابی متون^۱ از مؤسسه ملی استانداردها و فناوری‌های امریکا^۲ است. این مؤسسه همراه با لوکهد مارتن^۳، کنفرانس تفسیر و یکپارچه‌سازی اطلاعات^۴ را نیز برگزار کردند که بر ارزیابی ابزارهای طراحی هستی‌شناسی تمرکز داشت. این موضوع در کارگاه ارزیابی ابزارهای مرتبط با هستی‌شناسی^۵ در ۲۰۰۴ نیز مطرح شد. رقابت در ارزیابی هستی‌شناسی باید ادامه یابد تا هستی‌شناسی‌هایی را که در قالب زبان‌هایی به‌جز زبان‌های بازنمون دانش وب معنایی ارائه شده‌اند نیز در برگیرد. برخی از این زبان‌ها عبارتند از: منطق رایج استاندارد ISO^۶، زبان دانش عملی^۷ (مربوط به طرح بازنمون دانش عملی برای حمایت هوشمند^۸) و حتی آن دسته از زبان‌های بازنمون دانش مانند CycL که باید به یکی از زبان‌های استاندارد عمومی ترجمه شوند.

سؤالات اصلی که باید در طراحی چنین رقابتی پرسیده شوند، عبارتند از:

۱. کدام فعالیت‌ها و چالش‌ها سبب می‌شوند که محتوای هستی‌شناسی، روش‌های منطقی و استدلالی آن، و ایجاد ابزارهای هستی‌شناسی، بهتر بتوانند مسائل دنیای واقعی را حل کنند و مهارت‌ها و فعالیت‌های مربوط به آن را توسعه بدهند؟
۲. منابع مهم و اصلی برای چنین رقابت ارزیابی هستی‌شناسی از کجا تأمین می‌شود؟
۳. ساختار سازمانی چنین رقابت ارزیابی چگونه باید باشد (مانند ساختار دانشگاهی، تجاری، دولتی، با پشتیبانی کنسرسیوم استانداردهای بین‌المللی و ...)?

1. Text Retrieval Evaluation Conference (TREC)
 2. National Institute of Standards and Technology's (NIST)
 3. Lockheed Martin
 4. Information Interpretation and Integration Conference (I3CON)
 5. Evaluation of Ontology-based Tools Workshop (EON)
 6. Standard Common Logic (SCL)
 7. Interoperable Knowledge Language (IKL)
 8. Interoperable Knowledge Representation for Intelligence Support (IKRIS)

۴. چه معیارهایی باید هنگام ارزیابی هستی‌شناسی در نظر گرفته شود؟ آیا چنین معیارهایی به وضوح، ارزش یک هستی‌شناسی خوب و مناسب را نشان می‌دهند؟

۷-۴-۲. تأیید اصولی هستی‌شناسی توسط یک سازمان یا جامعه بررسی‌کننده

مرحله دیگری که باید انجام شود این است که فرآیند تصدیق و تأیید اصولی برای هستی‌شناسی انجام شود. در این رویکرد، سازمان یا انجمنی خاص (متمرکز یا توزیع‌شده)، باید هستی‌شناسی را با روشی اصولی و برطبق مجموعه‌ای از معیارهای ارزیابی تعیین شده مورد بررسی و بازبینی قرار دهد و در نهایت، نمره و رتبه‌ای برای کیفیت هستی‌شناسی تعیین کند. برخی از معیارهایی که در اینجا مورد بررسی قرار می‌گیرند عبارتند از: حوزه موضوعی مربوط، وسعت کاربرد یا میزان پوشش آن حوزه، میزان عمق رده‌بندی و تراکم نسبی گره‌ها، جامعیت خصوصیات، و ...

روش خاص ایجاد هستی‌شناسی نیز می‌تواند در این رویکرد، تشریح شود و مورد بررسی قرار گیرد.

۷-۴-۳. ایجاد مدل تکامل هستی‌شناسی

روش دیگر ارزیابی هستی‌شناسی، ایجاد مدل تکامل و پیشرفت هستی‌شناسی است (از اهداف طرح یکپارچه‌سازی جامع‌ترین مدل تکاملی^۱ توسط مؤسسه مهندسی نرم‌افزار^۲). در این روش، مراحل رشد و توسعه هستی‌شناسی و روش‌های تکامل تدریجی آن در آینده، هم برای محتوی و هم برای ابزارهای مربوط به هستی‌شناسی، در نظر گرفته می‌شود. چنین مدلی می‌تواند سطوح مختلف تکامل را که توسط ویژگی‌هایی تعریف می‌شوند، ترسیم کند. برخی از این ویژگی‌ها عبارتند از: میزان ارجاعات به منابع موضوعی مستند، سودمندی، میزان استفاده‌پذیری، حمایت ابزاری (از طریق ابزارهای تولید هستی‌شناسی)، حمایت از ابزارهای منطقی و استدلالی و ... (Obrst, Hughes, and Ray, 2006).

۸. سطوح ارزیابی هستی‌شناسی

هستی‌شناسی‌ها را می‌توان براساس سطوح ارزیابی نیز طبقه‌بندی نمود. هستی‌شناسی ساختار به نسبت پیچیده‌ای دارد و مناسب‌تر است که به جای ارزیابی هستی‌شناسی به طور کلی، بر ارزیابی سطوح مختلف آن به صورت جداگانه تمرکز کرد. دلیل دیگر برای انتخاب رویکرد مبتنی بر سطوح هستی‌شناسی این است که فنون خودکار یادگیری که در ساخت هستی‌شناسی‌ها استفاده شده است، برای سطوح مختلف هستی‌شناسی متفاوت هستند.

1. Capability Maturity Model Integration

2. Software Engineering Institute

- سطوح هستی‌شناسی به‌طور متفاوتی توسط نویسندگان تعریف شده‌اند، اما این تعاریف تا حد زیادی مشابه هستند و به‌طور معمول، سطوح زیر را شامل می‌شوند:
- ارزیابی براساس سطح واژگانی و داده‌ای^۱: ارزیابی در این سطح بر مفاهیم، نمونه‌ها، داده‌های حقیقی^۲ و ... موجود در هستی‌شناسی و واژه‌هایی که برای ارائه این مفاهیم استفاده شده‌اند، متمرکز است. این سطح ارزیابی شامل مقایسه هستی‌شناسی با منابع مختلف داده مرتبط با حوزه مورد نظر (مانند مجموعه متون مرتبط با حوزه) و نیز روش‌هایی همچون معیارهای مشابهت رشته کلمات است.
 - ارزیابی براساس ساختار سلسله مراتبی و رده‌بندی^۳: در هستی‌شناسی، رابطه سلسله مراتبی Is-a یکی از روابط تعریف شده در میان مفاهیم است. اگر چه روابط مختلف دیگری نیز میان مفاهیم وجود دارد، رابطه Is-a، اهمیت ویژه‌ای در هستی‌شناسی‌ها دارد و به‌طور معمول، این نوع رابطه در ارزیابی هستی‌شناسی‌ها مورد توجه قرار می‌گیرد.
 - ارزیابی براساس سایر روابط معنایی: هستی‌شناسی ممکن است روابط دیگری را علاوه بر Is-a شامل شود و ممکن است این روابط به‌صورت جداگانه مورد ارزیابی قرار گیرند. معیارهایی مانند جامعیت (بازیافت) و مانعیت (دقت) در این سطح بررسی می‌شوند.
 - ارزیابی براساس سطح بافت و کاربرد هستی‌شناسی^۴: یک هستی‌شناسی ممکن است بخشی از یک مجموعه هستی‌شناسی بزرگ باشد. ممکن است مفاهیم و تعاریفی را به هستی‌شناسی‌های دیگر ارجاع بدهد یا توسط سایر هستی‌شناسی‌ها مورد ارجاع و استفاده قرار گیرد. در هنگام ارزیابی هستی‌شناسی لازم است این زمینه و ارتباط با سایر هستی‌شناسی‌ها در نظر گرفته شود. مسأله دیگر در زمینه ارزیابی، کاربرد و هدف خاصی است که هستی‌شناسی برای آن ایجاد شده است. از طریق بررسی عملکرد و کاربرد مورد نظر، می‌توان نشان داد که تا چه حد نتایج به‌دست آمده ناشی از استفاده از هستی‌شناسی بوده است.
 - ارزیابی براساس سطح نحوی^۵ هستی‌شناسی: ارزیابی در این سطح ممکن است بیشتر مورد توجه آن دسته از هستی‌شناسی‌هایی باشد که به‌صورت دستی تهیه می‌شوند. هستی‌شناسی به‌طور معمول، در زبانی رسمی و خاص توصیف می‌شود و باید با ویژگی‌های نحوی آن زبان هماهنگ باشد. حضور اسناد زبان طبیعی و جلوگیری از

1. Lexical, vocabulary, or data layer
3. Hierarchy or taxonomy
5. Syntactic level

2. Facts
4. Context or application level

ایجاد حلقه میان تعاریف از مسائل دیگری است که باید در این زمینه مورد توجه قرار گیرند (Gomez-Perez 1994 cited in Brank, Grobelnic, and Mladenec 2005).

- ارزیابی براساس ساختار و معماری و طراحی هستی‌شناسی^۱: این نوع ارزیابی هم بیشتر مورد توجه آن دسته از هستی‌شناسی‌هایی است که به صورت دستی تهیه می‌شوند. در طراحی و ساخت هستی‌شناسی باید اصول و معیارهای از پیش تعیین شده خاصی مورد توجه قرار گیرند. سازماندهی هستی‌شناسی و قابلیت آن برای گسترش بیشتر در آینده، برخی اصول ساختاری مورد توجه در ارزیابی هستی‌شناسی‌ها هستند (Gomez-Perez 1994, 1996 cited in Brank, Grobelnic, and Mladenec 2005). این نوع ارزیابی به طور معمول، کامل به صورت دستی انجام می‌شود (Brank, Grobelnic, and Mladenec 2005).

۹. سخن پایانی

امروزه، هستی‌شناسی‌ها ابزارهای مهمی برای بازنمون حوزه‌های موضوعی مختلف محسوب می‌شوند. گروه‌های مختلفی نیز در طراحی هستی‌شناسی همکاری دارند یا آن را در فعالیت‌های خود مورد استفاده قرار می‌دهند. دانشمندان علوم رایانه، کتابداران و اطلاع‌رسانان، مهندسان دانش هوش مصنوعی و متخصصان حوزه‌های موضوعی برخی از گروه‌های تأثیرگذار و مرتبط با هستی‌شناسی هستند. از آنجا که برای بازنمون دانش در یک حوزه موضوعی، ممکن است چندین هستی‌شناسی با کیفیت و ویژگی‌های متفاوت ایجاد شده باشد، نیاز است تا با توجه به نیاز کاربران دانش، بهترین و مناسب‌ترین هستی‌شناسی از بین هستی‌شناسی‌های موجود انتخاب گردد. در این مقاله، برخی از مراحل، معیارها، رویکردها، و سطوح ارزیابی هستی‌شناسی‌ها مورد بررسی قرار گرفت. دو رویکرد کمی و کیفی در ارزیابی هستی‌شناسی‌ها مطرح شد. مراحل مختلف ارزیابی هستی‌شناسی شامل ارزیابی پیش از مدل‌سازی هستی‌شناسی، ارزیابی در روند مدل‌سازی و ارزیابی پس از طراحی و انتشار هستی‌شناسی مورد بررسی قرار گرفت. معیارهای ارزیابی هستی‌شناسی و نیز رویکردهای مختلف ارزیابی شامل رویکرد استاندارد طلایی، رویکرد وظیفه‌محور، رویکرد استخراج داده و رویکرد معیارمحور در بخش دیگری از مقاله ارائه شد. سطوح مختلف ارزیابی هستی‌شناسی شامل سطح واژگانی و داده‌ای، سطح ساختار سلسله‌مراتبی و رده‌بندی، سطح روابط معنایی، سطح بافت و کاربرد هستی‌شناسی، سطح نحوی هستی‌شناسی و سطح معماری و ساختار هستی‌شناسی، نیز در بخش پایانی مقاله مطرح شد.

1. Structure, architecture, design

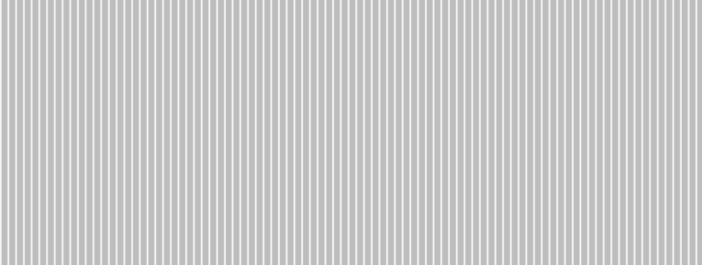
هرچند رویکردهای مختلفی برای ارزیابی هستی‌شناسی وجود دارد، شاید نتوان از میان رویکردهای مختلف، یک رویکرد را به‌عنوان بهترین و مناسب‌ترین رویکرد، انتخاب کرد. زیرا انتخاب یک رویکرد مناسب، به عوامل گوناگونی مانند هدف ارزیابی و کاربرد هستی‌شناسی و جنبه‌ای از هستی‌شناسی که هدف ارزیابی است، بستگی دارد. بنابراین، در انتخاب رویکرد مناسب برای ارزیابی هستی‌شناسی باید دقت نمود.

امید است با ارائه این معیارها و رویکردها بتوان هستی‌شناسی‌هایی با کیفیت و محتوای مناسب، طراحی نمود و یا هستی‌شناسی‌های موجود را بر این اساس ویرایش کرد. این مسأله در رشته کتابداری و اطلاع‌رسانی از اهمیت ویژه‌ای برخوردار است، زیرا سازماندهی مطلوب و درنهایت بازیابی سریع و مناسب، از مهمترین اهداف این رشته است و هستی‌شناسی‌ها ابزارهایی مناسب برای تحقق این اهداف هستند.

۱۰. منابع

- ژاکوب، الین ک. ۲۰۰۳. هستی‌شناسی‌ها و وب معنایی. ترجمه فاطمه شیخ‌شجاعی. ۱۳۸۴. فصلنامه کتاب ۱۶ (۴): ۱۸۹-۱۹۴.
- شریف، عاطفه. ۱۳۸۷. کاربرد هستی‌شناسی‌ها در نظام مدیریت دانش. کتابداری و اطلاع‌رسانی ۱۱ (۳): ۹۷-۱۱۶.
- صنعت‌جو، اعظم. ۱۳۸۴. ضرورت بازنگری در ساختار اصطلاحنامه‌ها: بررسی عدم کارایی اصطلاحنامه‌ها در محیط اطلاعاتی جدید و قابلیت‌های هستی‌شناسی‌ها در مقایسه با آن. فصلنامه کتاب ۱۶ (۴): ۷۹-۹۲.
- Brank, J., M. Grobelnic, and D. Mladenic. 2005. A survey of Ontology evaluation techniques. In *Proceedings of the Conference on Data Mining and Data Warehouses (SiKDD 2005), October 17. Ljubljana, Slovenia*. <http://kt.ijs.si/dunja/sikdd2005/Papers/BrankEvaluationSiKDD2005.pdf> (accessed 24 June 2009).
- Brewster, C., H. Alani, S. Dasmahapatra, and Y. Wilks. 2004. Data driven Ontology evaluation. In *Proceedings of the 4th International Conference on Language Resources and Evaluation, May 26- 28. Lisbon, Portugal*. <http://eprints.aktors.org/337/02/BrewsterLREC-final.pdf> (accessed 1 July 2009).
- Burton-Jones, A., V. C. Storey, V. Sugumaran, and P. Ahluwalia. 2005. A semiotic metrics suite for assessing the quality of ontologies. *Data & Knowledge Engineering* 55 (1): 84-102. <http://ksuseer1.ist.psu.edu/viewdoc/download?jsessionid=67750338FE8079372E31B13FF94AEF2F?doi=10.1.1.91.5586&rep=rep1&type=pdf> (accessed 4 Oct. 2009).
- Gómez-Pérez, A. 1995. Some ideas and examples to evaluate ontologies. In *Proceedings of the 11th Conference on Artificial Intelligence for Applications, February 20-23. Los Angeles, CA, USA*. ftp://ftp.ksl.stanford.edu/pub/KSL_Reports/KSL-94-65.ps.gz (accessed 5 Oct. 2009).
- Hartmann, J. et al. 2005. *Methods for Ontology evaluation*. (No. IST-2004-507482) IST Programme of the Commission of the European. <http://www.starlab.vub.ac.be/research/projects/knowledgeweb/KWeb-Del-1.2.3-Revised-v1.3.1.pdf> (accessed 23 June 2009).
- Labrou, Y., and T. Finin. 1999. Yahoo! as an Ontology: Using Yahoo! categories to describe documents. In *Proceedings of the 1999 ACM Conference on Information and Knowledge Management (CIKM'99), November 2 – 6. Kansas City, Missouri*. http://ebiquity.umbc.edu/_file_directory_/papers/171.pdf (accessed 3 Oct. 2009).

- Lozano-Tello, A., and A. Gómez-Pérez. 2004. ONTOMETRIC: A method to choose the appropriate Ontology. *Journal of Database Manager* 15 (2): 1-18. http://www.infosci-journals.com/downloadPDF/pdf/ITJ2538_IgXdYeT04Y.pdf (accessed 4 Oct. 2009).
- Lu, Q. 2006. OntoKBEval: a support tool for OWL Ontology evaluation. Unpublished master's thesis, Concordia University, Montreal, Quebec, Canada. www.webont.org/owlled/2006/acceptedPosition/submission_18.pdf (accessed 22 June 2009).
- Maedche, A., and S. Staab. 2002. Measuring similarity between Ontologies In *Proceedings Of the European Conference on Knowledge Acquisition and Management - EKAW-2002, October 1-4, Madrid, Spain*. <http://www.aifb.uni-karlsruhe.de/~sst/Research/Publications/ekaw2002-compare.pdf> (accessed 4 Oct. 2009).
- Murshed, A., and R. Singh. 2005. Evaluation and ranking of Ontology construction tools (No. DIT-05-013), Informatica e Telecomunicazioni, University of Trento. <http://eprints.biblio.unitn.it/archive/00000747/01/013.pdf> (accessed 4 Oct. 2009).
- Obrst, L., T. Hughes, and S. Ray. 2006. Prospects and possibilities for Ontology evaluation: the view from NCOR. In *Workshop on Evaluation of Ontologies for the Web (EON2006), May 22, Edinburgh, UK*. <http://km.aifb.uni-karlsruhe.de/ws/eon2006/eon2006obrstetal.pdf> (accessed 24 June 2009).
- Porzel, R., and R. Malaka. 2004. A task-based approach for Ontology evaluation. In *Proceedings of the ECAI Workshop on Ontology Learning and Population: Towards Evaluation of Text-based Methods in the Semantic Web and Knowledge Discovery Life Cycle, August 22- 23, Valencia, Spain*. P. Buitelaar, S. Handschuh, and B. Magnin (Ed.) <http://olp.dfki.de/ecai04/final-porzel.pdf> (accessed 24 June 2009).
- Yu, J., J. A. Thom, and A. Tam. 2007. Ontology evaluation using Wikipedia categories for browsing. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management, November 6-10, Lisbon, Portugal*. <http://goanna.cs.rmit.edu.au/~jyu/publications/YuEtal07.pdf> (accessed 1 July 2009).



Ontology Evaluation: Consideration of Criteria, Approaches and Layers

Akram Fathian Dastgerdi*

PhD Student in LIS, Ferdowsi University of Mashhad

Iranian Journal of
**Information
Processing &
Management**

Abstract: Ontology is commonly used as a structure capturing knowledge about a certain area via providing relevant concepts and relations between them. Nowadays, because of the increase in designing ontologies in different domains, it is important to describe some criteria for selecting the most appropriate ontology. The purpose of this paper is to discuss the ontology evaluation criteria, approaches and layers. At first, different evaluation stages in ontology evaluation were explained. Then the most important approaches to ontology evaluation were described: included gold standard, task-based, data-driven and criteria based evaluation. Another part of this paper was about ontology evaluation criteria such as those mentioned by experts of ontology domain as well as the criteria that proposed by US National Center for ontological Research. Lastly, the levels of ontology evaluation, involved lexical, vocabulary, or data layer, hierarchy or taxonomy layer, other semantic relations layer, context or application level, syntactic level, structure, architecture and design layer were characterized.

Keywords: Ontology, ontology evaluation criteria, ontology evaluation approaches, ontology evaluation layers, ontology evaluation stages

Iranian Research Institute
For Science and Technology
ISSN 2251-8223
eISSN 2251-8231
Indexed in LISA, SCOPUS & ISC
special issue: on Information Storage,
retrieval and Management (winter 2012)

*Corresponding author: fathian2000@gmail.com