

The Relation between the Combination Levels in the Structure of Concepts Entities and the Amount of Semantic Relations in Improving Information Retrieval Performance in UN, FAO's Agricultural Ontology

Maziar Amirhosseini

LIS PhD; Shiraz University; and Knowledge Technology and Management Ph.D., National University of Malaysia; Assistant Professor; Academic Relations and International Affairs (ARIA); Agricultural Research, Education and Extension Organization (AREEO); Tehran, Iran;
Email: mazi_lib@yahoo.com

Received: 05, Dec. 2020 | Accepted: 13, Feb. 2021

Abstract: This research aims to identify the relations between the combination levels in concepts structure with the amount of semantic relations through applying quantitative evaluation. Research hypotheses emphasize that there is a significant relations between the level of combination in the structure of concepts and the number of their semantic relationships. The results of the research background showed that there is no investigation on the relations between structure of concepts with different number of words in their structure and the semantic relations to increase IR performance. Research data are extracted from VocBench as an authoritative agricultural ontology, produced by Food and Agriculture Organization (FAO), United Nations. The size of the research population is 40,000 agricultural concepts. The sample size is 1,500 concepts selected through stratified random sampling. The analysis of the combination levels of concepts structure has been applied using Excel and SPSS software to operate Proportional Analysis and Frequency Analysis. The relations between concepts structure and semantic relations have been adapted inferential statistics, especially Compare Mean Analysis and Pearson Correlation method using SPSS software reports. The research results show that the results of Factoring Ratio is equal 0.58 between one and two-word concepts, meaning that simple concepts comprise of the dominant domain in ontology. The

**Iranian Journal of
Information
Processing and
Management**

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

ISSN 2251-8223

eISSN 2251-8231

Indexed by SCOPUS, ISC, & LISTA

Vol. 36 | No. 4 | pp. 1113-1136

Summer 2021



amount of two-word concepts includes 38.7% of the total concepts. The number of two-word concepts causes decrease of simplicity in concept structure. The average of semantic relations input and output in simple concepts is equal to 6.70 or 7. There is a significant inverse relations between the combination level in concepts structure and the number of taxonomic relations (p -value = 0.000, N = 1500 and r = -0.98) and the number of non-taxonomic relations (p -value = 0.035, N = 1500 and r = -0.54) based on Pearson correlation. It means that an increase in the number of words in concepts structure results in decreasing the number of semantic relations. The results showed that reducing the level of composition in the structure of concepts increases the semantic relations which results in increasing recall and precision in IR performance as well as increasing the integration of the conceptual relations network in ontology. The research results can be used in developing and evaluating the structure of knowledge organization systems (KOSs) such as thesauri regarding structural analysis in IR performance as well as in constructing agricultural ontology of Iran.

Keywords: Concept Structure, Semantic Relations, Taxonomic Relations, Non-Taxonomic Relations, Structural Analysis, Quantitative Evaluations, Agricultural Ontology (VocBench)



ارتباط بین سطح ترکیب در ساختار مدخل‌های مفاهیم و شمار روابط معنایی در تحلیل عملکرد بازیابی اطلاعات در هستی‌شناسی کشاورزی سازمان خواروبار کشاورزی سازمان ملل متحد

مازیار امیرحسینی

دکتری علم اطلاعات و دانش‌شناسی؛ و دکتری
مدیریت و تکنولوژی دانش؛ استادیار؛ مدیریت
اطلاعات و دانش؛ دفتر ارتباطات علمی و همکاری‌های
بین‌المللی؛ سازمان تحقیقات، آموزش و ترویج
کشاورزی؛ وزارت جهاد کشاورزی؛ تهران، ایران؛
mazi_lib@yahoo.com



مقاله برای اصلاح به مدت ۲ روز نزد پدیدآوران بوده است.

پذیرش: ۱۳۹۹/۱۱/۲۵

دریافت: ۱۳۹۹/۰۹/۱۵

نشریه علمی | رتبه بین‌المللی
پژوهشگاه علوم و فناوری اطلاعات ایران
(ایرانداک)

شاپا (چاپی) ۸۲۳۳-۲۲۵۱

شاپا (الکترونیکی) ۸۲۳۱-۲۲۵۱

نمایه در SCOPUS، ISC، LISTA و

jipm.irandoc.ac.ir

دوره ۳۶ | شماره ۴ | صص ۱۱۱۳-۱۱۳۶
تابستان ۱۴۰۰



چکیده: ارزیابی‌های ساختاری به شیوه سامان دادن و نمایش روابط بین مفاهیم در بدنه یک حوزه علمی یا دانش می‌پردازد که در واقع، مبنای شکل‌گیری هستی‌شناسی است. به این ترتیب، تحلیل سطح ترکیب در ساختار مدخل‌های مفاهیم و تأثیر آن بر میزان روابط بین آن‌ها می‌تواند در ساخت، حفظ و نگهداری هستی‌شناسی‌ها با تکیه بر ارزیابی‌های ساختاری حائز اهمیت فراوان باشد. این پژوهش با هدف تبیین ارتباط بین سطح ترکیب در ساختار مفاهیم با شمار روابط معنایی با استفاده از روش ارزیابی کمی صورت گرفته است. این هدف بر این فرضیه تکیه دارد که بین سطح ترکیب در ساختار مفاهیم و تعداد روابط معنایی آن‌ها رابطه‌ای معنادار وجود دارد. نتایج پیشینه پژوهی نشان می‌دهد که در ارتباط بین سطح ترکیب در ساختار مفهوم با روابط معنایی اختصاص یافته به آن‌ها در راستای افزایش عملکرد بازیابی اطلاعات، پژوهشی انجام نشده است. داده‌های این پژوهش از هستی‌شناسی معتبر کشاورزی «سازمان خواروبار کشاورزی سازمان ملل متحد (VocBench)» اخذ شده است. حجم جامعه پژوهش ۴۰۰۰۰ مفهوم کشاورزی است که حدود ۱۵۰۰ مفهوم از طریق

روش نمونه‌گیری تصادفی طبقه‌ای انتخاب شد. به‌منظور تحلیل سطح ترکیب در ساختار مفاهیم با استفاده از نرم‌افزار «اکسل» و «اس‌پی‌اس‌اس» تحلیل‌های نسبی و تحلیل بسامد صورت گرفت. در بررسی ارتباط بین ساختار مفهوم و روابط معنایی از آمار استنباطی، به‌ویژه از مقایسه تحلیل میانگین و روش همبستگی «پیرسون» با استفاده از گزارش‌های نرم‌افزار «اس‌پی‌اس‌اس» مدد گرفته شده. نتایج نشان می‌دهد که نسبت عامل‌بندی مفاهیم یک کلمه‌ای برابر با ۰/۵۸ است؛ به این مفهوم که از هر ۱۰۰ مفهوم یک و دو-کلمه‌ای، ۵۸ مفهوم یک کلمه‌ای در هستی‌شناسی کشاورزی وجود دارد. مفاهیم مرکب دو-کلمه‌ای ۳۸/۷ درصد از کل مفاهیم را به خود اختصاص داده است. دلیل کاهش نسبت سادگی در ساختار مفاهیم، شمار قابل توجه مفاهیم مرکب دو کلمه‌ای است. متوسط روابط معنایی ورودی به مفاهیم یک-کلمه‌ای و خروجی از آن‌ها برابر با ۶/۷۰ یا ۷ رابطه معنایی است. بر اساس آنالیز همبستگی «پیرسون» بین شمار روابط تاکسونومیک ($r = -0.098$ و $N = 1500$ و $p\text{-value} = 0.000$) و نیز شمار روابط غیر تاکسونومیک ($r = -0.054$ و $N = 1500$ و $p\text{-value} = 0.035$) وجود دارد. این بدان معناست که افزایش سطح ترکیب در ساختار مفاهیم، عامل کاهش شمار روابط معنایی آن‌ها می‌شود. بنابراین، کاهش سطح ترکیب در ساختار مفاهیم، عامل افزایش روابط معنایی در هستی‌شناسی است که در افزایش جامعیت و مانعیت در نظام بازیابی اطلاعات و ایجاد انسجام در شبکه روابط معنایی در هستی‌شناسی مؤثر است. نتایج این پژوهش می‌تواند به‌طور مستقیم در تدوین و ارزیابی ساختار سایر نظام‌های سازمان دانش^۱ مانند اصطلاحنامه‌ها، به‌منظور افزایش عملکرد در بازیابی اطلاعات و نیز در تدوین هستی‌شناسی کشاورزی ایران مورد استفاده قرار گیرد.

کلیدواژه‌ها: ساختار مفاهیم، روابط معنایی، روابط تاکسونومیک، روابط غیر تاکسونومیک، ارزیابی‌های ساختاری، ارزیابی کمی، هستی‌شناسی کشاورزی (VocBench)

۱. مقدمه

روش‌های ارزیابی هستی‌شناسی‌ها به سه گروه دسته‌بندی می‌شوند: ساختاری، عملکردی، و قابلیت استفاده (Gangemi et al. 2006). ارزیابی‌های ساختاری با توجه به اهمیت آن‌ها (Eynard, et al. 2012) کمتر مورد توجه قرار گرفته‌اند (Gangemi et al. 2005). این گونه از روش‌های ارزیابی‌ها، ساختار مفاهیم (Alani and Brewster 2005)، روابط معنایی (Assal et al. 2009) و ارتباطات شبکه‌ای بین مفاهیم را در شکل‌گیری هستی‌شناسی هدف قرار می‌دهند (Martín-Chozas 2018). ساختار مفهوم از دیدگاه گروهی از پژوهشگران مهندسی و ارزیابی هستی‌شناسی‌ها به‌عنوان ساختار مداخل^۲ مفاهیم است. این گروه بر آن

باورند که مفاهیم شامل مداخل و روابط بین آن‌هاست (Bergman 2015). در اینجا منظور از ساختار مفاهیم، ساختار مداخل مفاهیم است. اهمیت این روش در این نکته نهفته است که ارزیابی‌های ساختاری به‌شیوه سامان دادن و نمایش روابط بین مفاهیم در بدنه یک حوزه علمی یا دانش می‌پردازد که در واقع، مبنای شکل‌گیری هستی‌شناسی است. به این ترتیب، به‌رغم ظرفیت بالای ارزیابی‌های ساختاری در تحلیل حوزه‌ها یا دامنه‌های ساختاری در هستی‌شناسی، بیشتر شاخص‌ها و فرمول‌های ارزیابی در حوزه ارزیابی‌های عملکردی و قابلیت استفاده از هستی‌شناسی ارائه شده است.

حوزه‌های ساختاری می‌تواند دربرگیرنده ساختار مفاهیم و روابط معنایی بین آن‌ها باشد. به‌عبارت دیگر، مفاهیم که نقش بنیادین را در ساختار هستی‌شناسی بازی می‌کنند، دارای ساختار خود هستند (Alani and Brewster 2005). مفاهیم را می‌توان از دیدگاه ساختار پیچیده یا ساده آن‌ها بررسی کرد؛ مانند ترکیب نحوی یک، دو، سه-کلمه‌ای یا بیشتر در ساختار مفاهیم. استناداردهای بین‌المللی بر عامل‌بندی کلمات مرکب در شکل‌گیری کلمات ساده‌تر تأکید کرده‌اند (International Organization for Standardization (ISO) 2011). این امر برای افزایش جامعیت و مانعیت (Leveling et al. 2011) در بهینه‌سازی عملکرد بازیابی اطلاعات (Amirhosseini and Salim 2010) بوده است. مفاهیم ساده می‌توانند با شمار بیشتری از مفاهیم در ارتباط قرار گیرند (Amirhosseini and Salim 2010). این پژوهش به‌دنبال بررسی دو متغیر سطح ترکیب در ساختار مفاهیم و تحلیل ارتباط آن با شمار روابط معنایی است.

ارتباط پیش‌گفته می‌تواند از طریق یک مثال بیشتر روشن شود. مفهوم «اطلاعات» می‌تواند در ارتباطات متعدد معنایی با مفاهیمی مانند «منابع اطلاعات»، «مدیریت اطلاعات»، «فناوری اطلاعات»، «اطلاعات مدیریتی»، «اطلاعات منابع»، «منابع اطلاعات مدیریتی»، «مدیریت منابع اطلاعاتی»، «فناوری منابع اطلاعاتی»، «فناوری اطلاعات مدیریتی»، «فناوری مدیریت اطلاعات» و امثال آن‌ها مرتبط باشد. مفاهیمی که با مفهوم تک‌واژه‌ای «اطلاعات» ارتباط پیدا می‌کنند، لزوماً در یک حوزه موضوعی نبوده و دارای ساختار مرکب پیچیده هستند. این مفاهیم مرکب می‌توانند با شمار کمتری از مفاهیم پیچیده ذکر شده در ارتباط قرار گیرند. برای نمونه، مفهوم «منابع اطلاعات مدیریتی» از نظر معنایی می‌تواند به شکل سلسله‌مراتبی با «اطلاعات مدیریتی»، «فناوری اطلاعات مدیریتی» و «منابع اطلاعات مدیریتی» در ارتباط باشد. بنابراین، می‌توان گفت که مفاهیم ساده از دیدگاه معنایی در

مرحلهٔ بازیابی اطلاعات، با شمار بیشتری از مفاهیم در ارتباط قرار می‌گیرند و این امر می‌تواند به افزایش جامعیت و مانعیت در نظام بازیابی اطلاعات منجر شود.

تعریف‌های عملیاتی

روابط تاکسونومیک^۱: شامل روابط سلسله‌مراتبی اعم و اخص و نیز کل و جزء می‌شود که در اینجا برابر با ۱۶ رابطهٔ دوسویه است؛ برای نمونه:

«محصولات گیاهی» has sub-concept «دانه‌های روغنی»

«دانه‌های روغنی» is sub-concept of «محصولات گیاهی»

روابط تاکسونومیک ورودی^۲: روابط معنایی سلسله‌مراتبی دریافت‌شده توسط مفاهیم است؛ به‌عنوان مثال:

«دانه‌های روغنی» is sub-concept of «نارگیل»

روابط تاکسونومیک خروجی^۳: روابط معنایی سلسله‌مراتبی است که از هر مفهوم خارج می‌شود؛ مانند:

«نارگیل» has sub-concept «دانه‌های روغنی»

روابط غیر تاکسونومیک^۴: روابط معنایی همبسته یا وابستهٔ بین مفاهیم است که شامل ۳۴ رابطهٔ غیر تاکسونومیک دوسویه است؛ به‌عنوان مثال:

«دانه‌ها» has product «محصولات دانه‌دار»

«محصولات دانه‌دار» Product of «دانه‌ها»

روابط غیر تاکسونومیک ورودی^۵: روابط غیر تاکسونومیکی دریافت‌شده یا واردشده به یک مفهوم است؛ مانند:

«محصولات دانه‌دار» product of «دانه‌ها»

روابط غیر تاکسونومیک خروجی^۶: روابط وابستهٔ خارج‌شده از یک مفهوم است؛ مانند:

«دانه‌ها» has product «محصولات دانه‌دار»

1. taxonomic relations

2. taxonomic relations input

3. taxonomic relations output

4. non-taxonomic relations

5. non-taxonomic relations input

6. non-taxonomic relations output

روابط مفهومی یا معنایی^۱: مجموعه روابط تاکسونومیک و غیر تاکسونومیک است که در اینجا شامل ۵۰ نوع رابطه است.

روابط معنایی ورودی^۲: مجموعه روابط معنایی تاکسونومیک و غیر تاکسونومیک ورودی به مفهوم است.

روابط معنایی خروجی^۳: مجموعه روابط معنایی تاکسونومیک و غیر تاکسونومیک خروجی از مفهوم است.

ساختار مفهوم^۴: ساختار مفهوم از دیدگاه نحوی^۵ و معنایی^۶ می‌تواند از یک یا چند واژه به‌منظور تبیین مفهوم استفاده کند. در اینجا، ساختار مفاهیم از دیدگاه نحوی مد نظر است.

هستی‌شناسی^۷: ابزار برقراری ارتباط بین مفاهیم مرتبط در یک سیستم یا متن است (Simperl 2009).

هستی‌شناسی کشاورزی (VocBench): یک هستی‌شناسی پیچیده است، شامل انواع روابط معنایی و حدود ۴۰۰۰۰ مفهوم کشاورزی که توسط «سازمان خواروبار کشاورزی سازمان ملل متحد» تهیه شده و از مهندسی مجدد اصطلاحنامه کشاورزی «اگرووک»^۸ به‌وجود آمده است.

۲. پیشینه پژوهش

در این بخش به پژوهش‌هایی در داخل و خارج از کشور می‌پردازیم که محققان با تکیه بر روش ارزیابی ساختاری و سنجش‌های مرتبط با آن به ارزیابی ساختار مفاهیم و روابط معنایی در سازمان‌های دانش، به‌ویژه در هستی‌شناسی‌ها پرداخته‌اند.

۲-۱ پیشینه پژوهش در ایران

در سال ۱۳۸۶، با تکیه بر تحلیل‌های نسبی، فرمول‌ها و معیارهای نوینی در ارزیابی‌های کمی ساختار توصیفگرها و دامنه روابط معنایی در ارزیابی اصطلاحنامه‌های فارسی پیشنهاد شد (امیرحسینی ۱۳۸۶). در همین سال، ارزیابی کمی مقوله پیوستگی در شبکه توصیفگرهای

1. conceptual or semantic relations

2. semantic relations input

3. semantic relations output

4. concept structure

5. syntax

6. semantic

7. ontology

8. AGROVOC

اصطلاحنامه‌های علوم اسلامی بررسی شد (امیرحسینی ۱۳۸۷). «امیرحسینی» همچنین در یک طرح پژوهشی به ارزیابی ساختار اصطلاحات و دامنه روابط سلسله‌مراتبی و وابسته با استفاده از نسبت انسجام و وابستگی در اصطلاحنامه «کب»^۱ پرداخت (۱۳۸۷ ب). در همین سال، سطح پیش‌همارایی در ساختار توصیفگرهای اصطلاحنامه‌های «اصفا» و «علوم اسلامی» تعیین گردید (امیرحسینی و محبی ۱۳۸۷).

«فتحیان دستگردی» به توصیف روش‌های مختلف ارزیابی هستی‌شناسی‌ها پرداخت (۱۳۹۰). در همین سال، نمونه هستی‌شناسی «اصفا» از لحاظ قابلیت استفاده با اصطلاحنامه آن مقایسه شد (صنعت‌جو و فتحیان ۱۳۹۰). «ابارشی» با تکیه بر استانداردهای «ایزو» و روش‌های تحلیل نسبی به ارزیابی اصطلاحنامه‌های «اصفا»، «یونسکو» و «اصطلاحنامه جامع علوم» پرداخت (۱۳۹۳). «بهشتی و اژه‌ای» با استفاده از روش مت‌آنتولوژی به ارائه طرحی ارزشمند به منظور شکل‌گیری هستی‌شناسی جامع علوم پایه بر اساس مفاهیم و روابط موجود در اصطلاحنامه‌های تدوین‌شده در «ایرانداک» پرداخته‌اند (۱۳۹۴). در پژوهشی دیگر «ثروتی و ولوی» هستی‌شناسی‌های مرتبه بالا را به منظور تدوین هستی‌شناسی دامنه نظامی، امنیتی و مدیریت بحران مورد مقایسه قرار دادند (۱۳۹۷). در سال ۱۳۹۸، دامنه روابط تامسونومیک و غیر تامسونومیک به روش تحلیل‌های ساختاری در هستی‌شناسی «کشاورزی VocBench» ارزیابی شد (امیرحسینی ۱۳۹۸). وی همچنین، رفتار تعاملی بین ساختار ساده و پیچیده مفهوم با روابط معنایی در هستی‌شناسی کشاورزی VocBench را بررسی کرده است (امیرحسینی ۱۳۹۹).

۲-۲. پیشینه پژوهش در خارج

در سال ۲۰۰۳، سنج‌های ریاضی‌ای به منظور شناسایی ارتباط بین مفهوم و روابط معنایی در هستی‌شناسی‌های «آژانس پروژه‌های تحقیقاتی پیشرفته دفاعی ایالات متحده»^۲ پیشنهاد شد (Burton-Jones et al. 2003). «کنگ» و همکاران در ارزیابی پیچیدگی روابط معنایی بین مفاهیم هستی‌شناسی‌های منطبق با زبان مدل‌سازی یکپارچه^۳ (UML) معیارهایی ارائه کردند (Kang et al. 2004). در سال ۲۰۰۵، معیارهای تحلیل پیچیدگی در ساختار روابط معنایی، در سه گروه از هستی‌شناسی‌های ژن^۴ بررسی شد (Mungall 2005). در سال ۲۰۰۶،

1. CAB Thesaurus

2. DARPA Agent Markup Language (DAML)

3. Unified Modeling Language (UML)

4. gene ontologies (GO)

فرمول تحلیل ساختار مفاهیم و روابط سلسله‌مراتبی در ویرایش‌های مختلف هستی‌شناسی ژن ارائه شد (Zhang et al. 2006). «ژانگ، یه و یانگ» میزان پیچیدگی ساختار هستی‌شناسی بر اساس تحلیل‌های نسبی در ارتباط بین مفاهیم و روابط معنایی را در هستی‌شناسی ژن بررسی کردند (Zhang, Ye and Yang 2006). در سال ۲۰۰۹، «فورلتی» از طریق روش تحلیل پیوند بین مفاهیم و الگوریتم‌های داده‌کاوی، مهم‌ترین مفاهیم در هستی‌شناسی‌ها را مورد بررسی قرار داد (Furletti 2009). «امیرحسینی و سلیم» روش‌شناسی OntoAbsolute را با تکیه بر سنج‌های تحلیل نسبی به‌منظور تحلیل سادگی در ساختار مفهوم و تحلیل دامنه انواع روابط معنایی در هستی‌شناسی‌های سطح پایین، میانی و سطح بالا^۱ پیشنهاد کردند (Amirhosseini and Salim 2011). در سال ۲۰۱۶، روش ارزیابی ساختاری مستقل از گراف^۲ در ارزیابی رفتار تعاملی بین ساختار مفهوم و روابط معنایی پیشنهاد شد (Amirhosseini 2016). در سال ۲۰۱۷، با تکیه بر معیارهای شباهت معنایی^۳، شباهت بین روابط معنایی و خوشه‌های آن‌ها در هستی‌شناسی‌های تهدیدهای امنیت سایبری بررسی شد (Chmielewski and Stpor 2016). «زامنآ، بریگنول و ماگیتمان» ساختار هستی‌شناسی سلسله‌مراتبی DMOZ^۴ در شناسایی نشانی‌ها و منابع اینترنتی در بررسی قابلیت هدایت استفاده‌کننده به مفاهیم مورد نظر، از طریق معیار اتصال و مرکزیت را بررسی کردند (Xamena, Brignole and Maguitman 2017). در سال ۲۰۱۸، «جین و مایر» ساختار ارتباط بین مفاهیم و دامنه روابط مفهومی در نمایش مؤثر دانش در هستی‌شناسی وضعیت اضطراری را ارزیابی کردند (Jain and Meyer 2018). در سال ۲۰۱۹، «الغمدی» و همکاران مفاهیم و جایگاه آن‌ها در خوشه‌های روابط معنایی را با استفاده از سنج‌های کمی در هستی‌شناسی‌های علوم ژنتیک تبیین کردند (Alghamdi et al. 2019) و در سال ۲۰۱۹، «امیرحسینی و سلیم» روش نوین ارزیابی کمی ساختاری هستی‌شناسی‌ها در مقایسه با سایر روش‌های ارزیابی را ارائه کردند (Amirhosseini & Salim 2019).

نتایج پیشینه پژوهشی نشان داد که پژوهش‌های گوناگونی در تحلیل کمی ساختار مفاهیم و ارتباط آن با روابط معنایی در داخل و خارج از کشور صورت گرفته است. مقالات پژوهشی مشترک (Amirhosseini and Salim (2019; 2011)، پایان‌نامه‌های تحصیلی «امیرحسینی» (۱۳۸۶ و ۲۰۱۶)، مقالات پژوهشی «امیرحسینی» (۱۳۸۶ و ۱۳۸۷)، طرح پژوهشی

1. upper, middle and ower-level ontologies

2. graph-independent approach

3. semantic similarity

4. DMOZ is short form of "Directory Mozilla"

«امیر حسینی» (۱۳۸۶) و پایان‌نامه «ابارشی» (۱۳۹۳) از دیدگاه روش‌شناسی با این پژوهش در ارتباط هستند. تحلیل رفتار تعاملی بین ساختار مفهوم و روابط معنایی در طرح پژوهشی امیر حسینی (۱۳۹۹) از دیدگاه تحلیل‌های ساختاری با این پژوهش دارای سنخیت هستند. لیکن، نتایج حاصل از پیشینه پژوهی نشان می‌دهد که تاکنون پژوهشی به شکل دقیق و موشکافانه به تحلیل سطح ترکیب در ساختار مفهوم و میزان ارتباط آن با شمار روابط معنایی نپرداخته است.

۳. بیان مسئله

این پژوهش در پی ارزیابی ساختاری هستی‌شناسی‌ها با تکیه بر روش‌های کمی است. در تحلیل‌های ساختاری هستی‌شناسی‌ها، ساختار مدخل‌های مفاهیم و روابط معنایی بین آن‌ها که در واقع، مداخل اصلی هستی‌شناسی‌ها هستند، مورد مطالعه قرار می‌گیرند (Martín-Chozas 2018). ساختار مداخل مفاهیم را می‌توان از منظر سطح ترکیب کلمات در ساختار مفهوم از دیدگاه نحوی مورد ارزیابی قرار داد. استانداردهای بین‌المللی بر عامل‌بندی یا شکستن ساختار مفاهیم مرکب به منظور شکل‌گیری مفاهیم با ساختار ساده (International Organization for Standardization (ISO) 2011) که از طریق مفاهیم یک کلمه‌ای نمایش داده می‌شود، تأکید کرده‌اند (National Information Standards Organization (NISO) 2005). مفاهیم ساده عامل افزایش جامعیت و مانعیت (Lazarinis et al. 2009) در راستای افزایش عملکرد بازیابی اطلاعات (Leveling, Magdy & Jones 2011) در سازمان‌های دانش هستند (Monz and de Rijke 2002). به این ترتیب، مفاهیم ساده یا تک‌واژه‌ای نقش قابل توجهی در افزایش عملکرد یا کارایی ذخیره و بازیابی اطلاعات و دانش بازی می‌کنند (Amirhosseini and Salim 2010). بنابراین، ارزیابی‌های مبتنی بر تحلیل‌های ساختاری از ظرفیت قابل توجهی در ارزیابی دامنه‌های ساختاری و روابط متقابل بین آن‌ها در هستی‌شناسی‌ها در راستای افزایش جامعیت و مانعیت در ذخیره و بازیابی اطلاعات برخوردار هستند.

بر اساس حدس علمی این پژوهش که به دنبال اثبات آن هستیم، مفاهیم با سطح کمتری از ترکیب کلمات در ساختار، عامل افزایش تعداد روابط معنایی هستند. افزایش روابط معنایی بین توصیفگرها و مفاهیم ساده، عامل کاهش حجم سازمان‌های دانش و افزایش میزان اثر بخشی ذخیره و بازیابی اطلاعات است (Amirhosseini and Salim 2015).

نتایج پیشینه پژوهی نشان می‌دهد که در حوزه ارزیابی‌های ساختاری به مقوله تحلیل ارتباط بین ساختار مفهوم با سطوح ترکیب گوناگون کلمات در ساختار آن‌ها و شمار روابط معنایی پرداخته نشده است. بنابراین، نبود پژوهش درباره ارتباط بین حوزه‌های ساختاری در هستی‌شناسی‌ها به منظور بهبود عملکرد بازیابی اطلاعات، محسوس است. تأیید رابطه مورد نظر بین دامنه‌های ساختاری پیش‌گفته، عامل هدایت سازندگان و تدوین‌کنندگان هستی‌شناسی در گسترش کاربرد مفاهیم با سطوح ترکیب کلمات کمتر در ساختار در مقوله تدوین، حفظ و نگهداری اثربخش هستی‌شناسی در راستای تعیین اعتبار و صحت آن‌ها، جهت افزایش عملکرد بازیابی اطلاعات است. به این ترتیب، شناسایی ارتباط بین سطح ترکیب کلمات در ساختار مفاهیم از نقطه نظر سطح ترکیب واژه‌ها از نظر نحوی و تحلیل ارتباط این ساختار با میزان دریافت و ارسال انواع روابط معنایی، مانند روابط معنایی تاکسونومیک ورودی و خروجی و روابط معنایی غیر تاکسونومیک ورودی و خروجی، مسئله اصلی این پژوهش است. این بدان مفهوم است که اگر حدس‌های علمی این پژوهش در افزایش شمار روابط معنایی در مفاهیم با سطح ترکیب کمتر از کلمات در ساختار نحوی به اثبات برسد، می‌بایست تدوین‌کنندگان و ارزیابی‌کنندگان هستی‌شناسی‌ها در کاربرد مفاهیم پیچیده و مرکب تجدید نظر نمایند. از طرف دیگر، هستی‌شناسی‌هایی که از سطح ترکیب کمتر کلمات در ساختار مفاهیم برخوردار هستند، هستی‌شناسی‌های منسجم‌تر از نظر روابط معنایی بین مفاهیم خواهند بود. به همین دلایل است که مسئله این پژوهش در ارتباط بین ساختار مفاهیم و شمار روابط معنایی اختصاص یافته به آن‌ها می‌تواند در ارزیابی، ساخت و انتخاب هستی‌شناسی‌های مناسب نقش مهمی ایفا نماید.

۴. هدف‌های پژوهش

هدف‌های ویژه این پژوهش به شرح زیر است:

۱. بررسی سطح ترکیب در ساختار مفاهیم کشاورزی در هستی‌شناسی «کشاورزی VocBench».
۲. بررسی فراوانی مفاهیم کشاورزی که در ساختار آن‌ها از شمار متفاوتی از کلمات استفاده شده است.
۳. یافتن ارتباط بین متوسط تعداد انواع روابط معنایی با گروه مفاهیمی که در ساختار آن‌ها از شمار متفاوتی از واژه‌ها استفاده شده است.

۴. تفسیر همبستگی بین سطح ترکیب در ساختار مفاهیم با شمار روابط معنایی در ساختار مفاهیم با شمار روابط معنایی در هستی‌شناسی «VocBench».

۵. فرضیه‌های پژوهش

در این پژوهش، مطابق با اهداف سوم و چهارم تحقیق، دو فرضیه در تحلیل ارتباط بین ساختار مفهوم و شمار روابط معنایی ارائه می‌شود که سعی بر آزمون و سنجش آن‌ها در هستی‌شناسی «کشاورزی VocBench» داریم. اولین فرضیه که در ارتباط با هدف سوم پژوهش است، به‌عنوان فرضیه عمومی و فرضیه دوم که در راستای تحقق هدف چهارم پژوهش است، به‌عنوان فرضیه اصلی شناخته می‌شود. در واقع، فرضیه اول پژوهش به‌عنوان مکمل فرضیه اصلی یا فرضیه دوم در نظر گرفته می‌شود. فرضیه‌های عمومی و اصلی این پژوهش به‌شرح زیر هستند:

۱. فرضیه عمومی: حدس علمی این پژوهش به‌دنبال دستیابی به دانش عمومی در خصوص ارتباط بین سطح ترکیب در ساختار مفاهیم با شمار روابط معنایی است. این فرضیه که با هدف سوم پژوهش به‌منظور انجام استدلال‌ات مورد نظر در ارتباط است، ناظر بر وجود تفاوت بین میانگین شمار روابط معنایی با مفاهیم با سطوح مختلف ترکیب در ساختار آن‌هاست. این فرضیه ناظر بر این مقوله است که «مفاهیم با ساختارهای ساده به‌طور متوسط به شمار بیشتری از روابط معنایی دست می‌یابند».
۲. فرضیه اصلی: این فرضیه، که با هدف چهارم در این پژوهش در ارتباط است، عامل دستیابی به دانش تخصصی و دقیق در خصوص ارتباط میان ساختار مفهوم و شمار روابط معنایی است. حدس علمی در خصوص این فرضیه عبارت است از اینکه: «افزایش سطح ترکیب کلمات در ساختار مفاهیم، عامل کاهش شمار روابط معنایی اختصاص یافته به آن‌هاست».

۶. روش پژوهش

روش‌های به‌کاررفته در این پژوهش، کمی مبتنی بر تحلیل‌های نسبی^۱، تحلیل بسامد^۲، مقایسه متوسط داده^۳ و روش همبستگی «پیرسون»^۴ در تحلیل سطح ترکیب در ساختار مفاهیم

1. proportional analysis

2. frequency analysis method

3. compare mean data analysis

4. person correlation method

و ارتباط آن با شمار روابط معنایی اختصاص یافته به آن‌ها در هستی‌شناسی «کشاورزی هستی‌شناسی» است. به‌طور کلی، می‌توان گفت که رویکرد کمی در ارزیابی ساختار هستی‌شناسی‌ها (Brewster et al. 2004) به‌منظور کسب اطلاعات دقیق در راستای استخراج دانش عمومی و تخصصی از داده‌های ساختاری است (Velardi et al. 2005). از آنجا که این پژوهش سعی بر تحلیل دامنه‌های ساختاری در هستی‌شناسی‌ها دارد، رویکرد مناسب در روش این تحقیق، روش پژوهش کمی است. در این خصوص، از سازوکارهای تحلیل‌های نسبی، آمار توصیفی و آمار استنباطی برای سنجش داده‌ها استفاده می‌شود. مراحل انجام تحقیق با عنایت به روش‌شناسی و تکنیک‌های پیش‌گفته به‌شرح زیر است:

۱. تعیین متغیرهای مستقل و وابسته در پژوهش در راستای تعیین ماهیت داده‌های مورد بررسی در حوزه‌های ساختاری هستی‌شناسی کشاورزی
 ۲. تعیین جامعه نمونه از طریق روش نمونه‌گیری طبقه‌بندی‌شده. هستی‌شناسی کشاورزی «سازمان ملل متحد» شامل ۲۵ خوشه اصلی است که هر یک از این خوشه‌ها در هفت سطح (شامل خوشه‌های ذیل خوشه‌های اصلی) به شکل سلسله‌مراتبی گسترش یافته است. مفاهیم یا جامعه مورد تحقیق در این پژوهش از ۲۵ خوشه اصلی در هفت سطح به شکل تصادفی انتخاب شده است. تعداد جامعه نمونه با عنایت به جدول Morgan (1970) & Krejcie با در نظر گرفتن ۴۰۰۰۰ مفهوم کشاورزی در این هستی‌شناسی، با ۹۵ درصد اطمینان و مثبت و منهای ۲٫۵ درصد خطا، معادل ۱۵۰۰ مفهوم است.
 ۳. استخراج داده‌های پژوهش از هستی‌شناسی کشاورزی، مبتنی بر فاکتورهای مرتبط با دامنه‌های ساختاری در هستی‌شناسی مورد نظر که شامل ساختار مفهوم و شمار انواع روابط معنایی است.
 ۴. ورود اطلاعات به نرم‌افزار تحلیل آماری «اس‌پی‌اس‌اس» و «اکسل»
 ۵. در تحلیل‌های آماری و کمی از سنجش‌ها و تحلیل‌های متعدد و چندگانه، مانند تحلیل‌های نسبی و تحلیل بسامد که در ارائه یافته‌های پژوهش مکمل یکدیگر محسوب می‌شوند، به دانش عمومی و تخصصی پیرامون ساختار مفهوم از نقطه نظر سطح ترکیب کلمات در ساختار آن‌ها دست می‌یابیم.
- ۱-۵. رویکرد تحلیل نسبی: در این مرحله با اخذ گزارش از داده‌های وارد شده به نرم‌افزار «اکسل» زمینه تحلیل‌های نسبی در خصوص شناسایی سطح ترکیب کلمات در ساختار نحوی مفاهیم فراهم می‌آید.

۵-۲. رویکرد تحلیل بسامد: به منظور پشتیبانی از نتایج حاصل از تحلیل‌های نسبی و دستیابی به دانش عمیق‌تر در خصوص بررسی ساختاری مفاهیم، با تکیه بر روش تحلیل بسامد از نرم‌افزار «اس‌پی‌اس‌اس» گزارش گرفته می‌شود.

۶. سنجش فرضیه‌های پژوهش: در اینجا، دو نوع فرضیه عمومی و اصلی پژوهش که مکمل یکدیگر هستند، با تکیه بر روش‌های آمار توصیفی و استنباطی مورد تحلیل و بررسی قرار می‌گیرند که به شرح زیر است:

۱-۶. سنجش فرضیه عمومی: در این فرضیه به تحلیل ارتباط متوسط شمار روابط معنایی با ساختار مفاهیم با سطوح گوناگون کاربرد کلمات در ساختار نحوی آن‌ها با تکیه بر روش مقایسه متوسط داده می‌پردازیم. نتایج حاصل از سنجش این فرضیه که بر اساس گزارش نرم‌افزار «اس‌پی‌اس‌اس» است، به افزایش دانش عمومی نسبت به برقراری ارتباط پیش گفته منجر می‌شود.

۲-۶. سنجش فرضیه اصلی: این فرضیه ناظر بر تحلیل میزان همبستگی بین ساختار مفهوم با سطوح ترکیب گوناگون از کلمات با شمار روابط معنایی مرتبط با آن‌هاست. فرضیه اصلی با تکیه بر روش همبستگی «پیرسون» مورد سنجش قرار می‌گیرد و گزارش مربوطه از نرم‌افزار «اس‌پی‌اس‌اس» استخراج خواهد شد. در واقع، فرضیه عمومی به عنوان پشتیبان فرضیه اصلی عمل می‌کند و در نهایت، ارتباط بین ساختار مفهوم و شمار روابط معنایی اختصاص یافته به آن‌ها به شکل گزاره نظری ارائه می‌شود.

پژوهشگاه علوم انسانی و مطالعات فرهنگی

رتال جامع علوم انسانی

۷. تحلیل داده‌ها و یافته‌های پژوهش

۱-۷. تحلیل سطح ترکیب در ساختار مفاهیم

در تحلیل نسبی سطح ترکیب در ساختار مفاهیم به منظور تشریح هدف اول پژوهش، از فرمول نسبت عامل‌بندی^۱ (امیرحسینی ۱۳۸۷؛ Amirhosseini and Salim 2010) استفاده شده است. فرمول نسبت عامل‌بندی به شرح زیر است:

$$FR = \frac{a}{a + b}$$

1. factoring ratio

قابل ذکر است که در نامگذاری این نسبت که توسط امیرحسینی (۱۳۸۶) ابداع گردیده، از پیشنهاد استاد فقید جناب آقای دکتر عباس حری استفاده شده است.

$a =$ تعداد مفاهیم با (n کلمه) در آنها ساختار

$b =$ تعداد مفاهیم با ($n + 1$ کلمه) در ساختار آنها

جدول ۱. محاسبه سطح ترکیب گوناگون کلمات در ساختار مفاهیم

درصد	نسبت	شمار مفاهیم	ساختار مفهوم
۵۸	۰/۵۸	۸۱۰	مفاهیم یک کلمه‌ای
۸۵	۰/۸۵	۵۸۰	مفاهیم دو کلمه‌ای
۸۹	۰/۸۹	۹۷	مفاهیم سه کلمه‌ای
۹۲	۰/۹۲	۱۲	مفاهیم چهار کلمه‌ای
۱۰۰	۱	۱	مفاهیم پنج کلمه‌ای
		۱۵۰۰	جمع

بر اساس جدول فوق، تعداد مفاهیم یک کلمه‌ای (۸۱۰ مفهوم) بیشتر از سایر مفاهیم مرکب چند کلمه‌ای (۶۹۰ مفهوم) است. نسبت عامل‌بندی مفاهیم یک کلمه‌ای نیز معادل $۰/۵۸ ((۸۱۰ / (۸۱۰ + ۵۸۱))$ است. این بدان مفهوم است که از هر ۱۰۰ مفهوم یک و دو کلمه‌ای، ۵۸ مفهوم یک کلمه‌ای وجود دارد. این نسبت به شکلی چشمگیر در نسبت بین مفاهیم مرکب با بیش از دو کلمه در ساختار آنها افزایش می‌یابد و نتایج نسبت عامل‌بندی به شکلی تقریباً مشابه در این نسبت‌ها دیده می‌شود.

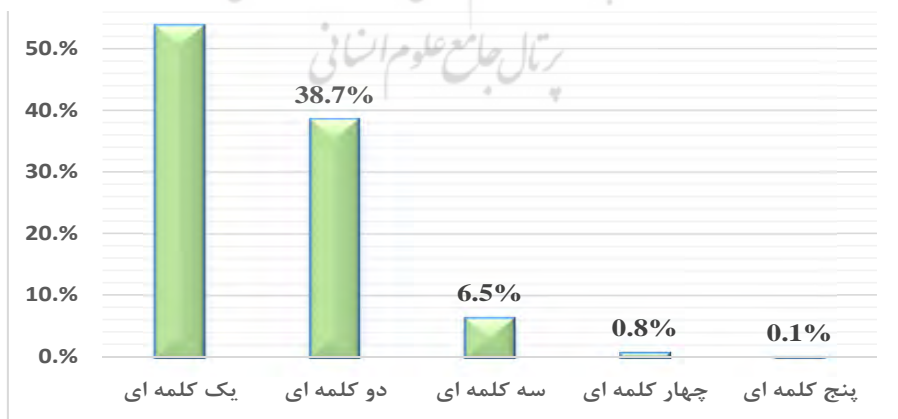
۲-۷. بسامد کاربرد مفاهیم با سطوح ترکیب گوناگون کلمات در ساختار آنها

در این بخش به منظور تبیین دقیق‌تر دامنه کاربرد این گونه مفاهیم در بررسی هدف دوم پژوهش از تحلیل بسامد استفاده شده است.

جدول ۲. تحلیل بسامد کاربرد مفاهیم با سطوح مختلف کاربرد کلمات در ساختار آن‌ها

ساختار مفهوم	بسامد	درصد بسامد	درصد معتبر	درصد تجمعی
داده معتبر	۸۱۰	۵۳/۹	۵۳/۹	۵۳/۹
مفاهیم یک کلمه‌ای	۵۸۰	۳۸/۷	۳۸/۷	۹۲/۶
مفاهیم دو کلمه‌ای	۹۷	۶/۵	۶/۵	۹۹/۱
مفاهیم سه کلمه‌ای	۱۲	۰/۸	۰/۸	۹۹/۹
مفاهیم چهار کلمه‌ای	۱	۰/۱	۰/۱	۱۰۰
مفاهیم پنج کلمه‌ای	۱۵۰۰	۱۰۰	۱۰۰	
جمع	۰			
داده ناقص سیستم	۱۵۰۰			
جمع				

در این جدول، شمار مفاهیم ساده یک کلمه‌ای برابر با ۸۱۰ مفهوم است که بیش از نیمی (۵۳/۹ درصد) از کل مفاهیم را در جامعه نمونه تشکیل می‌دهد. از سوی دیگر، مفاهیم دو کلمه‌ای ۳۸/۷ درصد (معادل ۵۸۰ مفهوم) از کل مفاهیم را شامل می‌شود. در مقابل، شمار مفاهیم سه، چهار و پنج کلمه‌ای، برابر با ۱۱۰ مفهوم است که معادل ۷/۴ درصد (به ترتیب، ۶/۵، ۰/۸ و ۰/۱ درصد) از کل جامعه نمونه است. درصد بسامد تجمعی مفاهیم نشان می‌دهد که مفاهیم یک و دو کلمه‌ای، دامنه غالب مفاهیم (۹۲/۶ درصد) در جامعه نمونه است. بنابراین، نتایج حاصل از تحلیل بسامد، نتایج تحلیل‌های نسبی را مورد تأیید قرار می‌دهد. نمودار زیر که درصد بسامد معتبر را به نمایش می‌گذارد، موضع را بیشتر روشن می‌کند.



نمودار ۱. درصد بسامد معتبر شمار مفاهیم با سطوح گوناگون کاربرد کلمات در ساختار آن‌ها

نمودار فوق آشکار می‌سازد که شمار کلمات ساده یک کلمه‌ای بیشتر از مفاهیم مرکب است و حجم غالب مفاهیم شامل مفاهیم یک و دو کلمه‌ای است.

۷-۳. ارتباط بین سطح ترکیب در ساختار مفاهیم با متوسط شمار روابط معنایی

برای آزمون فرضیه عمومی پژوهش، گزارش روش مقایسه متوسط داده‌ها در جدول ۳، آمده است.

جدول ۳. مقایسه متوسط شمار روابط معنایی با گروه مفاهیم با سطوح ترکیب کلمات گوناگون در ساختار نحوی آن‌ها

ساختار مفهوم	روابط تاکسونومیک	روابط غیر تاکسونومیک	روابط ورودی	روابط معنایی خروجی	روابط معنایی ورودی و خروجی
مفاهیم یک کلمه‌ای	متوسط ۵/۱۳	۱/۵۷	۲/۱۱	۴/۵۹	۶/۷۰
تعداد	۸۱۰	۸۱۰	۸۱۰	۸۱۰	۸۱۰
مفاهیم دو کلمه‌ای	متوسط ۲/۷۲	۱/۱۰	۱/۹۶	۱/۸۵	۳/۸۱
تعداد	۵۸۰	۵۸۰	۵۸۰	۵۸۰	۵۸۰
مفاهیم سه کلمه‌ای	متوسط ۲/۶۰	۱/۱۵	۲/۰۳	۱/۷۲	۳/۷۶
تعداد	۹۷	۹۷	۹۷	۹۷	۹۷
مفاهیم چهار کلمه‌ای	متوسط ۲/۰۰	۰/۴۲	۱/۵۰	۰/۹۲	۲/۴۲
تعداد	۱۲	۱۲	۱۲	۱۲	۱۲
مفاهیم پنج کلمه‌ای	متوسط ۲/۰۰	۱/۰۰	۲/۰۰	۱/۰۰	۳/۰۰
تعداد	۱	۱	۱	۱	۱
جمع	متوسط ۴/۰۰	۱/۳۵	۲/۰۴	۳/۳۱	۵/۳۵
تعداد	۱۵۰۰	۱۵۰۰	۱۵۰۰	۱۵۰۰	۱۵۰۰

بر اساس جدول فوق، مفاهیم ساده یک کلمه‌ای به نسبت مفاهیم مرکب، به تعداد بیشتری از انواع روابط معنایی دست می‌یابند. متوسط روابط معنایی تاکسونومیک، غیر تاکسونومیک، روابط معنایی ورودی و خروجی در این گونه مفاهیم به ترتیب، برابر با ۵/۱۳، ۱/۵۷، ۲/۱۱ و ۴/۵۹ است. نکته قابل توجه این است که متوسط روابط معنایی ورودی و خروجی مفاهیم یک کلمه‌ای، برابر با ۶/۷۰ یا ۷ رابطه معنایی است که تقریباً دو برابر روابط معنایی مفاهیم مرکب است. نکته دیگر این است که متوسط روابط معنایی خروجی

از مفهوم یک کلمه‌ای برابر با ۴/۵۹ است که تقریباً دو برابر متوسط روابط معنایی ورودی به مفهوم (معادل ۲/۱۱) است.

۷-۴. همبستگی بین سطح ترکیب در ساختار مفاهیم با شمار روابط معنایی

در تحلیل هدف چهارم تحقیق، جدول شماره ۴، به ارائه گزارش همبستگی «پیرسون» در آزمون فرضیه اصلی پژوهش می‌پردازد.

جدول ۴. تحلیل همبستگی بین سطح ترکیب در ساختار مفهوم با شمار روابط معنایی

همبستگی	سطح ترکیب در مفاهیم	روابط معنایی
	گزارش	همبستگی پیرسون
	۰/۶۵۰	همبستگی
	۰/۰۱۲	معناداری (آزمون دوطرفه)
	۱۵۰۰	تعداد
	۰/۹۰۳ **	همبستگی
	۰/۰۰۰	معناداری (آزمون دوطرفه)
	۱۵۰۰	تعداد
	۰/۰۳۵	همبستگی
	۰/۱۷۸	معناداری (آزمون دوطرفه)
	۱۵۰۰	تعداد
	۰/۷۵ **	همبستگی
	۰/۰۰۴	معناداری (آزمون دوطرفه)
	۱۵۰۰	تعداد
	۰/۹۸ **	همبستگی
	۰/۰۰۰	معناداری (آزمون دوطرفه)
	۱۵۰۰	تعداد
	۰/۵۴ *	همبستگی
	۰/۰۳۵	معناداری (آزمون دوطرفه)
	۱۵۰۰	تعداد

** معناداری همبستگی در سطح ۰/۰۵ (آزمون دوطرفه)

* معناداری همبستگی در سطح ۰/۰۱ (آزمون دوطرفه)

بر اساس آنالیز همبستگی «پیرسون» در مورد روابط تاکسونومیک خروجی $p\text{-value} = ۰/۰۰۴$ و $N = ۱۵۰۰$ و $r = -۰/۹۰۳$ ، غیر تاکسونومیک خروجی $p\text{-value} = ۰/۰۰۰$ و $N = ۱۵۰۰$ و $r = -۰/۷۵$ ، شمار روابط تاکسونومیک $p\text{-value} = ۰/۰۰۰$ و $N = ۱۵۰۰$ و $r = -۰/۹۸$ و شمار روابط غیر تاکسونومیک $p\text{-value} = ۰/۰۳۵$ و $N = ۱۵۰۰$ و $r = -۰/۵۴$ ، ارتباط معنادار معکوس بین روابط معنایی با سطح ترکیب در مفاهیم وجود دارد. به این مفهوم که هر چقدر سطح ترکیب کلمات در مفاهیم افزایش می‌یابد یا مفاهیم پیچیده‌تر می‌شوند، شمار روابط تاکسونومیک خروجی آن‌ها کاهش می‌یابد. در مقابل، رابطه معنادار بین روابط تاکسونومیک و غیر تاکسونومیک ورودی با ساختار مفهوم دیده نمی‌شود. لیکن، مهم‌ترین رابطه معنادار بین شمار روابط تاکسونومیک و غیر تاکسونومیک با ساختار مفهوم گزارش شده که شامل انواع روابط معنایی است.

۸. بحث و بررسی

۸-۱. بررسی دامنه کاربرد مفاهیم با سطوح مختلف ترکیب در ساختار آن‌ها

بر اساس رهنمودهای ایجاد نظام سازمان‌های دانش مانند اصطلاحنامه‌ها، ساخت مفاهیم پیچیده و مرکب ضروری نبوده و دسترس پذیر ساختن آن‌ها در یک نظام معنایی محل اشکال است (International Standards for Organization (ISO) 2011, 2013). از سوی دیگر، رهنمودهای سایر استانداردها بر کاربرد مفاهیم ساده یا تک‌واژه‌ای تأکید دارند (International Standards for Organization (ISO) 1986, 1985; British Standards Institution) (2005), (National Information Standards Organization (NISO), (2005), (BS) 1979, 1985, 2005-2008). کاربرد مفاهیم ساده عامل اصلی افزایش عملکرد بازیابی اطلاعات (Amirhosseini and Salim 2010) در افزایش جامعیت (Lazarinis et al. 2009) و مانعیت (Airio 2006) و نیز باعث کاهش حجم واژگان کنترل شده است (Amirhosseini and Salim 2015). نتایج نشان می‌دهد که بیش از ۵۰ درصد مفاهیم در هستی‌شناسی، مفاهیم ساده یا یک‌کلمه‌ای هستند. البته، حجم مفاهیم مرکب دو کلمه‌ای به مراتب بیشتر از سایر مفاهیم مرکب است. یعنی، هر چقدر ساختار مفاهیم پیچیده‌تر می‌شود، شمار این گونه مفاهیم کاهش می‌یابد. بنابراین، هستی‌شناسی کشاورزی در کاربرد مفاهیم ساده از رهنمودهای استانداردها تبعیت کرده و در مقابل، عامل افزایش مفاهیم پیچیده مفاهیم مرکب دو کلمه‌ای است. این است که جهت

افزایش سطح سادگی ساختار مفاهیم می‌بایست تا حد ممکن، این‌گونه مفاهیم از طریق روش‌های عامل‌بندی تجزیه شوند.

۲-۸. تحلیل بسامد کاربرد مفاهیم با سطوح مختلف ترکیب در ساختار آن‌ها

تحلیل بسامد، ضمن تأیید نتایج حاصل از تحلیل‌های نسبی نشان داد که مفاهیم یک و دو کلمه‌ای دامنه غالب مفاهیم را تشکیل می‌دهند. یعنی، با وجود پیروی هستی‌شناسی کشاورزی از رهنمودهای استانداردها در کاربرد مفاهیم ساده، شمار مفاهیم دو کلمه‌ای به شکلی قابل توجه بیش از سایر مفاهیم مرکب است. بنابراین، حرکت از پیچیدگی به سمت سادگی در ساختار مفاهیم وجود دارد. لیکن، عامل اصلی افزایش دامنه کاربرد مفاهیم پیچیده، مفاهیم مرکب دو کلمه‌ای است.

۳-۸. تحلیل ارتباط بین سطح ترکیب در ساختار مفاهیم با متوسط شمار روابط معنایی

آزمون فرضیه عمومی پژوهش نشان داد که شمار قابل توجهی از روابط معنایی به‌طور متوسط به مفاهیم ساده اختصاص یافته است. به این ترتیب، حدس علمی این پژوهش در ارتباط بین افزایش شمار روابط معنایی در مفاهیم ساده تأیید می‌شود. از سوی دیگر، افزایش شمار روابط معنایی خروجی از مفهوم به نسبت روابط معنایی ورودی به مفهوم، امکان وجود اعم‌نگری^۱ را در هستی‌شناسی کشاورزی افزایش می‌دهد. یعنی، وسعت یا پهنای خوشه‌های مفاهیم در سطح، بیشتر از عمق یا اخص‌نگری^۲ این خوشه‌ها بوده (Jamouille 2015) و عامل افزایش جامعیت در بازیابی است (Gregory et al. 2017).

۴-۸. تحلیل همبستگی بین سطح ترکیب در ساختار مفاهیم با شمار روابط معنایی

گزارش همبستگی نشان داد که هر چقدر تعداد کلمات در ساختار مفهوم افزایش می‌یابد، تعداد روابط معنایی آن با کاهش مواجه می‌شود. به بیان دیگر، مفاهیم یک کلمه‌ای با تعداد بیشتری از مفاهیم در ارتباط خواهند بود. به این ترتیب، حدس علمی این پژوهش مبنی بر ارتباط معنادار بین ساختار مفهوم با شمار روابط معنایی تأیید می‌شود که مؤید نتیجه حاصل از فرضیه عمومی پژوهش نیز است؛ هرچند این ارتباط در روابط تاکسونومیک و غیر تاکسونومیک ورودی دیده نشده است. دلیل این امر

1. exhaustivity

2. specificity

ممکن است در شمار زیاد مفاهیم مرکب دو کلمه‌ای یا افزایش ترکیب در ساختار مفاهیم باشد. برای نمونه، مفاهیمی چون، «قلمه پایه گیاه» یا «گیاه پایه قلمه»، «فنون کشاورزی احیاکننده»^۱ یا «احیاکننده فنون کشاورزی»، «پوشش سطحی محافظ» یا «محافظ پوشش سطحی» و «بیماری تحلیل برنده»^۲ یا «برنده تحلیل بیماری» که مفاهیمی با معانی کاملاً متفاوت هستند، به دلیل مشابهت در کلمات در ساختار نحوی آن‌ها می‌توانند به نادرستی با یکدیگر یا سایر مفاهیم مشابه در ارتباط قرار گیرند.

۹. نتیجه‌گیری

این پژوهش با تکیه بر روش‌های آماری گوناگون به دنبال ارائه دانش عمیق و کاربردی در خصوص ساختار مفهوم و ارتباط آن با روابط معنایی است که در زیر به آن‌ها اشاره می‌شود.

۱. تدوین کنندگان هستی‌شناسی کشاورزی، ضمن پیروی از استانداردها در حرکت از پیچیدگی در ساختار مفهوم به سمت کاربرد مفاهیم ساده‌تر موفق بوده‌اند. عامل افزایش مفاهیم مرکب، کاربرد زیاد مفاهیم مرکب دو کلمه‌ای است. البته، بسیط بودن ساختار مفهوم صرفاً در شکل‌گیری مفاهیم یک کلمه‌ای نیست (مانند مفهوم مرکب دو کلمه‌ای «ویتامین سی» که در واقع، یک مفهوم بسیط است)؛ هرچند که تک‌واژه‌ای بودن مفاهیم، معیار قابل اعتمادی در ارزیابی سادگی در ساختار مفهوم است.
۲. تأیید فرضیه‌های عمومی و اصلی به این مفهوم است که کاربرد کلمات کمتر در ساختار مفهوم، عامل افزایش شمار روابط معنایی بوده و انسجام در شبکه روابط معنایی بین مفاهیم در هستی‌شناسی افزایش می‌یابد. این مقوله عامل افزایش جامعیت و مانعیت در بازاریابی اطلاعات بوده و از حجم واژگان کنترل‌شده کاسته است.
۳. تأیید فرضیه‌های پژوهش به شکل‌گیری یک گزاره نظری در ارتباط بین ساختار مفهوم و شمار روابط معنایی منجر می‌شود. گزاره نظری بر این نکته تأکید دارد که «مفاهیم ساده یا تک‌واژه‌ای شامل شمار بیشتری از روابط معنایی خواهند بود». این مبنای نظری می‌تواند تدوین کنندگان سازمان‌های دانش، به ویژه هستی‌شناسی‌ها را به سمت ساخت مفاهیم ساده‌تر رهنمون باشد و مبنای قابل اعتمادی در ارزیابی انواع سازمان‌های دانش تلقی شود.

1. regenerative agriculture

2. degenerative disease

۴. امکان وجود اعم‌نگری در ساختار روابط معنایی هستی‌شناسی کشاورزی، از یک‌طرف دامنه گسترده خوشه‌های مفاهیم را در مقابل کاهش عمق هستی‌شناسی نشان می‌دهد که این امر در سازماندهی دانش کشاورزی حائز اهمیت است و از طرف دیگر، زمینه بسیار مناسبی را برای پژوهش‌های آینده در ارزیابی اعم‌نگری و اخص‌نگری در این هستی‌شناسی فراهم می‌آورد.

۵. در پایان می‌توان گفت که هستی‌شناسی کشاورزی «سازمان خواروبار کشاورزی سازمان ملل متحد»، یک نظام سازمان دانش کشاورزی منسجم در شبکه روابط معنایی است. به این دلیل که کاربرد قابل توجه مفاهیم ساده، که از طریق حرکت از پیچیدگی به سمت سادگی در ساخت مدخل‌های مفاهیم حاصل می‌شود، عامل اصلی ارتباطات بیشتر در شبکه روابط معنایی در این هستی‌شناسی شده است. این ویژگی، علاوه بر کاهش میزان واژگان در این هستی‌شناسی، نقش مؤثری در افزایش جامعیت و مانعیت در عملکرد نظام ذخیره و بازیابی اطلاعات دارد. با توجه به اینکه تاکنون هستی‌شناسی ملی و بومی کشاورزی ایران شکل نگرفته است، هستی‌شناسی «VocBench» می‌تواند به‌عنوان مبنای شکل‌گیری این هستی‌شناسی در نظر گرفته شود. بنابراین، با تکیه بر مقتضیات و ویژگی‌های بومی، ملی و خاص کشاورزی میهن‌مان می‌توانیم ضمن استفاده از هستی‌شناسی کشاورزی «سازمان خواروبار کشاورزی» به‌عنوان مبنای کار، از نتایج حاصل از این پژوهش در کاربرد یا ساخت مفاهیم با ساختارهای ساده به‌منظور افزایش انسجام درونی در شبکه روابط معنایی و مفهومی بین اصطلاحات و مدخل‌ها در هستی‌شناسی کشاورزی ایران استفاده نماییم.

فهرست منابع

- ابارشی، فاطمه. ۱۳۹۳. ارزیابی اصطلاحنامه‌های اصفاء، یونسکو و جامع علوم بر اساس استاندارد ISO ۲۵۹۶۴: رویکرد هستی‌شناختی. پایان‌نامه کارشناسی ارشد. دانشکده علوم تربیتی و روان‌شناسی. دانشگاه الزهراء.
- امیرحسینی، مازیار. ۱۳۸۶. ارزیابی عوامل کمی و کیفی مؤثر در نظام ذخیره و بازیابی اطلاعات در اصطلاحنامه‌های فارسی. پایان‌نامه دکتری. دانشکده علوم تربیتی و روان‌شناسی. دانشگاه شیراز.
- _____. ۱۳۸۶. ارزیابی کمی نسبت تعادل در اصطلاحنامه‌های علوم اسلامی حرکت به سمت تحلیل توانمندی واژگان مدخل. کلیات کتاب ماه. مهر ۱۳۸۶: ۱۶-۲۹.

_____ ۱۳۸۷. ارزیابی کمی مقوله پیوستگی در شبکه توصیفگرهای اصطلاحنامه اصول فقه، منطق، فلسفه اسلامی، علوم قرآنی و کلام. اصطلاحنامه و کاربردهای آن در محیط الکترونیکی: مجموعه مقالات همایش ملی. قم: مرکز اطلاعات و مدارک اسلامی.

_____ ۱۳۸۷. ب. ارزیابی کمی و کیفی اصطلاحنامه کب (CAB). طرح پژوهشی. شماره مصوب پروژه: ۸۵۰۰۱-۰۰۰۰-۰۱-۵۱۰۰۰۰-۲۰۰-۲. مرکز فناوری اطلاعات و اطلاع‌رسانی کشاورزی. سازمان تحقیقات، آموزش و ترویج کشاورزی. وزارت جهاد کشاورزی.

_____ ۱۳۹۶-۱۳۹۸. گزارش نهایی پروژه/ طرح پژوهشی: تحلیل ساختاری دامنه روابط مفهومی در هستی‌شناسی کشاورزی سازمان خواروبار کشاورزی سازمان ملل متحد VocBench. طرح پژوهشی. شماره مصوب پروژه: ۹۱-۲۷-۰۰۲-۹۶۰۸۶۴. دفتر ارتباطات علمی و همکاری‌های بین‌المللی، سازمان تحقیقات، آموزش و ترویج کشاورزی.

_____ ۱۳۹۹. طرح پژوهشی: تحلیل ساختاری رفتار تعاملی بین ساختار مفهوم و شمار روابط معنایی در هستی‌شناسی کشاورزی سازمان خواروبار کشاورزی سازمان ملل متحد VocBench. شماره مصوب پروژه: ۹۱-۲۷-۰۰۴-۲۲۳-۰۹۹. [در حال اجرا]. تهران: دفتر ارتباطات علمی و همکاری‌های بین‌المللی، سازمان تحقیقات، آموزش و ترویج کشاورزی.

_____ و رؤیا مجبی. ۱۳۸۷. ارزیابی نسبت عامل‌بندی در ساختار توصیفگرهای اصطلاحنامه اصفا و علوم اسلامی: تحلیل سطح پیش‌همارایی. فصلنامه مطالعات ملی کتابداری و سازماندهی اطلاعات ۲۰ (۳): ۴۵-۶۰.

بهشتی، ملوک‌السادات، و فاطمه اژه‌ای. ۱۳۹۴. طراحی و پیاده‌سازی هستی‌شناسی علوم پایه بر اساس مفاهیم و روابط موجود در اصطلاحنامه‌های مرتبط. پژوهشنامه پردازش و مدیریت اطلاعات ۳۰ (۳): ۶۷۷-۶۹۶.

ثروتی، لیلا، و محمدرضا ولوی. ۱۳۹۷. ارزیابی هستی‌شناسی‌های مرتبه بالا و ارائه الزامات هستی‌شناسانه مناسب برای کاربردهای نظامی، امنیتی و مدیریت بحران. فصلنامه علمی - پژوهشی فرماندهی و کنترل ۲ (۲): ۶۵-۸۱.

صنعت‌جو، اعظم، و اکرم فتحیان دستگردی. ۱۳۹۰. مقایسه کارآمدی اصطلاحنامه و هستی‌شناسی در بازنمون دانش (طراحی و ساخت نمونه هستی‌شناسی اصفا). پژوهشنامه کتابداری و اطلاع‌رسانی ۱ (۱): ۲۱۹-۲۴۰. فتحیان دستگردی، اکرم. ۱۳۹۰. ارزیابی هستی‌شناسی‌ها: بررسی معیارها، رویکردها و سطوح. پژوهشنامه پردازش و مدیریت اطلاعات. ۴: ۳-۲۵.

References

- Airio, E. 2006. Word normalization and decomposing in mono- and bilingual IR. *Information Retrieval* 9 (3)271-249 .:
- Aitchison, J, A. Gilchrist, and D. Bawden2000. . *Thesaurus construction and use: a practical manual*. Fourth edition. London: The Association for Information Management (Aslib).
- Alani, H., and C. Brewster. 2005. Ontology Ranking Based on The Analysis of Concept Structures. In *Proc. of the 3rd International Conference on Knowledge Capture (K-Cap)*, Banff, Canada, 2005: pp 51-58.

- Alghamdi, S. M., B. A. Sundberg, J. P. Sundberg, P. N. Schofield, and R. Hoehndorf. 2019. Quantitative evaluation of ontology design patterns for combining pathology and anatomy ontologies. *Scientific Reports* 9 (1): 1-12.
- Amirhosseini, M. 2016. Analysis of concept structure and semantic relations based on graph-independent structural analysis. Ph. D. Dissertation, Supervisor Juhana Salim. Faculty of Information Sciences and Technology, Universiti Kebangsaan Malaysia.
- _____, and J. Salim. 2010. Quantitative Evaluation of Simplicity Invisible Domain in Islamic Knowledge Organizations. In *2010 International Conference on Information Retrieval and Knowledge Management: CAMP 10, exploring the invisible word (17-18 March, 2010, Shah Alam, Malaysia)*. Institute of Electrical and Electronics Engineers, 119-124.
- _____. 2011. OntoAbsolute as an Ontology Evaluation Methodology in Analysis of the Structural Domains in Upper, Middle and Lower Level Ontologies. In *STAIR'11: International Conference on Semantic Technology and Information Retrieval 28th to 29th June 2011, Putrajaya, Kuala Lumpur, Malaysia*. Malaysia: Institute of Electrical and Electronics Engineers, 2011, 26-33.
- _____. 2015. Quantitative evaluation of the movement from complexity toward simplicity in the structure of thesaurus descriptors. *Malaysian Journal of Library & Information Science* 20 (3): 47-62.
- _____. 2019a. Structural Analysis of Semantic Relations regarding Integration and Association of Semantic Network in VocBench as an Agricultural Ontology. *International Journal of Engineering Technology and Management Research* 6 (3): 41-57.
- _____. 2019b. 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* 33 (1): 46-68.
- Assal, H., K. Pohl, and J. Pohl. 2009. The Representation of Context in Computer Software, In *Pre-Conference Proceedings, Focus Symposium on Knowledge Management Systems, InterSymp-2009, Baden-Baden, Germany, 4 August, 2009*.
- Bergman, M. 2015. Conceptual and Practical Distinctions in the Attributes Ontology. In *Adaptive Information, Innovation and Infrastructure (AI 3)*. <https://www.mkbergman.com/1842/conceptual-and-practical-distinctions-in-the-attributes-ontology/#:~:text=Attribute%20%E2%80%94%20an%20inherent%20characteristic%20of,share%20the%20same%20potential%20attributes> (accessed Feb. 14, 2020)
- Brewster, C., H. Alani, S. Dasmahapatra, Y. Wilks. 2004. Data driven ontology evaluation. In *Proc. of the 4th International Conference on Language Resources and Evaluation (LREC)*, Lisbon, Portugal, 2004. European Language Resources Association.
- British Standards Institution (BS). 1979. *BS 5723:1979: Guidelines for the establishment and development of monolingual thesauri*. London: British Standards Institution.
- _____. 1985. *BS 6723:1985. Guide to establishment and development of multilingual thesauri*. London: British Standards Institution.
- _____. 2005-8. *BS 8723: Structured vocabularies for information retrieval - Guide. (Published in five separate parts between 2005 and 2008)*. London: British Standards Institution.
- Burton-Jones, A., V. C. Storey, V. Sugumaran and P. Ahluwalia. 2003. Assessing the Effectiveness of the DAML Ontologies for the Semantic. In *Proc. of the 8th International Conference on Applications of Natural Language to Information Systems*. Burg (Spreewald), Germany, 56-69.
- Chmielewski, M., and P. Stpor. 2016. Medical Data Unification Using Ontology-Based Semantic Model Structural Analysis. witek J., Borzemski L., Grzech A., Wilimowska Z. (eds) In *Information Systems Architecture and Technology: Proceedings of 36th International Conference on Information Systems Architecture and Technology – ISAT 2015 – Part III*. Advances in Intelligent Systems and Computing, 431. Springer, Cham.
- Eynard, D., M. Matteucci and F. A. Marfa. 2012. *Modular Framework to Learn Seed Ontologies From Test. Semi-Automatic Ontology Development: Processes And Resources*. Hershey, PA: Information Science Reference, 2012.
- Furletti, B. 2009. *Ontology-driven knowledge discovery*. Lucca, Italy: IMT Institute for Advanced Studies.

- Gangemi, A., C. Catenacci, M. Ciaramita and J. Lehmann. 2005. A theoretical framework for ontology evaluation and validation. In *Proceedings of the 2nd Italian Semantic Web Workshop SWAP, Italy 2005*.
- Gangemi, A., C. Catenacci, M. Ciaramita and J. Lehmann. 2006. Modelling ontology evaluation and validation. In: Y. Sure, & J. Domingue (eds). *The Semantic Web: Research and Applications. The 3rd European Semantic Web Conference, ESWC 2006. Lecture Notes in Computer Science, vol 4011*. Berlin, Heidelberg: Springer.
- Gregory, K., P. T. Groth, H. Cousijn, A. Scharnhorst, S. Wyatt. 2017. Searching Data: A Review of Observational Data Retrieval Practices. *CoRR abs/1707.06937*. <http://arxiv.org/abs/1707.06937>. (accessed May 5, 2020)
- International Organization for Standardization (ISO). 1985. *ISO 5964:1985: Documentation — Guidelines for the establishment and development of multilingual thesauri*. Geneva: International Organization for Standardization.
- _____. 1986. *ISO 2788: Guidelines for the establishment and development of monolingual thesauri (2nd. ed.)*. Geneva: International Standards Organization.
- _____. 2011. *ISO/FDIS 25964-1: Information and documentation -thesauri and interoperability with other vocabularies - Part 1: Thesauri for information retrieval*. Geneva: International Organization for Standardization; Final Draft circulated April 2011.
- _____. 2013. *ISO/FDIS 25964-2: Information and documentation -thesauri and interoperability with other vocabularies - Part 2: Part 2: Interoperability with other vocabularies*. Geneva: International Organization for Standardization; Final Draft circulated April 2011.
- Jain, S., and V. Meyer. 2018. Evaluation and Refinement of Emergency Situation Ontology. *International Journal of Information and Education Technology* 8 (10): 713-719.
- Jamoulle, M. 2015. The Words of Prevention, Part II: Ten Terms in the Realm of Quaternary Prevention. *Revista Brasileira de Medicina de Família e Comunidade* 10 (35): 1-10.
- Kang, D., B. Xu, J. Lu and W. Chu. 2004. A Complexity Measure for Ontology Based on UML. In *Proceedings of the 10th IEEE International Workshop on Future Trends of Distributed Computing Systems (FTDCS'04 Suzhou, China)* pp.: 222-228.
- Krejcie, R. V., and D. W. Morgan. 1970. Determining sample size for research activities. *Educational and Psychological Measurement* 30: 607-610.
- Lazarinis, F., J. Vilares, J. Tait, and E.N. Efthimiadis. 2009. Current research issues and trends in non-English Web searching. *Information Retrieval* 12 (3): 230-250.
- Leveling, J., W. Magdy, G. J. F. Jones. 2011. An investigation of decompounding for cross-language patent search. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information, July 24-28*. Beijing, China, 1169-1170.
- Martin Chozas, P. 2018. *Towards a Linked Open Data Cloud of Language Resources in the Legal Domain*. Master Thesis, E.T.S. de Ingenieros Informáticos, Universidad Politécnica de Madrid (UPM), 2018.
- Monz, C. and M. de Rijke. 2002. Shallow morphological analysis in monolingual retrieval for Dutch, German, and Italian. *Accessing Multilingual Information Repositories, vol. 2406 of Lecture Notes in Computer Science*. Pp.: 262-277.
- Mungall, C. 2005. *Increased complexity in the GO*. Available at: <http://www.fruitfly.org/~cjm/obol/doc/go-complexity.html> (accessed Feb. 20, 2020)
- National Information Standards Organization (NISO). 2005. *Guidelines for the construction, format, and management of monolingual controlled vocabularies: ANSI/NISO Z39.19-2005*, Bethesda Md.: NISO Press.
- Simperl, E. 2009. Reusing ontologies on the Semantic Web: A feasibility study. *Data & Knowledge Engineering* 68: 905-925.
- Velardi, P., R. Navigli, A. Cucchiarelli, F. Neri, P. Buitelaar, P. Cimiano and B. Magnini (ed.). 2005. *Evaluation of OntoLearn, a methodology for automatic learning of domain ontologies. Ontology Learning from Text: Methods, Evaluation and Applications*. Amsterdam, Netherlands: IOS Press.

- Xamena, E., N. B. Brignole, and A. Maguitman. 2017. Structural Analysis of topic ontologies. *Information Sciences*. 421, 15–29. doi: <https://doi.org/10.1016/j.ins.2017.08.081> (accessed June 12, 2020)
- Zhang, D., C. Ye, and Z. Yang. 2006. An Evaluation Method for Ontology Complexity Analysis in Ontology Evolution. S. Staab and V. Svatek (Eds.). *EKAW 2006, LNAI 4248, International Conference on Knowledge Engineering and Knowledge Management No15, Podebrady, Tcheque, Republique*. 4248, 214-221.

مازیار امیرحسینی

متولد سال ۱۳۴۶ دارای مدرک دکتری در رشته علم اطلاعات و دانش‌شناسی از دانشگاه شیراز و همچنین، دارای مدرک دکتری در رشته مدیریت و تکنولوژی دانش از دانشگاه ملی مالزی است. ایشان هم‌اکنون استادیار مدیریت اطلاعات و دانش دفتر ارتباطات علمی و همکاری‌های بین‌المللی سازمان تحقیقات آموزش و ترویج کشاورزی، وزارت جهاد کشاورزی است.



حوزه‌هایی مانند سازماندهی اطلاعات و دانش، سندداری، نظام‌های اطلاع‌رسانی، ساخت و ارزیابی سازمان‌های دانش و ارزیابی روابط معاشناختی در هستی‌شناسی‌ها از جمله علایق پژوهشی وی است.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
رتال جامع علوم انسانی