

Разработка гибридной онтологической модели базы знаний для выбора абитуриентами направления обучения¹

В. С. Мошкин, И. А. Андреев, Е. С. Ваганова, А. А. Рейц

Ульяновский государственный технический университет, Ульяновск, Россия

Поступила в редколлегию 29.09.2025

Аннотация—В статье представлен подход к построению базы знаний экспертной системы, обеспечивающей поддержку принятия решений абитуриентов на этапе предобработки предоставляемых данных и расстановки приоритетности направлений. Модель базы знаний включает гибридизацию предметной онтологии OWL и набора правил SWRL. Предметная онтология включает в себя классы, объекты и отношения, необходимые для валидации входных данных и реализации логического вывода о возможности поступления абитуриента на отдельные направления подготовки. В статье приведены примеры правил SWRL, которые включены в онтологию. На основе интеграции правил и OWL-онтологии делается логический вывод. Также в работе представлены результаты использования разработанной базы знаний в рамках работы технического персонала приемной комиссии в рамках приемной кампании вуза 2024 г. Проведенные эксперименты показали эффективность применения разработанной базы знаний, применение которой на этапе предобработки данных и расстановки приоритетности направлений позволило снизить временные затраты на 27%.

КЛЮЧЕВЫЕ СЛОВА: логический вывод, онтология, база знаний, приемная комиссия, OWL, SWRL.

DOI: 10.53921/18195822_2025_25_3.1_550

1. ВВЕДЕНИЕ

Основная цель любой базы знаний – сокращение времени на поиск информации и получение ответа на типовой вопрос в любой момент. Полноценные базы знаний содержат в себе фактическую информацию и правила вывода и позволяют делать автоматические умозаключения об уже имеющихся или вновь вводимых фактах и тем самым производить семантическую обработку информации.

На данный момент разработаны различные модели представления знаний практически для любой предметной области. К ним относятся: продукционные модели; семантические сети; фреймы; логические модели; нейронные сети и пр.

Онтологией можно считать интеллектуальные средства поиска, представления и обработки знаний и запросов. Широкое практическое применение онтологий, прежде всего, обусловлено таким преимуществом, как эффективное описание семантики данных для выбранной предметной области, которое позволяет при обращении к поисковой системе получать в ответ ресурсы, семантически релевантные запросу.

¹ Работа выполнена в рамках государственного задания № 075-03-2023-143 по проекту «Исследование интеллектуального предиктивного мультимодального анализа больших данных, конструирование признаков гетерогенных динамических данных для машинного обучения»

Приёмная комиссия – подразделение высшего или среднего учебного заведения, осуществляющее приёмную кампанию на очередной учебный год [1]. Основной целью создания приёмной комиссии является организация приёмной кампании: приём документов от лиц, поступающих в вуз, и обеспечение зачисления в учебное заведение, ознакомление абитуриентов с правилами и положениями учебного заведения, документацией по аккредитации направлений и специальностей, критериями приёма. Она также оказывает консультации по вопросам организации обучения и, что самое главное, принимает решение о зачислении или отказе в зачислении абитуриента в учебное заведение [2]. В настоящее время приёмная комиссия структурирована и состоит из определённых подразделений, которые выполняют установленные обязанности.

В настоящее время приемная комиссия является структурированной и состоит из определенных единиц, выполняющих установленные обязанности. Анализируя состав приемной комиссии различных вузов, установили, что в основном она состоит из:

- председателя приемной комиссии; заместителя председателя комиссии;
- заведующих кафедрами;
- председателей предметных экзаменационных комиссий;
- секретариата;
- технического персонала [3] [4].

Технический персонал принимает документы, консультирует по составу пакета, необходимого при поступлении; производит запись данных об абитуриенте в регистрационный журнал; оформляет документы личных дел абитуриентов; отвечает на вопросы поступающих в форме звонков, электронных писем, вопросов в устной форме абитуриентов, пришедших за консультацией по вопросам, связанным с организацией учебной деятельности, зачисления [5].

В настоящее время наблюдается тенденция на увеличение объема знаний и соответственно требования руководства к уровню подготовки персонала [6].

Возможными аналогами разрабатываемой базы знаний являются:

- текстовые данные по направлениям подготовки, представленные в неструктурированном виде. Такой тип представления информации не обладает эффективными алгоритмами поиска.
- вики-ресурсы, в том числе взаимосвязи между страницами с использованием гиперссылок. Основным недостатком является сложность поиска и отсутствие возможности вывода ответов.
- базы данных. Базы данных не хранят условия и правила, необходимые для вывода.

Таким образом, за счёт использования специализированной базы знаний можно решить задачу поддержки работы технического персонала в области предварительной обработки данных и ориентирования абитуриентов при выборе направления обучения [7].

2. МОДЕЛЬ БАЗЫ ЗНАНИЙ ЭКСПЕРТНОЙ СИСТЕМЫ

Для решения поставленной задачи была предложена модель базы знаний экспертной системы поддержки работы технического персонала приемной комиссии. Формальное описание объектов [8] онтологической базы знаний можно представить следующим образом:

$$O = \langle D, MB, SB, F \rangle, \text{ где}$$

D – документ об образовании;

MB – минимальные баллы по каждому учитываемому предмету при поступлении;

SB – сумма баллов абитуриента;

F – желаемый факультет для обучения;

$$D = \langle A, DSPO, DVO, DS, DSp \rangle, \text{ где}$$

A – аттестат;

$DSPO$ – диплом о среднем профессиональном образовании;

DVO – диплом о высшем образовании;

DS – диплом специалиста, полученный до 2015 года;

DSp – диплом специалиста, полученный после 2015 года.

Наличие минимальных баллов по каждому учитываемому предмету и соответствие документа об образовании определяет условия соотнесения абитуриента к анонимным классам «Может поступать на программы бакалавриата и специалитета» и «Может поступать на программы магистратуры».

Анонимный класс содержит все индивиды, удовлетворяющие ограничению, т.е. все индивиды, которые связаны заданными свойствами с индивидами заданных классов по заданным правилам [9] [10].

MB - минимальные баллы по предметам.

SB - сумма баллов при поступлении от 140 до 240 с шагом в 20 баллов: 140-160, 160-180, 180-200, 200-220, 220-240, более 240 баллов.

F – факультеты подготовки вуза.

Предлагаемая структура базы знаний – это интеграция OWL-онтологии и набора SWRL-правил. [11] [12]

Набор интерпретативных функций (аксиом) онтологической базы знаний представляется в виде:

$$F = \langle FC, FIn, FR \rangle,$$

$FC: \{CRule\} \rightarrow \{C\}$ – это функция, связывающая набор классов онтологии OWL с набором классов, используемых в системе продукционных правил SWRL;

$FIn: \{InRule\} \rightarrow \{In\}$ – это функция, связывающая набор классов объектов онтологии OWL с набором объектов классов, используемых в системе продукционных правил SWRL;

$FC: \{RRule\} \rightarrow \{RD\}$ – это функция, связывающая набор отношений между объектами и типами данных, используемых в системе продукционных правил SWRL, с набором отношений классов онтологии OWL [13] [14]. Правила SWRL являются дизъюнктами Хорна и обычно записываются следующим образом:

$$Ax : C1(?x) \wedge C2(?y) \wedge r1(?x, ?y) \wedge C3(?x, ?z) \rightarrow C2(?z, ?y), \text{ где}$$

$(C1, C2, C3) \in C, r1 \in R,; x, y$ – экземпляры или переменные; z – переменные или значения [15].

При формировании SWRL правил из каждого блока используются данные и предлагается трек поступления, в котором пронумерованы направления в порядке расстановки приоритетности с первого до пятого [16].

Для формирования треков поступления используются статистические данные из отчетных материалов приемной комиссии и учебного управления за последние пять лет.

Блок-схема логического вывода итогов анализа по расстановке приоритетности направлений представлена на рисунке 1.

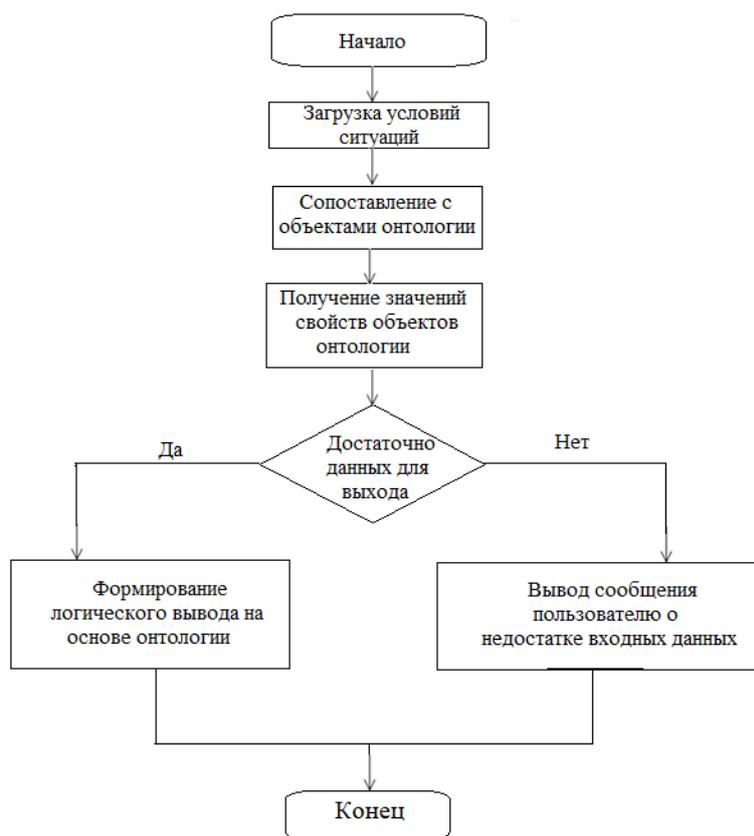


Рис. 1. Схема логического вывода итогов анализа

Схема взаимодействия форматов знаний OWL и SWRL в процессе логического вывода результатов анализа представлена на рисунке 2 [17].

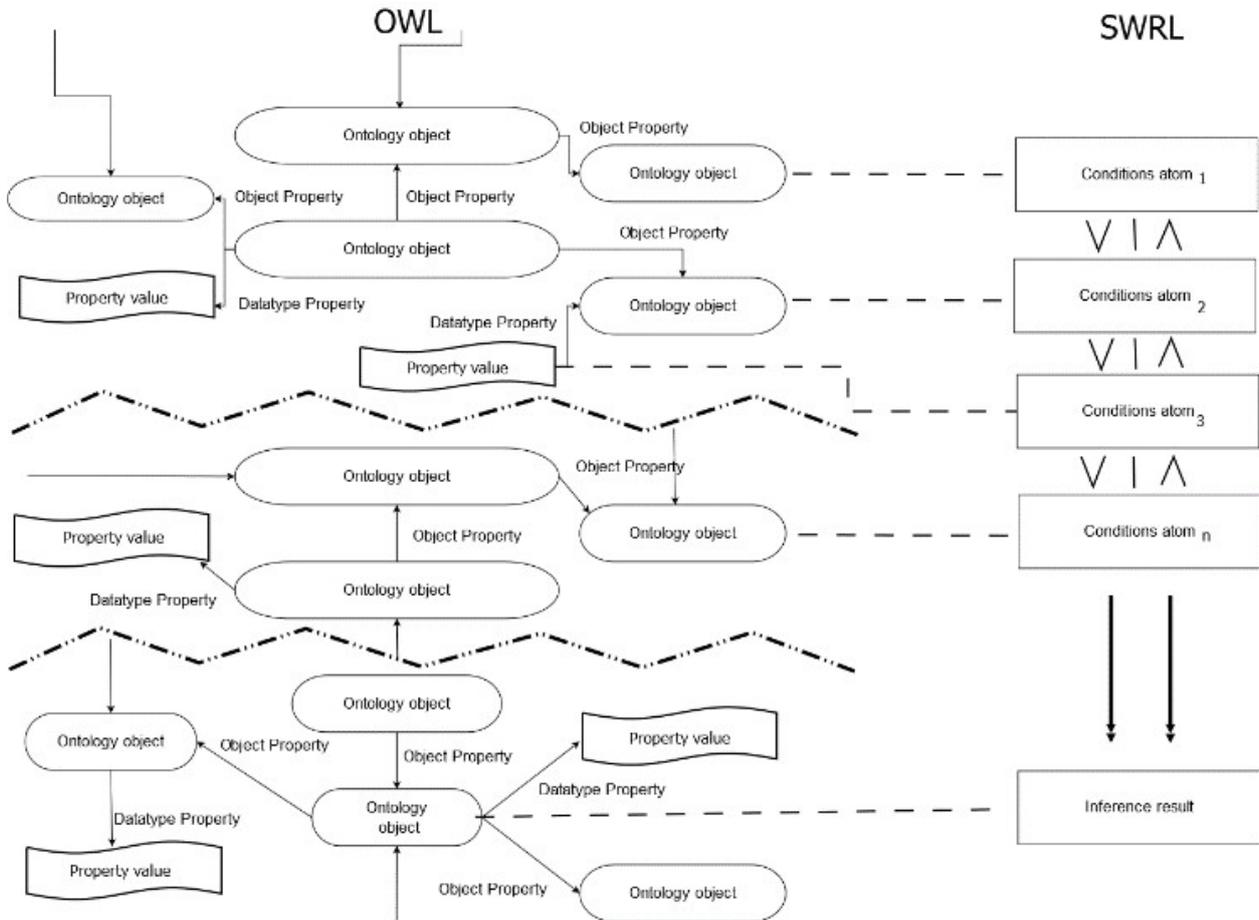


Рис. 2. Схема взаимодействия форматов знаний OWL и SWRL в процессе логического вывода результатов анализа.

3. РЕАЛИЗАЦИЯ БАЗЫ ЗНАНИЙ И РЕЗУЛЬТАТЫ ЭКСПЕРИМЕНТОВ

Основой разработанной в рамках проекта экспертной системы является база знаний на основе описанного ранее механизма интеграции онтологии и набора SWRL-правил [18].

Для разработки базы знаний был использован редактор Protege. Разработанная экспертная система на основе онтологической модели базы знаний для расстановки приоритетности направлений при поступлении в УлГТУ состоит из 10 классов верхнего уровня, 7 – классов второго уровня (рис.3).

На этапе проектирования был определен базовый класс «Абитуриент» и его подклассы – специальности (уровень бакалавриата/специалитета: гуманитарные, технические, экономические; и «поступающий в магистратуру»), сформулированы свойства объектов (topObjectProperty) и созданы несколько примеров экземпляров [19]. Экземпляры взаимодействуют друг с другом через свойства объектов или Object Property. В разработанной онтологии свойства экземпляров задаются с помощью раздела Data Properties [20].

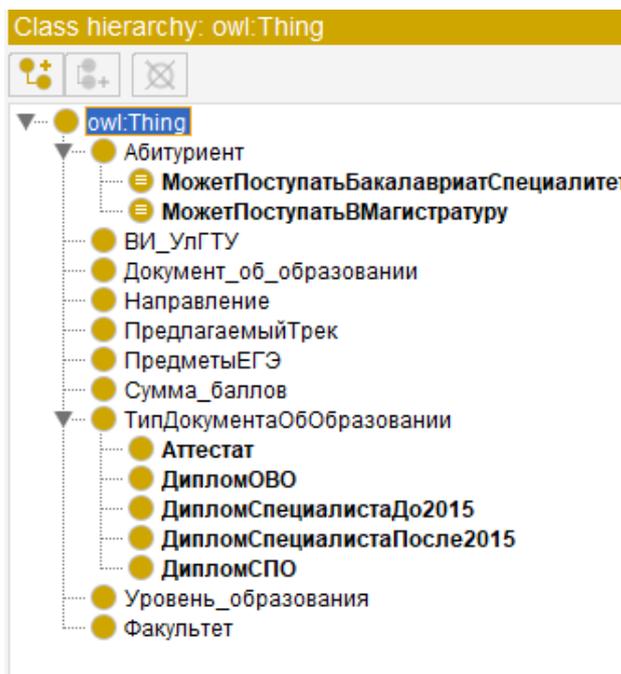


Рис. 3. Дерево классов онтологии.

В онтологии также существуют анонимные классы. Пример экзистенциального ограничения для анонимного класса «МожетПоступитьБакалавриатСпециалитет» представлен на рисунке 4.

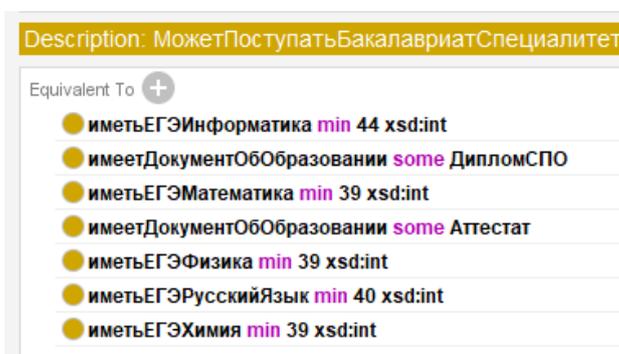


Рис. 4. Экзистенциальное ограничение для анонимного класса «МожетПоступитьБакалавриатСпециалитет».

Всего в данную онтологию экспертами было добавлено 21 SWRL-правило. Пример разработанного SWRL-правила представлен ниже:

$Абитуриент(?y) \wedge \text{МожетПоступитьБакалавриатСпециалитет}(?y) \wedge \text{имеетСуммуБаллов}(?y, ?p) \wedge \text{имеетДокументОбОбразовании}(?y, \text{Аттестат}) \wedge \text{имеетЖелаемыйФакультет}(?y, \text{ФакультетИнформационныхСистемИТехнологий}) \rightarrow \text{ПредлагаемыйТрек}(\text{Трек1})$

Данное SWRL правило описывает условия, при которых абитуриенту предлагается определенный трек обучения. По условиям правила:

1. $Абитуриент (?y)$ - объект, который является абитуриентом.
2. $МожетПоступитьБакалавриатСпециалитет(?y)$ - абитуриент может поступить на бакалавриат или специалитет.

3. *Имеет Сумму Баллов*(?*y*, ?*p*) - абитуриент имеет определенную сумму баллов.
4. *Имеет Документ Об Образовании*(?*y*, *Аттестат*) - абитуриент имеет документ об образовании в виде аттестата.
5. *Имеет Желаемый Факультет*(?*y*, *Факультет Информационных Систем И Технологий*) - абитуриент выбрал желаемый факультет.

Если все эти условия выполнены, то абитуриенту предлагается определенный трек обучения (*Предлагаемый Трек (Трек1)*).

Был проведен ряд экспериментов по применению разработанной базы знаний. Эксперименты проводились в рамках приемной кампании 2024 года в Ульяновском государственном техническом университете.

База знаний была апробирована техническим персоналом приемной комиссии. Алгоритм действия технического персонала заключался в следующем:

- абитуриент предоставляет информацию о документе, об образовании, имеющихся результатах экзаменов, а также о приоритетных направлениях для поступления;
- технический персонал вносит все полученные данные в соответствующие вкладки в редакторе онтологии. Для разработки обучающей онтологии использовали редактор Protege 5.1.
- запускается машина вывода для проверки согласованности и автоматического извлечения знания о возможности «быть рекомендованным к зачислению» (согласно документу об образовании и имеющимся результатам экзаменов), либо же абитуриент будет отнесен к подклассу «Не прошел контроль».
- экспертная система в соответствии с введенными данными абитуриента формирует набор треков, на которые рекомендуется подача заявления этого абитуриента.

Мониторинг консультаций абитуриентов осуществлялся в течение четырёх рабочих дней технического консультанта. За это время было проконсультировано 253 абитуриента, на что было затрачено 14 часов 45 минут 30 секунд, т.е. в среднем на предварительную работу по ориентации абитуриентов было потрачено 3,5 минуты.

Далее техническому специалисту была предоставлена разработанная база знаний для помощи в проведении консультаций. Замерялось время работы со следующими 253 абитуриентами. Работа заняла 10 часов 45 минут 9 секунд. Время обработки данных сократилось до 2,55 минут за счёт использования разработанной базы знаний.

Таким образом, экономия времени, без учёта времени, затраченного на разработку базы знаний, составила в среднем 27% для каждого абитуриента на этапе приоритизации направлений. Также использование базы знаний сэкономило 4 часа времени за 4 рабочих дня. Этот результат, безусловно, свидетельствует о высокой эффективности предлагаемого подхода.

4. ЗАКЛЮЧЕНИЕ

Таким образом, в рамках данной работы рассмотрена проблема определения приоритетности направлений при поступлении в вуз, проанализирована предметная область и проведен обзор текущего состояния исследований и разработок.

Для достижения цели исследования была спроектирована и построена экспертная система, основанная на интеграции OWL и SWRL, для определения приоритетности направлений при поступлении в вуз. Разработанная экспертная система позволила существенно сократить время абитуриента на определение приоритетности направлений при поступлении в вуз (на 27%).

Поскольку план приема ежегодно утверждается Министерством науки и высшего образования Российской Федерации и администрацией вуза, следует отметить, что основным преимуществом данной базы знаний является ее доступность для внесения изменений.

Основные недостатки/ограничения использования данной гибридной формы представления знаний:

- сложность составления онтологии. Необходимость привлечения экспертов на начальном этапе наполнения базы знаний;
- невозможность четкой формулировки некоторых взаимосвязей и аксиом.

Таким образом, дальнейшее развитие исследования может заключаться в расширении наборов правил и корректировке существующих. Также одним из направлений развития проекта является разработка методики автоматизированного расширения базы знаний за счёт обработки слабоструктурированных данных нормативной документации.

СПИСОК ЛИТЕРАТУРЫ

1. Butsyk, S.V., Kolbin, R.V., Ruzakov, A.A. The problem of the construction of an automated control system in a Russian higher-education institution. *Sci. Tech. Inf. Proc.* 43, 194–198 (2016). <https://doi.org/10.3103/S0147688216030084>
2. Ишоева В. А., Сорокина Ю. В., Горностаева Ж. В. РОЛЬ РАБОТЫ ПРИЕМНОЙ КОМИССИИ, КАК ЭЛЕМЕНТ ПРЕДПРОДАЖНОГО СЕРВИСА В СФЕРЕ ОБРАЗОВАТЕЛЬНЫХ УСЛУГ // Теория и практика современной науки. 2015. №2 (2).
3. Свердлов, Д. А. Автоматизация приемной кампании в Омский государственный университет путей сообщения на платформе «1С:Университет ПРОФ» // Информационные технологии и автоматизация управления: Материалы XI Всероссийской научно-практической конференции студентов, аспирантов, работников образования и промышленности, Омск, 29–30 мая 2020 г. – Омск: ОмГТУ, 2020.
4. Федина, Л. А. Актуальные аспекты автоматизации приемной кампании в вуз // Современные технологии хранения, обработки и анализа больших данных: Сборник научных трудов кафедры автоматизированных систем обработки информации и управления. – М. : РГУ имени А. Н. Косыгина (Технологии. Дизайн. Искусство), 2023. – С. 114–118.
5. Сапожников, Д. Е. Проблемы автоматизации приёмной кампании в российских вузах // XLVIII Огарёвские чтения : Материалы научной конференции. В 3 частях, Саранск, 06–13 декабря 2019 г. / Составитель А. В. Столяров. Ответственный за выпуск П. В. Сенин. Том Часть 1. – Саранск: Национальный исследовательский Мордовский государственный университет имени Н. П. Огарёва, 2020. – С. 466–471.
6. E. Yu. Kirpicheva, T. O. Makhalkina, and E. A. Rusakova, “Automation of the university admissions campaign in the context of a multitasking approach”, *Systems analysis in science and education*, issue. 3, pp. 26–37, Nov. 2023.
7. Moshkin V, Kurilo D, Yarushkina N. Integration of Fuzzy Ontologies and Neural Networks in the Detection of Time Series Anomalies. *Mathematics*. 2023; 11(5):1204. <https://doi.org/10.3390/math11051204>
8. Ye, P.N.J., Coyle, L., Dobson, S.: *Ontology-based models in pervasive computing systems*. *Knowl. Eng. Rev.* 22, 315–347 (2007)
9. Djakhdjakha, L., Boukara, D., Hemam, M., Boufaida, Z. (2020). An Extended Business Process Representation for Integrating IoT Based on SWRL/OWL. In: Hemanth, D., Kose, U. (eds) *Artificial Intelligence and Applied Mathematics in Engineering Problems*. ICAIAME 2019. *Lecture Notes on Data Engineering and Communications Technologies*, vol 43. Springer, Cham. https://doi.org/10.1007/978-3-030-36178-5_29
10. Zhai, Z., Martínez Ortega, J.F., Lucas Martínez, N., Castillejo, P.: A rule-based reasoner for underwater robots using OWL and SWRL. *Sensors* 18(10), 3481 (2018)
11. SWRL: A Semantic Web Rule Language Combining OWL and RuleML <https://www.w3.org/Submission/SWRL>

12. A. Filippov, V. Moshkin, A. Namestnikov, G. Guskov, M. Samokhvalov Approach to Translation of RDF/OWL-Ontology to the Graphic Knowledge Base of Intelligent Systems / Proceedings of the II International Scientific and Practical Conference "Fuzzy Technologies in the Industry – FTI 2018" // Ulyanovsk, Russia, 23-25 October, 2018. pp. 44-49
13. S. Gan, D. Shi, J. Zurada and J. Guan, "Construction of Learning Process Ontology and Application of SWRL Rules," 2023 16th International Conference on Advanced Computer Theory and Engineering (ICACTE), Hefei, China, 2023, pp. 1-9, doi: 10.1109/ICACTE59887.2023.10335228.
14. E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur and Y. Katz, "Pellet: A practical OWL-DL reasoner Journal of Web Semantics, vol. 5, no. 2, pp. 51-53, Jun. 2007.
15. Shalfeeva E.A., Gribova V.V. The Issues of Creation of Machine-Understandable Smart Standards Based on Knowledge Graphs // Informatics and Automation. - 2024. - Vol. 23. - N. 4. - pp. 969-988.
16. B. I. Geltser, T. A. Gorbach, V. V. Gribova, O. V. Karpik, E. S. Klyshinsky, N. A. Kochetkova, D. B. Okun, M. V. Petryaeva, K. I. Shakhgeldyan, "Ontology-based syntactic analysis of domain-specific texts", Proceedings of ISP RAS, 33:4 (2021), 99–116
17. Moshkin, V., Yarushkina, N., Kurilo, D., Dyrnochkin, A. (2023). Hybrid Approach to Time Series Anomaly Detection Using LSTM Networks and Ontology. In: Kovalev, S., Sukhanov, A., Akperov, I., Ozdemir, S. (eds) Proceedings of the Sixth International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'22). IITI 2022. Lecture Notes in Networks and Systems, vol 566. Springer, Cham. pp 181–191 https://doi.org/10.1007/978-3-031-19620-1_18
18. B. E. Idrissi, S. Baina, A. Mamouny and M. Elmaallamac, "RDF/OWL storage and management in relational database management systems: A comparative study Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 9, pp. 7604-7620, Oct. 2022.
19. Qu J. A Review on the Application of Knowledge Graph Technology in the Medical Field // Scientific Programming. 2022. vol. 2022. 12 p.
20. Novais P., Oliveira T., Satoh K., Neves J. The Role of Ontologies and Decision Frameworks in Computer-Interpretable Guideline Execution // Synergies between Knowledge Engineering and Software Engineering. 2018. vol. 626. pp. 197–216.

Development of a hybrid ontological model of a knowledge base for applicants to select a field of study

V. S. Moshkin, I. A. Andreev, E. S. Vaganova, A. A. Reitz

The article presents an approach to building a knowledge base of an expert system that provides support for applicants' decision-making at the stage of pre-processing the provided data and prioritizing the areas. The knowledge base model includes a hybridization of the subject OWL ontology and a set of SWRL rules. The article provides examples of SWRL rules that are included in the ontology. Based on the integration of rules and OWL ontology, a logical conclusion is made. The paper also presents the results of using the developed knowledge base in the work of the technical staff of the Admissions Committee as part of the 2024 University Admissions Campaign. The experiments showed the effectiveness of using the developed knowledge base, the use of which at the stage of data pre-processing and prioritizing the areas allowed to reduce time costs by 27%.

KEYWORDS: inference, ontology, knowledge base, admissions committee, OWL, SWRL.