Library and Information Science Research

https://Infosci.um.ac.ir Vol. 13, No 2, 2023 Research Article

ISSN(Online): 2538-4112 ISSN(Print): 2251-6344

Methods of the OntologyAutomatic Learning in the Fields of Quranic Concepts: a Scoping Review Study

Ali Mirarab
Assistant Professor, Department of Knowledge Dissemination, Islamic Sciences and Culture Academy, Qom, Iran. Email: alimirarab@isca.ac.ir
Morteza Mohammadi Ostani
Assistant Professor, Department of Knowledge and Information Science, University of Isfahan, Isfahan, Iran (Corresponding Author), Email: m.ostani@edu.ui.ac.ir
Faezehsadat Tabatabaei Amiri
PhD. Student, Knowledge and Information Science, Shahid Chamran University, Ahvaz, Iran, Email: faeze.tabatabai@gmail.com
Somayeh Dehghanisanij
PhD. Student, Medical Library and Information Science, Iran University of Medical Sciences, Tehran, Iran, Email: so.dehghani93@gmail.com

Recived: 2023-07-01 Revised: 2023-09-07 Accepted: 2023-11-11 Published: 2023-12-21

Citation: Mirarab, A., Mohammadi Ostani, M., Tabatabaei Amiri, F., & Dehghanisanij, S. (2023). Methods of the OntologyAutomatic Learning in the Fields of Quranic Concepts: a Scoping Review Study. Library and Information Science Research, 13(2), 27-49. doi: 10.22067/infosci.2023.83083.1169

Abstract

Objective: Today, semantic technology offers a new approach in organizing Quranic knowledge with the aim of providing meaningful information and representing Quranic teachings. Ontologies are a tool to formally express concepts and relationships in a specific domain. In the same way, the development of ontology as a tool for representing the effulgence and extracting the knowledge of the Quran is not only valuable, but also necessary. Ontology learning and its methods automatically to extract concepts are important topics in the field of Semantic Web and its technologies. Recently, the development and application of ontologies learning for the extraction of Quranic concepts has been considered. Therefore, the aim of the current research is to comprehensively investigate the ontologies automatic learning in the field of extracting knowledge and Quranic concepts in order to clarify the current and future situation. The investigated criteria were data set, learning methods, evaluation methods, results and future suggestions of studies in the field of ontologies automatic learning of the Quran.

Methodology: The research was conducted by the scoping review method in accordance with PRISMA guidelines and based on Arksey & O'Malley procedure. This process describes a protocol for matching the results of existing studies with research questions and criteria. The five steps suggested by Arksey & O'Malley are as follows: 1. Identify and design the research question(s), 2. Conduct search strategies advocate for relevant studies through the selection of appropriate keywords and Boolean operators, 3. Final selection of relevant studies, considering the inclusion and exclusion criteria, 4. Tabulating the data, and finally, 5. Reporting its results. Sources were searched in seven scientific databases including Emerald, Science Direct, IEEE Xplore Digital Library, Google Scholar, Web of Science, and Scopus. The search process has

©2022 The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, as long as the original authors and source are cited. No permission is required from the authors or the publishers.

been done in April 2023. A number of 811 articles, regardless of the time limit, were evaluated and selected. In order to organize the retrieved articles, EndNote resource management software was used and after matching the titles in different databases, 317 duplicate articles were removed. After reviewing the abstracts, the entry and exit criteria and the quality of the articles were applied. Also, in order to avoid bias in the selection of articles, during a random review, two independent researchers in the field of ontology automatic learning were evaluated and finally 25 articles were selected as review criteria.

Findings: Most of the study in the field of Quranic data set were in English and Arabic languages, and most of them used the English translation of Al-Hilali and Khan's Quran. The use of a limited data set was the most important limitation of the research conducted in the field of automatic learning of Quranic ontologies. Most of the studies have used normalization methods, text clustering and categorization, text summarization, information extraction, similarity and finding famous entities. Of course, in some studies, artificial intelligence methods such as neural network have also been used. In addition, the findings showed that data mining algorithms based on statistics and probability methods for learning and constructing automatic ontologies was apparently surging in popularity among researchers. Evaluation methods includes calculating accuracy, recall and F criteria in the application of automatic learning algorithms in Quranic ontologies. The studies that have used artificial intelligence techniques, by Semantic analysis, inference, modeling and validation of inferred data have achieved results such as sound recognition for teaching Quran reading, recognition of literary arrays and creating thematic connections in Quranic concepts as well as creating connections between these concepts and concepts in other religions. The evaluation of the presented methods for ontology automatic learning shows that the combined use of data mining methods and artificial intelligence brings better results. Most of the results of this field are in two general categories. The first category was based on the use of data mining, text mining and machine learning methods to automatically extract three concepts and dimensions (subject-predicate-object) along with Semantic relationships from the text of the Quran. The other category compares the performance of methods and algorithms based on statistics and similarity, such as TF, TF-IDF, AVE-TF, Ridf, TIM, N-gram, FREyA, Pos Taggin, Levenshtein, Log Likelihod, Herset, etc. in extracting concepts for the construction of the Quranic ontologies. The findings of the future studies review show the researchers' interest in artificial intelligence algorithms and their use in ontology learning and the automatic and semi-automatic development of Quranic ontologies. The lack of correct data sets is the reason for the inability of the world's advanced artificial intelligence systems such as GPT 4, which must be addressed in the future.

Discussion and conclusion: The results of this study can help to direct future research about the best practices in the automatic development of Quranic ontologies. This issue can be taken into consideration by designing a comprehensive Quranic ontology that covers all topics and concepts according to the context of the Quran, and by creating a comprehensive ontology of the Quranic concepts, it will guide users towards the retrieval of Quranic knowledge. Also, more use of artificial intelligence and natural language processing methods, such as GPT as a machine learning model for natural language text generation by deep neural network, it seems essential in the development of automatic learning of Quranic ontologies. Machine learning requires the existence of big data in the field of the Qur'an, hence the creation of standard data sets is one of the future studies.

Keywords: Ontology automatic learning, Quran ontologies, Semantic technology, Knowledge extraction, Data mining.

پژوهش‌نامه کتابداری و اطلاع رسانی
مقاله پژوهشی https://infosci.um.ac.ir دسترسی آزاد

روش‌های یادگیری خودکار هستی‌نگاشت‌ها در حوزهٔ مفاهیم قرآنی: مطالعهٔ مروری دامنه‌ای

علی میرعرب
استادیار گروه اشاعهٔ اطلاعات و تبادل دانش، پژوهشگاه علوم و فرهنگ اسلامی، قم، ایران، alimirarab@isca.ac.ir
مرتضی محمدی استانی
استادیار گروه علم اطلاعات و دانش‌شناسی، دانشگاه اصفهان، اصفهان، ایران (نویسنده مسئول)، m.ostani@edu.ui.ac.ir
فائزه‌سادات طباطبایی امیری
دانشجوی دکتری علم اطلاعات و دانش‌شناسی، دانشگاه شهید چمران اهواز، اهواز، ایران، faeze.tabatabai@gmail.com
سمیه دهقانی سانیج
دانشجوی دکتری کتابداری و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی ایران، تهران، ایران، so.dehghani93@gmail.com

تاریخ دریافت: ۱۴۰۲/۰۴/۱۰ تاریخ بازنگری: ۱۴۰۲/۰۶/۱۶ تاریخ پذیرش: ۱۴۰۲/۰۸/۲۰ تاریخ انتشار: ۱۴۰۲/۰۹/۳۰

استناد: میرعرب، علی؛ محمدی استانی، مرتضی؛ طباطبایی امیری، فائزه سادات؛ دهقانی سانیج، سمیه. (۱۴۰۲). روش‌های یادگیری خودکار هستی‌نگاشت‌ها در حوزهٔ مفاهیم قرآنی: مطالعهٔ مروری دامنه‌ای. پژوهش‌نامه کتابداری و اطلاع رسانی، ۱۳(۲)، ۴۹-۲۷. doi:10.22067/infosci.2023.83083.1169

چکیده

مقدمه: امروزه فناوری‌های معنایی رویکرد جدیدی را در پردازش و بازنمون معارف قرآنی با هدف ارائهٔ اطلاعات معنادار ارائه می‌دهند. هستی‌نگاشت‌ها به‌عنوان یکی از فناوری‌های معنایی، ابزاری جهت بیان رسمی مفاهیم و روابط موجود در حوزهٔ خاصی بوده که توسعه و کاربرد آن جهت استخراج معارف و علوم قرآنی مورد توجه قرار گرفته است. یادگیری هستی‌نگاشت‌ها و روش‌های آن به‌صورت خودکار جهت استخراج مفاهیم از مباحث مهم در حوزهٔ وب معنایی و فناوری‌های آن است. به‌تازگی توسعه و کاربرد یادگیری هستی‌نگاشت‌ها جهت استخراج مفاهیم قرآنی مورد توجه قرار گرفته است. ازاین‌رو، هدف پژوهش حاضر، بررسی جامع یادگیری خودکار هستی‌نگاشت‌ها در حوزهٔ استخراج مفاهیم قرآنی به‌منظور شفاف‌سازی وضعیت فعلی و آینده است. معیارهای مورد بررسی مجموعه داده‌ها، روش‌های یادگیری، روش‌های ارزیابی، نتایج و پیشنهادهای آتی پژوهش‌ها در حوزهٔ یادگیری خودکار هستی‌نگاشت‌های قرآنی بود.

روش‌شناسی: روش بررسی پژوهش حاضر، مرور دامنه‌ای بر اساس دستورالعمل‌های پریزما و بر اساس رویهٔ استفاده‌شده توسط آرکسی و امالی (۲۰۰۵) است. این فرآیند پروتکلی را به‌منظور تطبیق نتایج پژوهش موجود با سؤالات و معیارهای تحقیق توصیف می‌کند. پنج مرحلهٔ پیشنهادی آرکسی و امالی عبارت‌اند از: ۱. شناسایی و طراحی سؤال(ها) پژوهش، ۲. انجام استراتژی‌های جستجو برای استخراج مطالعات مرتبط از طریق انتخاب واژه‌های کلیدی مناسب و عملگرهای بولی، ۳. انتخاب نهایی پژوهش‌های مرتبط با تعیین معیارهای ورود و خروج، ۴. خلاصه‌سازی و گزارش یافته‌ها و درنهایت، ۵. گزارش و بحث پیرامون نتایج حاصل. جستجوی منابع در هفت پایگاه دادهٔ علمی مشتمل بر Emerald, Science Direct, IEEE Xplore Digital Library, Google Scholar, Web of Science, Scopus انجام شد. فرایند جستجو در فروردین ۱۴۰۲ صورت گرفت. تعداد ۸۱۱ مقاله، بدون توجه

به محدودهٔ زمانی، مورد ارزیابی و انتخاب قرار گرفت. به‌منظور سازماندهی مقالات بازیابی‌شده، از نرم‌افزار مدیریت منابع اطلاعاتی اندنوت استفاده شد و پس از تطبیق عناوین در پایگاه‌های اطلاعاتی مختلف، تعداد ۳۱۷ مقاله تکراری حذف گردید. پس از بررسی چکیده‌ها، معیارهای ورود و خروج و کیفیت مقالات اعمال گردید. همچنین به‌منظور جلوگیری از سوگیری در انتخاب مقالات، طی بررسی تصادفی مجددی، توسط دو پژوهشگر مستقل در حوزهٔ یادگیری خودکار هستی‌نگاشت نیز ارزیابی صورت گرفت و درنهایت تعداد ۲۵ اثر به‌عنوان ملاک مرور انتخاب گردید.

یافته‌ها: یافته‌ها نشان داد اغلب پژوهش‌ها در حوزهٔ مجموعهٔ داده‌های قرآنی به زبان‌های انگلیسی و عربی بودند و بخش عمده آن‌ها نیز از ترجمهٔ انگلیسی قرآن الهلالی و خان استفاده کرده‌اند. استفاده از مجموعه داده‌های بسیار محدود، مهم‌ترین محدودیت پژوهش‌های انجام شده بود. بخش عمدهٔ پژوهش‌ها از روش‌های نرمال‌سازی، خوشه‌بندی و دسته‌بندی متن، خلاصه‌سازی متن، استخراج اطلاعات، تشابه و یافتن موجودیت‌های نامدار استفاده کرده‌اند. البته در برخی پژوهش‌ها، روش‌های هوش مصنوعی نظیر شبکهٔ عصبی نیز به کار گرفته شده است. علاوه بر این، یافته‌ها نشان داد که الگوریتم‌های داده‌کاوی مبتنی بر روش‌های آمار و احتمال برای یادگیری و ساخت هستی‌نگاشت‌های خودکار در میان محققان با محبوبیت روبرو شده است. همچنین از روش‌های محاسبهٔ دقت، فراخوانی و معیار F برای ارزیابی نتایج کاربرد الگوریتم‌های یادگیری خودکار در هستی‌نگاشت‌های قرآنی استفاده کرده‌اند. پژوهش‌هایی که از روش‌های هوش مصنوعی بهره‌برداری کرده‌اند، با تحلیل معنایی، استنتاج، مدل‌سازی و تأیید اعتبار داده‌های استنتاج‌شده به نتایجی مانند تشخیص صوت برای آموزش قرائت قرآن، تشخیص آرایه‌های ادبی و ایجاد ارتباط‌های موضوعی در مفاهیم قرآنی و همچنین ایجاد ارتباط بین این مفاهیم با مفاهیم سایر ادیان نائل شده‌اند. ارزیابی روش‌های ارائه‌شده برای یادگیری خودکار هستی‌نگاشت‌های قرآنی نشان می‌دهد استفاده توأمان از روش‌های داده‌کاوی و هوش مصنوعی نتایج بهتری را به‌همراه دارد. بخش عمدهٔ نتایج این حوزه در دو دسته کلی قرار دارد. دستهٔ اول مبتنی بر به‌کارگیری روش‌های داده‌کاوی، متن‌کاوی و یادگیری ماشین جهت استخراج خودکار مفاهیم و ابعاد سه‌گانه (فعل، فاعل، مفعول) به‌همراه روابط معنایی از متن قرآن بود. دستهٔ دیگر به مقایسه عملکرد روش‌ها و الگوریتم‌های مبتنی بر آمار و مشابهت‌یابی نظیر TF، TF-IDF، AVE-TF، Ridf، TIM، N-gram، FREyA، Pos Taggin، Levenshtein، Log Likelihod، هِرست، و جز این‌ها در استخراج مفاهیم خودکار جهت ساخت هستی‌نگاشت قرآنی پرداخته‌اند. یافته‌های حاصل از بررسی کارهای آینده نشان از علاقهٔ محققان به الگوریتم‌های هوش مصنوعی و استفاده در یادگیری هستی‌نگاشت و توسعهٔ خودکار و نیمه‌خودکار هستی‌نگاشت‌های قرآنی دارد. فقدان مجموعه داده‌های صحیح، دلیل عجز سامانه‌های هوش مصنوعی پیشرفتهٔ دنیا مانند جی‌پی‌تی ۴ است که در آینده باید به این مهم پرداخته شود.

نتیجه‌گیری: نتایج این مطالعه می‌تواند به جهت‌دهی پژوهش‌های آتی درباره بهترین روش‌ها در توسعهٔ خودکار هستی‌نگاشت‌های قرآنی کمک کند. این مسئله می‌تواند با طراحی هستی‌نگاشت جامع قرآنی که تمام موضوعات و مفاهیم را با توجه به بافت قرآن، پوشش دهد، مدنظر قرار گرفته و با ایجاد هستی‌نگاشتی جامع از مفاهیم قرآن، کاربران را به‌سمت بازیابی دانش قرآنی رهنمون سازد. همچنین بهره‌برداری بیشتر از روش‌های هوش مصنوعی و پردازش زبان طبیعی نظیر جی.پی.تی. به‌عنوان مدل یادگیری ماشینی برای تولید متن به زبان طبیعی با استفاده از شبکهٔ عصبی عمیق، در توسعهٔ خودکار هستی‌نگاشت‌های قرآنی ضروری به نظر می‌رسد. با توجه به اینکه یادگیری ماشین مستلزم وجود داده‌های کلان در حوزهٔ قرآن است، ساخت مجموعه داده‌های استاندارد ازجمله کارهای آتی محققان است.

کلیدواژه‌ها: یادگیری خودکار هستی‌نگاشت، هستی‌نگاشت قرآن، فناوری معنایی، استخراج دانش، داده‌کاوی.

مقدمه

امروزه روش‌های حوزهٔ فناوری معنایی، رویکردهای نوینی را در توصیف و پردازش مفاهیم قرآنی با هدف ارائه اطلاعات معنادار ارائه می‌دهند. دراین‌بین، هستی‌نگاشت‌ها[1] در بازنمون دانش پیچیدهٔ مفاهیم قرآن، ابزار مناسبی محسوب می‌شوند (Saad, Noah, Salim, & Zainal, 2013). ازاین‌رو، ایجاد هستی‌نگاشت‌ها برای محتوای قرآنی، نه‌تنها مفید، بلکه ضروری می‌نماید. هستی‌نگاشت‌ها ابزار بیان رسمی مفاهیم و روابط موجود در قلمروی خاصی هستند (Sánchez & Moreno, 2008). گروبر، هستی‌نگاشت را توصیف یا بینشی صریح از جهان به‌همراه تشریح مفاهیم و روابط میان آن‌ها می‌داند (Gruber, 1993). هستی‌نگاشت‌ها کاربردهای وسیعی در حوزه‌های مختلف از قبیل وب معنایی، موتورهای کاوش، طراحی پایگاه‌های داده، استخراج و بازیابی اطلاعات، کتابخانه‌های دیجیتالی، مهندسی دانش و جز این‌ها دارند. یکی از حوزه‌های کاربرد هستی‌نگاشت‌ها، پردازش زبان طبیعی شامل درک عمیق متون و استخراج داده و اطلاعات از متون به‌منظور ایجاد پایگاه دانش متنی است (Castro-Schez, Jennings, Luo, & Shadbolt, 2004؛ Tomás & García, 2005).

در جهت ایجاد هستی‌نگاشت و تحقق کاربردهای آن، حوزه‌ای پژوهشی به نام مهندسی هستی‌نگاشت شکل گرفته است که در آن، با واکاوی مجموعه روش‌های مربوط به فرآیند توسعهٔ هستی‌نگاشت، بر راهکارهای مناسب جهت ایجاد، چرخهٔ حیات، یکپارچه‌سازی، قواعد، روش‌ها، یادگیری، استفاده و ارزیابی هستی‌نگاشت، تمرکز شده است (Corcho, Fernandez-Lopez, & Gomez-Perez, 2007). دراین‌بین، به‌منظور مقابله با چالش اکتساب و استخراج دانش در روند توسعهٔ هستی‌نگاشت و نیز زمان‌بری ایجاد آن، ساخت خودکار هستی‌نگاشت‌ها پیشنهاد شده است (Shamsfard & Abdollahzadeh Barforoush, 2003). برخی از رویکردهایی که می‌توان جهت استخراج دانش از متون زبان طبیعی مورد توجه قرار گیرد عبارت‌اند از: رویکرد استخراج اطلاعات (Etzioni, Fader, Christensen, Soderland, & Mausam, 2011)، رویکرد استخراج اطلاعات مبتنی بر هستی‌نگاشت (Wimalasuriya & Dou, 2010)، حاشیه‌نویسی معنایی (Uren, Cimiano, Iria, Handschuh, Vargas-Vera, Motta, & Ciravegna, 2006) و یادگیری هستی‌نگاشت (Maedche & Staab, 2001).

یادگیری هستی‌نگاشت به‌معنای فرآیندی جهت ساخت خودکار و نیمه‌خودکار هستی‌نگاشت (Gulla, Borch, & Ingvaldsen, 2007)، روشی است که جهت استخراج کارآمد دانش و ساخت هستی‌نگاشت به کار می‌رود. یادگیری هستی‌نگاشت سعی بر استخراج مفاهیم و روابط از متون دارد. در

۱. اصطلاحات «آنتولوژی» و «هستی‌شناسی» در علوم رایانه و فلسفه رایج است و در متون این حوزه از اصطلاحات دیگری نظیر «هستی‌نگار، هستان‌نگار و هستان‌شناسی» استفاده می‌شود، اما به‌زعم نگارندگان در بافت میراث فرهنگی (کتابخانه، آرشیو و موزه) عبارت «هستی‌نگاشت» تناسب بیشتری دارد.

سال‌های اخیر به‌موازات گسترش وب معنایی و درخواست فزاینده به برقراری روابط معنایی میان موجودیت‌های داده‌ای وب و نقش هستی‌نگاشت در تحقق این امر از یک‌سو، و سخت و زمان‌بر بودن ساخت دستی هستی‌نگاشت‌ها از سوی دیگر، توجه بسیاری از پژوهش‌ها به‌سمت یادگیری خودکار هستی‌نگاشت‌ها معطوف شده است. این پژوهش‌ها سعی دارند تا با ارائهٔ روش‌ها و الگوریتم‌هایی، این فرآیند را تسهیل کنند.

یکی از حوزه‌هایی که می‌توان از روش‌های یادگیری خودکار هستی‌نگاشت‌ها در استخراج اطلاعات از آن بهره برد، حوزهٔ معارف و مفاهیم قرآنی است. بااین‌حال، تفاوت قرآن با ادبیات بشر و توجه به ویژگی خاص آن در عرضهٔ اطلاعات گسترده در آیاتِ پراکنده و درعین‌حال مرتبط با یکدیگر ازلحاظ مفهومی، موجب شده که استخراج دانش از قرآن، کاری چالش‌برانگیز قلمداد شود. بسیاری از پژوهش‌ها در حوزهٔ ادبی و بلاغی قرآن نیز مؤید همین نکته است. بنابراین، چیدمان منحصربه‌فرد قرآن، دستیابی مستقیم به‌معنای آن را که ازنظر زبانی، غنی و ازلحاظ معنایی، چندلایه بوده، دشوار نموده است (Saad et al., 2013).

بررسی متون درزمینهٔ استخراج مفاهیم و دانش از قرآن حاکی از انجام پژوهش‌های متعددی است. هر چند نتایج پژوهش‌ها، منتج به طراحی و ارائهٔ روش‌های مناسب یادگیری خودکار هستی‌نگاشت مبتنی بر پردازش زبان طبیعی و آماری در حوزه‌های مختلف ازجمله قرآن شده است؛ اما هر یک از آن‌ها دارای کاستی‌ها و ضعف‌هایی نیز هستند. به‌منظور بهبود و ارتقای عملکرد این حوزه، انجام پژوهش‌های بیشتر در این زمینه، ضروری به نظر می‌رسد. بنابراین در راستای تحقق این امر، پیش‌ازاین بایسته است تصویری روشن از فعالیت‌ها و پیشرفت‌های صورت‌گرفته در حوزهٔ طراحی و ارائهٔ روش‌های مناسب استخراج دانش و یادگیری خودکار هستی‌نگاشت‌ها در حوزهٔ مفاهیم قرآن ارائه گردد. ازاین‌رو هدف پژوهش حاضر، بررسی و واکاوی الگوهایی که در پژوهش‌های پیشین درزمینهٔ استخراج مفاهیم قرآنی نظیر مجموعه داده، روش‌های توسعهٔ یادگیری خودکار، ارزیابی هستی‌نگاشت‌ها، نتایج و پیشنهادهای آن‌ها با استفاده از روش مرور دامنه‌ای[1] است. پژوهش حاضر، با بررسی کلی، دیدی جامع از این حوزه، که تاکنون پژوهشی مستقل به آن نپرداخته است، در اختیار پژوهشگران قرار می‌دهد و با شناسایی نتایج و پیشنهادهای آتیِ ارائه‌شده، زمینه را برای بهبود و ارتقای الگوریتم‌های مورد استفاده در استخراج مفاهیم قرآنی و یادگیری خودکار هستی‌نگاشت‌ها، مهیا می‌سازد.

روش‌ها

پژوهش حاضر، مروری نظام‌مند از نوع مرور دامنه‌ای بوده که سعی دارد الگوهایی که در پژوهش‌های پیشین درزمینهٔ استخراج مفاهیم قرآنی، معرفی و به کار گرفته شده‌اند را مورد توجه قرار دهد. در مرور

1. scoping review

نظام‌مند، بعد از شناسایی، انتخاب و ارزیابی پژوهش‌های صورت‌گرفته در حیطه‌ای موضوعی با استفاده از روشی نظام‌مند، یافته‌ها و نتایج مستخرج از پژوهش‌ها، مورد استفاده قرار می‌گیرد (Shannon, 2002). مطالعات مرور نظام‌مند اغلب روی سؤال تعریف‌شده‌ای بوده و بر اساس آن، مطالعات مناسب شناسایی و جستجو شده؛ درحالی‌که در مطالعات مرور دامنه‌ای، تمایل به در نظر گرفتن موضوعات گسترده‌تری است. به‌عبارت‌دیگر در این نوع مطالعات، حول موضوعی تمام حواشی آن در نظر گرفته می‌شود (Jalali & Mohammadi, 2021).

مرور دامنه‌ای یکی از انواع مرورها است که چهارچوب روش‌شناختی پنج‌گانهٔ آن توسط آرکسی و امالی[1] (۲۰۰۵) ارائه گردیده است. این فرآیند پروتکلی را به‌منظور تطبیق نتایج مطالعات و پژوهش‌های موجود با پرسش‌ها و معیارهای پژوهش، توصیف می‌نماید. پنج گام پیشنهادی عبارت است از: «۱. شناسایی و طراحی پرسش(های) پژوهش، ۲. فرآیند جستجو جهت شناسایی و استخراج مطالعات و پژوهش‌های مرتبط از طریق انتخاب کلمات کلیدی مناسب، ۳. انتخاب مطالعات و پژوهش‌های مرتبط به‌واسطهٔ تعیین معیارهای ورود و خروج، ۴. جدول‌بندی، جمع‌بندی و خلاصه‌سازی داده‌ها، و درنهایت، ۵. گزارش نتایج».

سؤالات پژوهش

پژوهش حاضر، مطالعات و پژوهش‌هایی را که در حوزهٔ طراحی و ارائهٔ روش‌ها و الگوهای مناسب جهت استخراج مفاهیم قرآنی و یادگیری هستی‌نگاشت، به رشته تحریر درآمده‌اند را مورد بررسی قرار داده است. اولین گام در پروتکل آرکسی و امالی شناسایی و طراحی پرسش‌های پژوهش است. با توجه به مطالب پیش‌گفته و هدف پژوهش، سؤالات ذیل مدنظر قرار گرفت:

۱) چه نوع مجموعه داده‌ای به‌منظور یادگیری خودکار هستی‌نگاشت‌های قرآنی استفاده شده است؟
۲) از چه روش‌هایی به‌منظور یادگیری خودکار هستی‌نگاشت‌های قرآنی استفاده شده است؟

۴) نتایج پژوهش‌های حوزهٔ یادگیری خودکار هستی‌نگاشت‌های قرآنی چیست؟
۵) پیشنهادهای آتی پژوهش‌های حوزهٔ یادگیری خودکار هستی‌نگاشت‌های قرآنی چیست؟

فرآیند جستجو

این فرآیند جهت شناسایی و استخراج پژوهش‌های مرتبط از طریق انتخاب کلمات کلیدی مناسب و درنهایت، شناسایی هرگونه مطالعهٔ مرتبط بالقوه بر اساس پرسش‌های پژوهش، انجام می‌شود. فرآیند

1. Arksey & O'Malley

جستجو شامل چهار مرحلهٔ کامل به شرح شکل (۱) است.

شکل ۱. فرایند جستجو

مقالات از پایگاه‌های اطلاعاتی بین‌المللی علمی که قابلیت جستجوی پیشرفته و کنترل کیفیت نتایج دارند، مشتمل بر Emerald, Science Direct, Google Scholar, IEEE Xplore Digital Library, Web of Science و Scopus بازیابی شده‌اند. همچنین به‌منظور جامعیت و غنی‌سازی منابع از بخش منابع و مآخذ مقالات بازیابی‌شده نیز استفاده شد. در این پژوهش، از راهبرد جستجوی زیر در پایگاه‌های اطلاعاتی منتخب استفاده گردید.

(Quran OR Al-Quran OR “Al- Quran” OR Al-Qur'an OR “Al- Qur'an” OR Koran) AND (ontolog* OR thesaur*)

فرایند جستجو در فروردین ۱۴۰۲ صورت گرفت. نتایج جستجو، بدون توجه به محدودهٔ زمانی، مورد ارزیابی و انتخاب قرار گرفت. در این مرحله تعداد ۸۱۱ مقاله، شناسایی شد. به‌منظور سازماندهی مقالات بازیابی‌شده، از نرم‌افزار مدیریت منابع اطلاعاتی اندنوت استفاده شد. با استفاده از این نرم‌افزار و تطبیق عنوان، تعداد ۳۱۷ مقاله تکراری حذف گردید.

تعیین معیارهای ورود و خروج

مقالات اولیه بر اساس معیارهای ورود و خروج انتخاب شده‌اند. این معیارها برای دستیابی به الزامات مقالات اولیه با تأکید بر پرسش‌های پژوهش و کیفیت مقالات، تدوین و بیان شده‌اند. نقش این معیارها در تأیید مقالات اولیه و رد هرگونه مقالات نامرتبط، ضروری است. جدول (۱)، معیارهای ورود و خروج را نمایش می‌دهد. معیارهای ورود، انتخاب مقالات اولیه را تأیید می‌کند. با بررسی چکیده، مقالاتی که معیارهای خروج از جدول را دارا بودند، به تعداد ۳۶۲ مقاله از مطالعه حذف شدند.

جدول ۱. معیارهای ورود و خروج مقالات حوزه یادگیری هستی‌نگاشت‌های قرآنی

ردیف	معیارهای ورود	معیارهای خروج
۱	ارائهٔ الگویی جهت استخراج دانش و یادگیری هستی‌نگاشت در حوزه قرآن به‌عنوان موضوع اصلی	ارائه نکردن الگویی جهت استخراج دانش از قرآن
۲	مقالات منتشره در مجلات و کنفرانس‌های بین‌المللی	کتاب، فصلی از کتاب، آثار خاکستری، و گزارش‌های کوتاه

۳	دربرگیرندهٔ گزارشی کامل و جامع از فرآیند کار	فقدان اجرای عملیاتی الگوی ارائه‌شده یا فقدان دسترسی به فایل تمام متن

استخراج داده‌ها و ارزیابی کیفیت

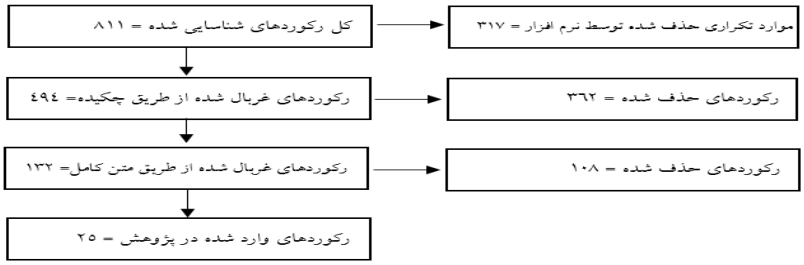
مطابق جدول (۲)، معیارهای ارزیابی کیفیت پژوهش‌ها به‌منظور تحلیل جامعیت بافت پژوهش‌های مورد مطالعه و حصول اطمینان از کیفیت مقالات اولیه مطابق با پرسش‌های پژوهش، تدوین گردید. در تدوین معیارهای ارزیابی و کیفیت، از پژوهش‌های کیتچنهام[1] (۲۰۰۴) و آزنی، احمد، نوح، هزوانی و حیاتی[2] (۲۰۱۵) در مرور ادبیات پژوهش الهام گرفته شد. به‌منظور جلوگیری از سوگیری در انتخاب مقالات، طی بررسی تصادفی مجدد، توسط دو پژوهشگر مستقل در حوزهٔ یادگیری خودکار هستی‌نگاشت نیز ارزیابی صورت گرفت. نتایج ارزیابی کیفیت، مقایسه و موارد متناقض جهت حصول اجماع، مورد بحث قرار گرفت. پس از مطالعهٔ متن کامل مقالات و اعمال معیارهای ارزیابی کیفیت (جدول ۲)، تعداد ۱۰۷ مقالهٔ دیگر از پژوهش حذف گردید. بعد از ارزیابی دقیق پژوهش‌ها، تعداد ۲۵ مقاله نهایی گردید.

جدول ۲. معیارهای ارزیابی کیفیت مقالات حوزهٔ یادگیری خودکار هستی‌نگاشت‌های قرآنی

ردیف	معیارها
۱	بیان صریح اهداف تحقیق
۲	ارائهٔ روشمند و توصیف رویکرد به‌کاررفته در طراحی و ارائهٔ روش‌های مناسب استخراج مفاهیم قرآنی و یادگیری خودکار هستی‌نگاشت
۳	تبیین کامل نتایج ارائه‌شده
۴	امکان استخراج مستقیم اطلاعات و داده‌های مورد نیاز

به‌طور کلی، شکل (۲) الگوریتم جستجو بر اساس فلوچارت پریزما[3]، بازیابی و اعمال معیارهای ورود و خروج از مطالعه و تعداد مقالاتی که در هر مرحله باقی ماندند را به تصویر کشیده است.

1. Kitchenham
2. Azni, Ahmad, Noh, Hazwani & Hayaati
3. PRISMA



شکل ۲. الگوریتم جستجو بر اساس فلوچارت پریزما

یافته‌ها

پرسش اول: چه نوع مجموعه داده‌ای به‌منظور یادگیری خودکار هستی‌نگاشت‌های قرآنی استفاده شده است؟

مرور منابع این حوزه نشان داد که اغلب پژوهش‌ها در حوزهٔ مجموعهٔ داده‌های قرآنی به زبان‌های انگلیسی و عربی بودند. پژوهش‌هایی نظیر خازانی[1] و همکاران (۲۰۲۱) و کدیر و یائوری[2] (۲۰۱۷) با تمرکز بر ترجمهٔ قرآن به زبان انگلیسی انجام شده، و بیشتر نیز از ترجمهٔ انگلیسی قرآن الهلالی و خان استفاده کرده‌اند.

چندین پژوهش بر بخشی از سوره‌ها نظیر اسماعیل، عبدالرحمن و ابوبکر[3] (۲۰۱۸)، اسماعیل و همکاران (۲۰۱۷) و اسماعیل و همکاران (۲۰۱۶) و برخی نیز به‌صورت موضوعی نظیر سد و همکاران (۲۰۱۳) متمرکز بودند.

گروه دوم، شامل پژوهش‌هایی نظیر الرحیلی، القحطانی، و آتول[4] (۲۰۱۸) است که بر روی مجموعه‌های زبان عربی انجام شده‌اند. در این گروه استفاده هم‌زمان از زبان عربی کلاسیک و ترجمهٔ قرآن به زبان عربی مدرن و یا هستی‌نگاشت‌های موجود به زبان عربی مدنظر بوده است.

در این میان، تعدادی از پژوهش‌ها نظیر اوتومو، سوریانا و سانوسی عظمی[5] (۲۰۱۹ و ۲۰۲۰) از ترجمه، تفاسیر و داده‌های علوم قرآنی به زبان اندونزیایی برای توسعهٔ روش‌های خودکار یادگیری در این حوزه بهره‌برداری کرده‌اند. استفاده از وردنت به زبان انگلیسی و اندونزیایی نیز در چند پژوهش به کار رفته است. یافته‌های این بخش به‌تفصیل در قالب جدول (۳) ذکر شده است.

جدول ۳. مجموعه داده‌های به‌کاررفته به‌منظور یادگیری خودکار هستی‌نگاشت‌های قرآنی

1. Khazani
2. Kadir & Yauri
3. Ismail, Abd Rahman & Abu Bakar
4. Alrehaili, Alqahtani & Atwell
5. Utomo, Suryana & Sanusi Azmi

مجموعه داده	نویسنده/نویسندگان (سال)
آیات مرتبط با «حج» از ترجمهٔ انگلیسی الهلالی و خان	Ismail, Abd Rahman & Abu Bakar (2018); Ismail, Abd Rahman, Abu Bakar & Makhtar (2018); Ismail, Abd Rahman & Abu Bakar (2017); Ismail, Abd Rahman & Abu Bakar (2016); Ismail, Abu Bakar & Abd Rahman (2015)
۱۰۰ آیه از مجموعه دادهٔ ترجمهٔ انگلیسی یوسف علی و الهلالی و خان	Weaam & Saad (2016)
ترجمهٔ انگلیسی قرآن	Khazani, Mohamed, Sembok, Yusop, Wani, Gulzar, Halip, Marzukhi & Yunos (2021); Petiwala & Sathya (2011)
استفاده از ۳۰۰ مفهوم و ۱۴۷۵ فعل از ترجمهٔ انگلیسی قرآن و وُردنِت	Kadir & Yauri (2017)
ترجمهٔ انگلیسی آیات نماز	Saad et al. (2013)
ترجمهٔ مالایی از سورهٔ ضحی تا سورهٔ ناس	Husin, Saad & Noah (2017)
ترجمه و تفسیر اندونزیایی قرآن و وُردنِت اندونزیایی	Utomo, Suryana & Sanusi Azmi (2020); Utomo, Suryana & Sanusi Azmi (2019)
متن عربی قرآن (کلاسیک و مدرن) و احادیث	Saber, Abdel-Galil & El-Fatah Belal (2022); Alshammeri, Atwell & Alsalka (2021); Elkhammash & Abdessalem (2019); Al-Zamil & Al-Radaideh (2014); Sharaf & Atwell (2012); Baqai, Basharat, Khalid, Hassan & Zafar (2009)
ترجمهٔ پیکرهٔ القرآن	Irwanto, Bijaksana & Adiwijaya (2018)
ترجمهٔ پیکرهٔ عربی نگارش‌شده به‌صورت Buckwalter	Bentrcia, Zidat & Marir (2018)
هستی‌نگاشت Qurany و پیکرهٔ قرآنی عربی	Alrehaili, Alqahtani & Atwell (2018)
اسامی عربی مرتبط با حوزه معنایی «زمان»	Al-Khalifa, Al-Yahya, Bahanshal, Al-Odah & Al-Helwah (2010)
اصطلاحات منتخب از هستی‌نگاشت موضوعات مصحف التجوید و قرائات هفت‌گانهٔ قرآن	Mahmoud & Hassan (2013)
متون مقدس	Varghese & Punithavalli (2019)

پرسش دوم: از چه روش‌هایی به‌منظور یادگیری خودکار هستی‌نگاشت‌های قرآنی استفاده شده است؟

مرور و بررسی یافته‌های مرتبط با روش‌های یادگیری خودکار هستی‌نگاشت‌های قرآنی نشان داد که هرچند بسیاری از پژوهش‌ها نظیر الزمیل و الردیده[1] (۲۰۱۴) با استفاده از الگوریتم‌های داده‌کاوی نظیر خوشه‌بندی و طبقه‌بندی، برچسب‌گذاری مفهومی، روش‌های آماری و مدل‌های بُرداری، انتخاب ویژگی‌ها و تجزیهٔ زبان، به استخراج و کشف الگوها و پیش‌بینی روابط بین مفاهیم قرآنی پرداخته‌اند، اما تنها تعداد معدودی از آن‌ها، گام را فراتر گذاشته و مبادرت به ادغام الگوریتم‌های مناسب برای داده‌کاوی متون با روش‌های پردازش زبان طبیعی کرده‌اند (Husin, Saad, & Noah, 2017). به‌طور کلی می‌توان روش‌های یادگیری خودکار هستی‌نگاشت‌ها در حوزهٔ مفاهیم قرآنی را به چند دستهٔ اصلی به شرح ذیل

1. Al Zamil & Al Radaideh

تقسیم‌بندی نمود:

۱) روش‌های نرمال‌سازی و طبقه‌بندی داده‌ها مشتمل بر استفاده از الگوریتم‌های طبقه‌بندی و خوشه‌بندی نظیر الگوریتم نزدیک‌ترین همسایه، ماشین‌بردار پشتیبان، انتخاب ویژگی‌ها، معیارهای شباهت چندگانه (نظیر وارگس و پونیثاوالی[1] (۲۰۱۹) و ویام و سد[2] (۲۰۱۶)).

۲) روش‌های خلاصه‌سازی مشتمل بر تکنیک‌های آماری، برآورد درست‌نمایی بیشینه و روش TF-IDF، سنجش میزان ربط بین عبارات اسمی با استفاده از ضریب همبستگی پیرسون، و روش بهبودیافتهٔ مبتنی بر مسیر[3] با استفاده از ارزش درجهٔ ارتباط (نظیر الخاماش و عبدالسالم[4] (۲۰۱۸) و ویام و سد (۲۰۱۶)).

۳) روش‌های استخراج الگوی نحوی‌واژگانی (زبان‌شناسی) مشتمل بر پردازش زبان طبیعی، پیش‌بینی شباهت معنایی، پیش‌پردازش خودکار، الگوریتم هرست[5]، روش یادگیری هم‌آیندها، **n**-گرام و محاسبات هم‌رخدادی، روش‌های مشابهت‌یابی، ابزار GATE به‌منظور برچسب‌گذاری کلمات[6]، ریشه‌یابی، ریخت‌شناسی، روش کیسه‌کلمات[7] و بُن‌واژه‌سازی، استخراج مشابهت‌ها، تحلیل معنایی، شبکهٔ عصبی[8]، قواعد معنایی، استخراج و استنتاج از سه‌گانه‌های معنایی[9]، یافتن موجودیت‌های نامدار، و تشخیص صوت (نظیر خازنی و همکاران (۲۰۲۱) و کدیر و یائوری (۲۰۱۷)).

برخی از پژوهش‌ها نظیر اوتومو و همکاران (۲۰۲۰) و اسماعیل و همکاران (۲۰۱۸) از دو روش جهت یادگیری خودکار هستی‌نگاشت‌ها استفاده کرده‌اند. نتایج این مرور در جدول (۴) ذکر شده است. همان‌طور که مشاهده می‌شود استفاده از الگوهای نحوی‌واژگانی یا به‌نوعی الگوهای مبتنی بر پردازش زبان طبیعی دارای تعداد و تنوع روشی بیشتری است.

جدول ۴. روش‌های یادگیری خودکار هستی‌نگاشت‌های قرآنی

روش	نویسنده/نویسندگان (سال)
نرمال‌سازی و طبقه‌بندی داده	Varghese & Punithavalli (2019); Weaam & Saad (2016)
خلاصه‌سازی	Utomo, Suryana & Sanusi Azmi (2020); Elkhammash & Abdessalem (2019); Bentrcia, Zidat & Marir (2018); Irwanto, Bijaksana & Adiwijaya (2018); Ismail et al. (2018); Ismail, Abd Rahman & Abu Bakar (2018); Ismail, Abd Rahman & Abu Bakar (2017); Weaam & Saad (2016)

1. Varghese & Punithavalli
2. Weaam & Saad
3. Path-based Integrated Method
4. Elkhammash & Abdessalem
5. Herst Algorithm
6. Part of Speech (POS) taggers
7. Bag of Words (BoW) Technique
8. neural network

۳. به‌کارگیری مدلی مبتنی بر قانون (rule-based) برای استخراج خودکار ابعاد سه‌گانهٔ جمله شامل فعل، فاعل و مفعول (گزاره) از متن عربی.

استخراج الگوهای نحوی‌واژگانی (زبان‌شناسی)	Saber, Abdel-Galil & El-Fatah Belal (2022); Alshammeri, Atwell & Alsalka (2021); Khazani et al. (2021); Utomo, Suryana & Sanusi Azmi (2020); Utomo, Suryana & Sanusi Azmi (2019); Alrehaili, Alqahtani & Atwell (2018); Ismail et al. (2018); Husin, Saad & Noah (2017); Ismail, Abd Rahman & Abu Bakar (2017); Kadir & Yauri (2017); Ismail, Abd Rahman & Abu Bakar (2016); Ismail, Abu Bakar & Abd Rahman (2015); Mahmoud & Hassan (2013); Al-Zamil & Al-Radaideh (2014); Saad et al. (2013); Sharaf & Atwell (2012); Petiwala & Sathya (2011); Al-Khalifa et al. (2010); Baqai et al. (2009)

پرسش سوم: چه روش‌هایی به‌منظور ارزیابی روش‌های یادگیری خودکار در هستی‌نگاشت‌های قرآنی استفاده شده است؟

روش‌های ارزیابی در اغلب پژوهش‌هایی که با تکنیک‌های داده‌کاوی صورت گرفته مشتمل بر محاسبهٔ فراخوانی، دقت، صحت، معیار F به‌عنوان ترکیبی از سنجه‌های دقت‌وصحت، معیارهای مثبت واقعی، مثبت کاذب و منفی کاذب، مقایسه با استاندارد طلایی، محاسبهٔ مشابهت با استفاده از ابزار SimLex-999 و محاسبهٔ تشابه کسینوسی، اقلیدسی، ژاکارد و تحلیل احساسات است. همچنین یافته‌ها حاکی از اجرای پرس‌وجو به زبان معنایی اسپارکل (Utomo, Suryana, & Sanusi Azmi, 2019)، ارزیابی و اعتبارسنجی توسط عامل انسانی و متخصص موضوعی، که تنها در دو مورد از پژوهش‌ها نظیر الخاماش و عبدالسالم (۲۰۱۹) مشاهده شد. در برخی از پژوهش‌ها نظیر اسماعیل و همکاران (۲۰۱۵)، توسعهٔ سنجه‌های جدید ارزیابی در پژوهش‌های مبتنی بر استفادهٔ هم‌زمان از داده‌کاوی و هوش مصنوعی مشاهده شد.

مرور پیشینه‌ها نشان داد که بخش عمدهٔ روش‌های ارزیابی یادگیری خودکار هستی‌نگاشت‌های قرآنی مبنی بر محاسبهٔ دقت، صحت و معیار صحت و F بوده که روشی مرسوم در این زمینه است. در چندین پژوهش نظیر محمود و حسن[1] (۲۰۱۳) و شرف و آتول[2] (۲۰۱۲) روشی جهت ارزیابی و اندازه‌گیری عملکرد هستی‌نگاشت‌ها ارائه نشده است. روش‌های ارزیابی یادگیری خودکار هستی‌نگاشت‌های قرآنی به‌تفصیل در جدول (۵) ذکر شده است.

جدول ۵. روش‌های ارزیابی یادگیری خودکار هستی‌نگاشت‌های قرآنی

روش‌های ارزیابی یادگیری خودکار هستی‌نگاشت	نویسنده/نویسندگان (سال)
محاسبهٔ دقت، صحت و معیار F	Saber, Abdel-Galil & El-Fatah Belal (2022); Alshammeri, Atwell & Alsalka (2021); Khazani et al. (2021); Utomo, Suryana & Sanusi Azmi (2020); Alrehaili, Alqahtani & Atwell (2018); Bentrcia, Zidat & Marir (2018); Ismail, Abd Rahman & Abu Bakar (2018); Ismail et al. (2018); Kadir & Yauri (2017); Weaam & Saad (2016); Al-Zamil & Al-Radaideh (2014); Saad et al. (2013)
محاسبهٔ تشابه کسینوسی، اقلیدسی، ژاکاد و تحلیل	Varghese & Punithavalli (2019)

1. Mahmoud & Hassan
2. Sharaf & Atwell

احساسات	
اجرای پرس و جوی SPARQL	Utomo, Suryana & Sanusi Azmi (2019)
محاسبهٔ ارزش همبستگی و میزان مشابهت با SimLex-999	Irwanto, Bijaksana & Adiwijaya (2018)
آزمون T و خبرگان (متخصص انسانی)	Elkhammash & Abdessalem (2019); Bentrcia, Zidat & Marir (2018); Husin, Saad & Noah (2017)
مقایسه با سیستم یادگیری هستی‌نگاشت Text2Ont	Ismail et al. (2018)
محاسبه با استفاده از فرمول PoE	Ismail, Abu Bakar & Abd Rahman (2015)
توسعهٔ سنجهٔ جدید پوشش عبارات هستی‌نگاشت	Petiwala & Sathya (2011)
منطق توصیف	Baqai et al. (2009)
معیارهای قابلیت گسترش و حداقل تعهد هستی‌نگاشت	Al-Khalifa et al. (2010)

پرسش چهارم: نتایج پژوهش‌های حوزهٔ یادگیری خودکار هستی‌نگاشت‌های قرآنی چیست؟

یافته‌های حاصل از مرور ادبیات نشان داد که یافته‌های پژوهش‌های حوزهٔ یادگیری خودکار هستی‌نگاشت‌های قرآنی را می‌توان از چندین منظر مورد بررسی قرار داد. از منظر کلی، پژوهش‌ها را می‌توان در سه گروه به شرح ذیل طبقه‌بندی نمود:

۱) پژوهش‌هایی که با انتخاب الگوریتم‌های پیشنهادی، منجر به ساخت پیکرهٔ جدیدی (Alshammeri, Atwell, & Alsalka, 2021) و یا نمونهٔ اولیه‌ای (Husin, Saad, & Noah, 2017) شده است.

۲) پژوهش‌هایی که با اصلاح و بهبود مدل‌ها، الگوها و معیارهای قبلی و ارائه روش پیشنهادی (Irwanto, Bijaksana, & Adiwijaya, 2018)، منجر به گسترش پیکرهٔ موجود شده است.

۳) پژوهش‌هایی که با ساخت خودکار هستی‌نگاشت با استفاده از روش‌های موجود (Ismail, Abd Rahman, & Abu Bakar, 2017)، تنها به مقایسهٔ روش‌های یادگیری قبلی اکتفا کرده‌اند.

همچنین از منظر نتایج، امکان دسته‌بندی‌هایی وجود دارد. بخش عمدهٔ نتایج در دو دستهٔ کلی قرار دارند. دستهٔ اول با به‌کارگیری روش‌های داده‌کاوی، متن‌کاوی و یادگیری ماشین به استخراج خودکار مفاهیم و ابعاد سه‌گانه (فعل، فاعل، مفعول) به‌همراه روابط معنایی از متن قرآن اقدام کرده‌اند. بیشتر نتایج نیز حاکی از سودمندی این روش‌ها دارد. پژوهش‌هایی نظیر صابر، عبدالجلیل و الفتاح بلال[1] (۲۰۲۲)، الخاماش و عبدسالم (۲۰۱۹)، وارگس و پونیثاولی[2] (۲۰۱۹)، الرحیلی و همکاران (۲۰۱۸)، بنتریکا، زیدت و ماریر[3] (۲۰۱۸) و حسین، سعد و نوح[4] (۲۰۱۷) از این دسته هستند. در دستهٔ دیگر، به

1. Saber, Abdel-Galil & El-Fatah Belal
2. Varghese & Punithavalli
3. Bentrica, Zidat & Marir
4. Husin, Saad & Noah

مقایسهٔ عملکرد روش‌ها و الگوریتم‌های مبتنی بر آمار و مشابهت‌یابی نظیر TF، TF-IDF، AVE-TF، Ridf، TIM، N-gram، FREyA، Pos Taggin، Levenshtein، Log Likelihod، هِرسِت، و جز این‌ها در استخراج مفاهیم خودکار جهت ساخت هستی‌نگاشت قرآنی مبادرت کردند. پژوهش‌های اوتومو و همکاران (۲۰۲۰)، اسماعیل و همکاران (۲۰۱۶، ۲۰۱۷ و ۲۰۱۸)، کدیر و یائوری (۲۰۱۷)، ویام و سد (۲۰۱۶) و سد و همکاران (۲۰۱۳) در این دسته هستند. همچنین اسماعیل و همکاران (۲۰۱۷) در پژوهشی به تأثیر مثبت استفاده از روش‌های ترکیبی آماری و زبانی اشاره داشت و الزمیل و الردیده (۲۰۱۴) با بهبود الگوریتم هِرسِت، روشی متفاوت را جهت استخراج مفاهیم خودکار پیشنهاد داد.

برخی پژوهش‌ها که به آن‌ها اشاره می‌گردد به نتایجِ و یافته‌هایی دست یافتند که جنبه‌های خاصی را در حوزهٔ استخراج خودکار مفاهیم قرآنی مدنظر داشتند. یافته‌های پژوهش اسماعیل و همکاران (۲۰۱۵) به اصلاح الگوهای پردازش زبان طبیعی به‌منظور استخراج رابطهٔ «بخشی_از» با استفاده از Pos و ابزار GATE منجر شد. محمود و حسن (۲۰۱۳) نیز با ایجاد حاشیه‌نویسی و استفاده از پایگاه داده، مشکل یادگیری و تلاوت فردی قرآن با هفت قرائت مختلف را از طریق توسعهٔ نرم‌افزار و ارائهٔ مدل تحلیل متن حل کردند.

در دستهٔ محدود دیگری، امکان یادگیری و تشخیص خودکار جنبه‌های ادبی موجود در قرآن نظیر آرایه‌های ادبی استعاره، تشبیه، روایت و قسم مورد بررسی قرار گرفت. پتیوالا و ساثیا[1] (۲۰۱۱) معماری عامل محوری را جهت یادگیری و تشخیص خودکار آرایه‌های ادبی قرآن، با ترکیب الگوریتم‌های وردنت و الگوهای نحوی واژگانی پیشنهاد کردند.

در پژوهش‌های سال‌های اخیر نیز توجه به ابزارهای معنایی نظیر گراف (Khazani et al., 2021) و ایجاد سیستم‌های پاسخگویی معنایی با بهره‌جویی از زبان پرس‌وجوی اسپارکل (Utomo, Suryana, & Sanusi Azmi, 2020) انجام شده است. یافته‌ها نشان از بهبود کارایی و دقت هستی‌نگاشت‌های قرآنی در جهت یادگیری خودکار به‌منظور جستجو و پاسخگویی دارد.

پرسش پنجم: پیشنهادهای آتی پژوهش‌های حوزهٔ یادگیری خودکار هستی‌نگاشت‌های قرآنی چیست؟

آیندهٔ پژوهش‌های حوزهٔ یادگیری خودکار هستی‌نگاشت‌های قرآنی به‌طور کلی به سمت‌وسوی استفاده از هستی‌نگاشت‌ها در کاربردهای مختلف نظیر موتورهای جستجوی معنایی خواهد رفت. همچنین با تمرکز بیشتر بر یادگیری ماشین و هوش مصنوعی در دنیا و کاربرد فراوان هستی‌نگاشت‌ها در این زمینه، استفاده از هستی‌نگاشت‌ها در یادگیری ماشین بسیار مورد اقبال قرار خواهد گرفت. از سویی دیگر، استفاده از الگوریتم‌های هوش مصنوعی در یادگیری هستی‌نگاشت‌ها نظیر مدل‌های BERT و یادگیری عمیق، برای تولید خودکار و نیمه‌خودکار آن‌ها به‌شدت پُرکاربرد خواهد شد.

1. Petiwala & Sathya

پیشنهادهای آتی پژوهش‌های حوزۀ یادگیری خودکار هستی‌نگاشت‌های قرآنی به‌تفصیل در جدول (۶) ذکر شده است. لازم به ذکر است که برخی از پژوهش‌ها دارای چندین پیشنهاد بودند و در دسته‌های متفاوت قرار گرفتند و برخی نیز پیشنهاد مشخصی را ارائه نکرده بودند.

جدول ۶. پیشنهادهای آتی پژوهش‌های حوزۀ یادگیری خودکار هستی‌نگاشت‌های قرآنی

پیشنهادهای پژوهشی آتی	نویسنده/نویسندگان (سال)
به‌کارگیری مدل‌های یادگیری عمیق نظیر ELMO، BERT، شبکۀ عصبی بازگشتی و شبکه‌های حافظۀ طولانی کوتاه‌مدت	Alshammeri, Atwell & Alsalka (2021); Varghese & Punithavalli (2019)
تمرکز و توجه به استخراج مفاهیم با قواعد و الگوهای ترکیبی	Alrehaili, Alqahtani & Atwell (2018); Ismail, Abd Rahman & Abu Bakar (2016, 2017, 2018)
طراحی و آزمون سیستم پاسخگویی معنایی به پرسش‌ها بر اساس هستی‌نگاشت‌های قرآنی	Khazani et al. (2021); Utomo, Suryana & Sanusi Azmi (2019)
ارائۀ الگو و معیارهای جدید جهت یادگیری و استخراج خودکار مفاهیم و آرایه‌ها	Sharaf & Atwell (2012); Petiwala & Sathya (2011); Al-Khalifa et al. (2010)
توجه به سایر متون دینی نظیر کتب فقهی و حدیثی (توسعۀ پیکرۀ بزرگ اسلامی) و همچنین سایر پایگاه‌های دانشی نظیر وردنت	Elkhammash & Abdessalem (2019); Irwanto, Bijaksana & Adiwijaya (2018); Kadir & Yauri (2017); Weaam & Saad (2016)
توسعۀ روش‌های نیمه‌خودکار	Elkhammash & Abdessalem (2019)
توجه به تکنیک‌های مبتنی بر مدل‌های زبانی جهت استخراج اطلاعات چندگانه به‌عنوان مفاهیم	Ismail, Abd Rahman & Abu Bakar (2018)
به‌کارگیری الگوریتم‌های ریشه‌شناسی در استخراج مفاهیم	Husin, Saad & Noah (2017)
تمرکز و توجه به واژگان غیرمجاز و قرار دادن حرف تعریف «ال» در لیست واژگان غیرمجاز	Husin, Saad & Noah (2017)
استفاده از طبقه‌بندی‌کننده‌ها و یادگیری خودکار جهت استخراج روابط در عبارات ربطی	Bentrcia, Zidat & Marir (2018); Ismail, Abu Bakar & Abd Rahman (2015)
بازیابی هوشمند مبتنی بر عامل و حاشیه‌نویسی معنایی اصوات مرتبط با قرآن	Baqai et al. (2009)

نتیجه‌گیری

پژوهش حاضر راهکارهای مختلف یادگیری و ساخت خودکار و نیمه‌خودکار هستی‌نگاشت‌ها را به‌منظور استخراج دانش حوزۀ مفاهیم قرآنی مورد مطالعه قرار داده است. مطالعۀ انجام‌شده به‌روش مرور

دامنه‌ای بود که مقایسهٔ بین راهکارهای ارائه‌شده را فراهم می‌کند. معیارهای مورد بررسی مشتمل بر مجموعه داده، روش‌های یادگیری و ارزیابی هستی‌نگاشت‌ها و ارائه یافته‌ها و پیشنهادهای پژوهش‌های این حوزه بود. اگرچه روش‌های مختلف یادگیری هستی‌نگاشت برای توسعهٔ خودکار هستی‌نگاشت‌های قرآنی از طریق روش‌های مهندسی هستی‌نگاشت‌ها ارائه شده، اما هیچ بررسی مروری مرتبطی تا به امروز برای این موضوع گزارش نشده است.

بررسی یافته‌ها پیرامون پرسش نخست پژوهش نشان داد که اغلب پژوهش‌ها در حوزهٔ مجموعه داده‌های قرآنی به زبان‌های انگلیسی و عربی بودند که یافته‌های متداولی است. پژوهش‌های انجام‌شده درزمینهٔ یادگیری هستی‌نگاشت با استفاده از مجموعه دادهٔ بسیار محدودی انجام شده‌اند که خود پژوهشگران نیز به آن اذعان کرده‌اند و در اغلب این پژوهش‌ها برای کارهای بعدی، استفاده از مجموعه داده‌های بزرگتر پیشنهاد شده است. استفاده از مجموعه داده‌های بسیار محدود، عمده‌ترین محدودیت پژوهش‌های این حوزه بود. این مسئله می‌تواند با طراحی هستی‌نگاشت جامع قرآنی که تمام موضوعات و مفاهیم را با توجه به بافت قرآن، پوشش دهد، مدنظر قرار گرفته و با ایجاد هستی‌نگاشت جامعی از مفاهیم قرآن، کاربران را به‌سمت بازیابی دانش قرآنی رهنمون سازد.

بیشتر پژوهش‌هایی که با هدف یادگیری خودکار هستی‌نگاشت صورت گرفته، ازنظر انتخاب مجموعه داده، استفاده از دادهٔ پایگاه‌ها یا مجموعه داده‌های بزرگتر جهت مقایسه نتایج بر روی داده‌های آموزش و آزمایش[1] مرسوم‌تر است. به‌عنوان‌مثال در آن‌ها از متن کامل و یا ترجمه و تفاسیر متن کامل قرآن کریم استفاده شده که همگام با یافته‌های پژوهش حاضر، مشاهده می‌شود که در پژوهش‌هایی که هدف پژوهشگر، تنها توسعهٔ هستی‌نگاشت با استفاده از روش‌های دستی و غیرخودکار بوده، از دامنه‌های محدودتری نظیر یک یا چند سوره، یک جزء، یا آیات مرتبط با موضوعی نظیر اسامی خاص استفاده شده است (Mirarab, Tabatabai Amiri, Dehghanisanij, & HosseinKhalili, 2023).

در اغلب راهکارهای ارائه‌شده در پژوهش‌ها، از روش‌های داده‌کاوی برای یادگیری و ساخت خودکار هستی‌نگاشت و توسعهٔ آن در حوزهٔ مفاهیم قرآنی استفاده شده است. بخش عمده پژوهش‌ها از روش‌های نرمال‌سازی، خوشه‌بندی و دسته‌بندی متن، خلاصه‌سازی متن، استخراج اطلاعات، تشابه و یافتن موجودیت‌های نامدار استفاده کرده‌اند. البته در برخی پژوهش‌ها، روش‌های هوش مصنوعی نظیر شبکهٔ عصبی نیز به کار گرفته شده است. یافته‌ها نشان می‌دهد راهکارهای مبتنی بر آمار برای یادگیری و ساخت خودکار و نیمه‌خودکار هستی‌نگاشت‌ها بسیار مورد اقبال پژوهشگران بوده است.

داده‌کاوی، فرایند استفاده از الگوریتم‌های محاسباتی و یادگیری ماشینی جهت کسب بینش از طریق کشف الگوها و روندها در مجموعه‌ای از داده‌ها است. در فرایند داده‌کاوی اطلاعات مفیدی که از

1. train and test data

مسیر پردازش داده‌ها استخراج شده، به ساختار قابل درکی برای تحلیل‌های بعدی تبدیل می‌شود. از سوی دیگر، هوش مصنوعی، حوزۀ وسیع‌تری است که هدف آن، هوشمندسازی ماشین برای انجام اموری نظیر درک تصویر، تشخیص صدا، استنتاج و تصمیم‌گیری، و پردازش زبان طبیعی است. در هوش مصنوعی از روش‌های متنوعی نظیر یادگیری ماشین، یادگیری عمیق، شبکه‌های عصبی، ایجاد گراف معنایی و پردازش زبان طبیعی برای یادگیری و اتخاذ تصمیم توسط ماشین، همانند انسان بهره‌جویی می‌شود. بر همین اساس، یافته‌های پژوهش حاضر نشان می‌دهد که در پژوهش‌هایی نظیر وارگس و پونیثاولی (۲۰۱۹) و حسین، سد و نوح (۲۰۱۷) که از تکنیک‌های هوش مصنوعی و یادگیری ماشین در توسعۀ خودکار هستی‌نگاشت‌های قرآنی استفاده کرده‌اند، اغلب از متون ترجمه و تفسیر به زبان طبیعی و متون بدون ساختار بهره‌برداری شده است. اما در روش‌های داده‌کاوی از اصطلاحنامه‌ها، وردنت و پیکره‌های موجود که دارای محتوای ساختارمند بوده‌اند، به‌منظور آماده‌سازی مجموعه دادۀ اولیه جهت ساخت خودکار هستی‌نگاشت استفاده شده است (Weaam & Saad, 2016). در همین راستا، بررسی یافته‌های پژوهش حاضر حاکی از آن است که نتایج پژوهش‌هایی که در آن‌ها از الگوریتم‌های داده‌کاوی استفاده شده، درنهایت، منجر به انتخاب خودکار مفاهیم و تشخیص و دسته‌بندی روابط معنایی بین مفاهیم منتخب، شده است. در سوی دیگر، پژوهش‌هایی که از روش‌های هوش مصنوعی بهره‌برداری کرده‌اند، با تحلیل معنایی، استنتاج، مدل‌سازی و تأیید اعتبار داده‌های استنتاج‌شده به نتایجی مانند تشخیص صوت برای آموزش قرائت قرآن، تشخیص آرایه‌های ادبی و ایجاد ارتباط‌های موضوعی در مفاهیم قرآنی و همچنین ایجاد ارتباط بین این مفاهیم با مفاهیم سایر ادیان نائل شده‌اند.

در پژوهش‌های داده‌کاوی، اثبات شده است که انتخاب و به‌کارگیری هم‌زمان الگوریتم‌ها به‌صورت ترکیبی با توجه به نوع و حوزۀ علمی دادۀ پایگاه مورد نظر، منجر به بهبود نتایج می‌شود (Ismail et al., 2018). مطابق با این یافته، در بحث پیرامون روش‌های یادگیری خودکار هستی‌نگاشت‌های قرآنی مشاهده شد که اغلب پژوهش‌ها مبادرت به اتخاذ روش‌هایی با ترکیب الگوریتم‌های مختلف با یکدیگر به‌منظور افزایش میزان دقت و ثبات[1] در نتایج حاصل کرده‌اند. یافته‌های پژوهش حاضر نشان می‌دهد که برای ارزیابی نتایجی که با روش‌های آموزش هوش مصنوعی بر روی مجموعه داده‌های به‌نسبت بزرگ حاصل‌شده از انواع معیارها و سنجه‌های موجود یا پیشنهادی استفاده شده است. این یافته قابل مقایسه با روش‌های اعتباریابی انسانی در نتایج حاصل از داده‌کاوی بر مجموعه داده‌های کوچک‌تر مثل اسامی موجودات در قرآن است.

یافته‌های پرسش سوم پژوهش نشان داد که پژوهش‌های انجام‌شده برای ارزیابی راهکارهای ارائه‌شده از روش‌های محاسبۀ دقت، فراخوانی و F-MEASURE استفاده کرده‌اند. اما برخی پژوهش‌ها از

1. robustness

ارزیابی انسانی استفاده کرده‌اند. محاسبهٔ مثبت واقعی، مثبت کاذب و منفی کاذب نیز مورد استفاده برخی از پژوهش‌ها قرار گرفته است. روش‌های ارائه‌شده برای ارزیابی یادگیری خودکار هستی‌نگاشت‌ها نشان می‌دهد استفاده هم‌زمان از روش‌های داده‌کاوی و هوش مصنوعی نتایج بهتری را به‌همراه دارد.

نتیجه‌گیری حاصل از ارزیابی نتایج حائز این نکته است که توجه به سازوکار الگوریتم‌ها و میزان ثبات آن‌ها در استنتاج از مجموعه داده‌های زبانی نسبت به مجموعه داده‌های سایر علوم غیرقرآنی بیشتر است. زیرا همان‌طور که در روش‌های ارزیابی هستی‌نگاشت‌ها مشاهده می‌شود، در محاسبهٔ انواع معیارهای دقت، فراخوانی و معیارF، افزایش یک سنجه با کاهش سنجهٔ دیگر همراه خواهد بود. درنتیجه آنچه محقق را در ترجیح هر یک از این معیارها و سنجه‌ها بر دیگری یاری می‌کند، حصول ثبات در نتایج نهایی استنتاج از داده‌ها است، و لازم است که میزان قابل قبولی از ثبات در نتایج رعایت شود. شایان ذکر است که محتوای قرآن کریم، کلام‌الله بوده و نسبت به سایر مجموعه داده‌های زبانی از ارزشی متعالی برخوردار است، توجه به این نکته که در کنار ارزیابی‌های خودکار نیاز به اعتباریابی هم‌زمان نتایج توسط متخصص انسانی ضرورت دارد، بیش‌ازپیش حائز اهمیت خواهد بود. باید توجه داشت که این رویکرد نیاز به منابع محاسباتی قابل توجهی داشته و توجه دقیق به پیامدهای اخلاقی مربوط به محتوای دینی تولیدشده توسط هوش مصنوعی ضرورت دارد. مؤید این مطلب نیز ظهور فناوری جدید موسوم به ترنسفورمر از پیش آموزش‌دیدهٔ تولیدگر (جی.پی.تی.)[1] و گاهی پاسخ‌های مشکل‌داری است که باید جانب احتیاط را در مورد محتواهای دینی رعایت کرد.

یافته‌ها و نتایج پژوهش‌های این حوزه نشان داد که بخش عمدهٔ آن با بهره‌جویی از روش‌های داده‌کاوی، متن‌کاوی و یادگیری ماشین به استخراج خودکار مفاهیم و ابعاد سه‌گانه (فعل، فاعل، مفعول) به‌همراه روابط معنایی از متن قرآن اقدام کرده‌اند. بیشتر نتایج نیز حاکی از سودمندی این روش‌ها داشت. دستهٔ دیگر نیز به مقایسهٔ عملکرد روش‌های اشاره‌شده و الگوریتم‌های مبتنی بر آمار پرداخته‌اند. برخی پژوهش‌های جدیدتر نیز به استفاده از الگوریتم‌های هوش مصنوعی مانند یادگیری عمیق روی آورده‌اند. با گسترش استفاده از روش‌های مبتنی بر هوش مصنوعی و یادگیری ماشین، اقبال محققان به این حوزه‌ها، بیشتر شده و یادگیری هستی‌نگاشت‌ها به این سمت سوق پیدا می‌کند.

یافته‌های حاصل از بررسی کارهای آینده نشان از علاقهٔ محققان به الگوریتم‌های هوش مصنوعی و استفاده در یادگیری هستی‌نگاشت و توسعهٔ خودکار و نیمه‌خودکار هستی‌نگاشت‌های قرآنی دارد. با توجه به اینکه یادگیری ماشین مستلزم وجود داده‌های کلان در حوزهٔ قرآن است، ساخت مجموعه داده‌های استاندارد ازجمله کارهای آتی محققان است. فقدان مجموعه داده‌های صحیح، دلیل عجز سیستم‌های هوش مصنوعی پیشرفتهٔ دنیا مانند جی‌پی‌تی ۴ است که در آینده باید به این مهم پرداخته شود. برخی

1. Generative Pre-trained Transformer (GPT)

پژوهش‌ها نیز به استفاده از هستی‌نگاشت‌ها در برنامه‌های کاربردی مختلف نظیر موتورهای جستجوی معنایی یا سامانه‌های بازیابی اطلاعات معنایی به‌عنوان کارهای آینده اشاره داشتند. پیشنهادی که کاربرد هستی‌نگاشت‌های قرآنی را در عمل مورد هدف قرار داده است. پژوهش‌های آتی، عملیاتی شدن هستی‌نگاشت‌های قرآنی را مدنظر قرار داده تا پژوهش‌ها از فضای آزمایشگاهی به فضای واقعی وارد شود.

پیشنهاد می‌شود متناسب با ویژگی‌های بافتی منحصربه‌فرد قرآن کریم، ترکیبی از تکنیک‌های یادگیری ماشین و یادگیری عمیق مورد آزمایش قرار گرفته و با ارزیابی نتایج بهترین تکنیک‌ها معرفی شوند. به‌عنوان‌مثال بررسی حضور «الـ» در کلمات و همچنین در نظر گرفتن وقف‌های واجب در استخراج روابط بین مفاهیم قرآنی برای بهبود الگوریتم‌های موجود مدنظر قرار گیرد. همچنین در بحث استفاده از روش یادگیری هم‌آیندها می‌توان با پیاده‌سازی این روش بر روی عبارات حاوی کلمهٔ جلالهٔ «الله» چالش بالاترین رخداد اصطلاح «الله» در کل متن را مرتفع کرده تا بتوان درنتیجه، گراف دقیق‌تری از این اصطلاح طراحی شود. پیشنهاد دیگر، ساخت برنامه‌های کاربردی در حوزهٔ استخراج مفاهیم قرآنی با استفاده از پیکره‌های ساخته‌شده در پژوهش‌های موجود نظیر پیکرهٔ QurAna است (Sharaf & Atwell, 2012). بر همین اساس، بهره‌برداری بیشتر از روش‌های هوش مصنوعی و پردازش زبان طبیعی نظیر جی.پی.تی. به‌عنوان مدل یادگیری ماشینی برای تولید متن به زبان طبیعی با استفاده از شبکهٔ عصبی عمیق، در توسعهٔ خودکار هستی‌نگاشت‌های قرآنی ضروری به نظر می‌رسد. در ادامه با توجه به یافته‌های به‌دست‌آمده، پیشنهادهایی برای پژوهش‌های کاربردی به شرح زیر ارائه می‌شود:

- ارزیابی فنی و جزئی روش‌های داده‌کاوی و هوش مصنوعی مورد استفاده در توسعهٔ خودکار و نیمه‌خودکار هستی‌نگاشت‌های قرآنی با توجه به پیچیدگی متن قرآن؛
- ارزیابی مجموعه داده‌های مورد استفاده در هستی‌نگاشت‌های تولیدشده به‌روش خودکار و نیمه‌خودکار براساس معیارهای ساخت مجموعه داده؛
- ایجاد هستی‌نگاشت جامع قرآن‌کریم به زبان فارسی با استفاده از روش‌های خودکار و نیمه‌خودکار.

References

Al-Khalifa, H. S., Al-Yahya, M., Bahanshal, A., Al-Odah, I., & Al-Helwah, N. (2010). *An approach to compare two ontological models for representing quranic words* Proceedings of the 12th International Conference on Information Integration and Web-based Applications & Services, Paris, France. https://doi.org/10.1145/1967486.1967593

Al-Zamil, M. G. H., & Al-Radaideh, Q. (2014). Automatic extraction of ontological relations from Arabic text. *Journal of King Saud University - Computer and Information Sciences*, *26*(4), 462-472. https://doi.org/https://doi.org/10.1016/j.jksuci.2014.06.007

Alrehaili, S. M., Alqahtani, M., & Atwell, E. (2018, March). A hybrid methods of aligning

Arabic Qur'anic semantic resources. *2018 IEEE 2nd International Workshop on Arabic and Derived Script Analysis and Recognition (ASAR)*, London (pp. 108-113). IEEE. doi:2010.1109/ASAR.2018.848.309

Alshammeri, M., Atwell, E., & Alsalka, M. a. (2021). Detecting Semantic-based Similarity Between Verses of The Quran with Doc2vec. *Procedia Computer Science*, *189*, 351-358. https://doi.org/https://doi.org/10.1016/j.procs.2021.05.104

Baqai, S., Basharat, A., Khalid, H., Hassan, A., & Zafar, S. (2009). *Leveraging semantic web technologies for standardized knowledge modeling and retrieval from the Holy Qur'an and religious texts* Proceedings of the 7th International Conference on Frontiers of Information Technology, Abbottabad, Pakistan. https://doi.org/10.1145/1838002.1838050

Bentrcia, R., Zidat, S., & Marir, F. (2018). Extracting semantic relations from the Quranic Arabic based on Arabic conjunctive patterns. *Journal of King Saud University - Computer and Information Sciences*, *30*(3), 382-390. https://doi.org/https://doi.org/10.1016/j.jksuci.2017.09.004

Castro-Schez, J. J., Jennings, N. R., Luo, X., & Shadbolt, N. R. (2004). Acquiring domain knowledge for negotiating agents: a case of study. *International Journal of Human-Computer Studies*, *61*(1), 3-31. https://doi.org/https://doi.org/10.1016/j.ijhcs.2003.09.006

Corcho, O., Fernandez-Lopez, M., & Gomez-Perez, A. (2007). Ontological engineering: what are ontologies and how can we build them? In *Semantic web services: Theory, tools and applications* (pp. 44-70). IGI Global.

Elkhammash, E., & Abdessalem, W. B. (2019). A Holy Quran Ontology Construction with Semi-Automatic Population. *Bioscience Biotechnology Research Communications*, *12*(1), 229-234. https://doi.org/http://dx.doi.org/10.21786/bbrc/12.1/26

Etzioni, O., Fader, A., Christensen, J., Soderland, S., & Mausam, M. (2011, July). Open information extraction: the second generation. In Toby Walsh (Ed.), *IJCAI'11: Proceedings of the Twenty-Second international joint conference on Artificial Intelligence - Volume One*, Barcelona, Spain (pp 3-10). AAAI Press.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199-220. https://doi.org/https://doi.org/10.1006/knac.1993.1008

Gulla, J. A., Borch, H. O., & Ingvaldsen, J. E. (2007). Ontology Learning for Search Applications. In R. Meersman & Z. Tari, *On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS* Berlin, Heidelberg.

Husin, M. Z., Saad, S., & Noah, S. A. M. (2017, November). Syntactic rule-based approach for extracting concepts from quranic translation text. *2017 6th International Conference on Electrical Engineering and Informatics (ICEEI)*, Langkawi, Malaysia (pp. 1-6). IEEE. doi:2010.1109/ICEEI.2017.8312367

Irwanto, Y., Bijaksana, M. A., & Adiwijaya. (2018). Semantic text relatedness on Al-Qur'an translation using modified path based method. *Journal of Physics: Conference Series*, *971*(1), 012047. https://doi.org/10.1088/1742-6596/971/1/012047

Ismail, R., Abd Rahman, N., & Abu Bakar, Z. (2016, 10-12 Oct. 2016). Identifying concept from English translated Quran. 2016 IEEE Conference on Open Systems (ICOS),

Ismail, R., Abd Rahman, N., & Abu Bakar, Z. (2017). A Pattern for Concept Identification from English Translated Quran. *MATEC Web Conf.*, *135*, 00067. https://doi.org/10.1051/matecconf/201713500067

Ismail, R., Abd Rahman, N., & Abu Bakar, Z. (2018). Single Term Concepts from English Translated Qur'an Using Statistical Methods. *International Journal of Engineering & Technology*, *7*(2.14), 13-16.

Ismail, R., Abd Rahman, N., Abu Bakar, Z., & Makhtar, M. (2018, May). Concepts extraction in ontology learning using language patterns for better accuracy. *2018 4th International Conference on Computer and Technology Applications (ICCTA)*, Istanbul, Turkey (pp. 122-126). IEEE. doi: 10.1109/CATA.2018.8398668

Ismail, R., Abu Bakar, Z., & Abd Rahman, N. (2015). Extracting knowledge from english translated quran using nlp pattern. *Jurnal Teknologi*, *77*(19), 67-73. https://doi.org/10.11113/jt.v77.6515

Jalali, R., & Mohammadi, M. (2021). Scoping review, missing loop in chain of review articles: Letter to the Editor [Original Article]. *Tehran University Medical Journal*, *78*(10), 710-711. http://tumj.tums.ac.ir/article-1-10938-fa.html. [In Persian]

Kadir, R., & Yauri, R. (2017). Resource description framework triples entity formations using statistical language model. *Journal of Fundamental and Applied Sciences*, *9*(4S), 710-729.

Khazani, M. M. M., Mohamed, H., Sembok, T. M. T., Yusop, N. M. M., Wani, S., Gulzar, Y., Halip, M. H. M., Marzukhi, S., & Yunos, Z. (2021). Semantic graph knowledge representation for Al-Quran verses based on word dependencies. *Malaysian Journal of Computer Science*, 132-153.

Maedche, A., & Staab, S. (2001). Ontology learning for the semantic web. *IEEE Intelligent systems*, *16*(2), 72-79.

Mahmoud, M., & Hassan, I. (2013, December). Artificial intelligence techniques for extracting individuals recitation of the Holy Quran from its combinations. *2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences*, Madinah, Saudi Arabia (pp. 292-297). IEEE. doi:10.1109/NOORIC.2013.65

Mirarab, A., Tabatabai Amiri, F. S., Dehghanisanij, S., & HosseinKhalili, N. (2023). Development of quranic ontologies: A domain review study. *International Journal of Information Science and Management*, *21*(3), 229-241. https://doi.org/https://doi.org/10.22034/ijism.2023.1977928.0

Petiwala, A. J., & Sathya, S. S. (2011, September). A multi-agent system to learn literature ontology: An experiment on English Quran corpus. *2011 2nd International Conference on Intelligent Agent & Multi-Agent Systems*, Chennai, India (pp. 46-51). IEEE. doi:10.1109/IAMA.2011.6049002

Saad, S., Noah, S. A. M., Salim, N., & Zainal, H. (2013, December). Rules and natural language pattern in extracting Quranic knowledge. *2013 Taibah University International Conference on Advances in Information Technology for the Holy Quran and Its Sciences*, Madinah, Saudi Arabia (pp. 381-386). IEEE. doi:10.1109/NOORIC.2013.80

Saber, Y. M., Abdel-Galil, H., & El-Fatah Belal, M. A. (2022). Arabic ontology extraction model from unstructured text. *Journal of King Saud University - Computer and Information Sciences*, *34*(8, Part B), 6066-6076. https://doi.org/https://doi.org/10.1016/j.jksuci.2022.02.007

Sánchez, D., & Moreno, A. (2008). Learning non-taxonomic relationships from web documents for domain ontology construction. *Data & Knowledge Engineering*, *64*(3),

600-623. https://doi.org/https://doi.org/10.1016/j.datak.2007.10.001

Shamsfard, M., & Abdollahzadeh Barforoush, A. (2003). The state of the art in ontology learning: a framework for comparison. *The Knowledge Engineering Review*, *18*(4), 293-316. https://doi.org/10.1017/S0269888903000687

Shannon, S. (2002). Critical appraisal of systematic reviews. *Canadian Association of Radiologists Journal*, *53*(4), 195.

Sharaf, A.-B., & Atwell, E. (2012, May). QurAna: Corpus of the Quran annotated with Pronominal Anaphora, In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)* Istanbul, Turkey (pp. 130-137). European Language Resources Association (ELRA).

Tomás, V. R., & García, L. A. (2005). Agent-Based Management of Non Urban Road Meteorological Incidents. In M. Pěchouček, P. Petta, & L. Z. Varga, *Multi-Agent Systems and Applications IV* Berlin, Heidelberg.

Uren, V., Cimiano, P., Iria, J., Handschuh, S., Vargas-Vera, M., Motta, E., & Ciravegna, F. (2006). Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Journal of Web Semantics*, *4*(1), 14-28. https://doi.org/https://doi.org/10.1016/j.websem.2005.10.002

Utomo, F. S., Suryana, N., & Sanusi Azmi, M. (2019). New instances classification framework on Quran ontology applied to question answering system. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, *17*, 139-146. https://doi.org/10.12928/TELKOMNIKA.v17i1.9794

Utomo, F. S., Suryana, N., & Sanusi Azmi, M. (2020). Stemming Impacr Analysis on indonesian quran translation and their tafsir classification for ontology instances. *IIUM Engineering Journal*, *21*(1), 33 -50. https://doi.org/10.31436/iiumej.v21i1.1170

Varghese, N., & Punithavalli, M. (2019). Lexical and semantic analysis of sacred texts using machine learning and natural language processing. *International Journal of Scientific & Technology Research*, *8*(12), 3133-3140.

Weaam, T., & Saad, S. (2016). Ontology population from quranic translation texts based on a combinationof linguistic patterns and association rules. *Journal of Theoretical and Applied Information Technology*, *86*(2), 250-257.

Wimalasuriya, D. C., & Dou, D. (2010). Ontology-based information extraction: An introduction and a survey of current approaches. *Journal of Information Science*, *36*(3), 306-323. https://doi.org/10.1177/0165551509360123