

Обзор и анализ представлений графов знаний

^{1*}**РАМАЗАНОВА Валия Сапкиреевна**, докторант, ravasa@mail.ru,

¹**САМБЕТБАЕВА Мадина Аралбаевна**, PhD, ассоциированный профессор, madina_jgtu@mail.ru,

²**ЗАГОРУЛЬКО Юрий Алексеевич**, к.т.н., зав. лабораторией, zagor@iis.nsk.su,

¹НАО «Евразийский национальный университет имени Л.Н. Гумилева», ул. Сатпаева, 2, Астана, Казахстан,

²Институт систем информатики имени А.П. Ершова СО РАН, пр. академика Лаврентьева, 6, Новосибирск, Россия,

*автор-корреспондент.

Аннотация. С развитием искусственного интеллекта графы знаний становятся все более значимыми в информационных системах. В статье проводится обзор определений и представлений графов знаний у разных авторов. Цель исследования заключается в формировании представления о графах знаний для последующих исследований. Для определения сути понятия графа знаний рассматриваются его история и эволюция, а также проводятся обзор и сравнительный анализ существующих определений за последние 10 лет с применением методов индукции и дедукции. Также выделяются основные акценты, на которые опирались исследователи в своих работах. В результате методами анализа и синтеза выводится авторское представление о графах знаний.

Ключевые слова: графы знаний, базы знаний, классификация баз знаний, представление графа знаний, обзор определений графов знаний, анализ представлений графов знаний, онтологии, структура графов знаний, стандарты графов знаний.

Введение

В настоящее время в условиях «посткризиса» одной из первостепенных задач современного общества стала разработка и применение новейших технологий в образовании. Решение этой задачи в значительной

степени зависит от грамотного внедрения искусственного интеллекта в деятельность образовательных учреждений.

С бурным развитием искусственного интеллекта все большее значение и применение в современных информационных

системах получают графы знаний. Они предоставляют мощный и гибкий инструмент для представления, анализа и использования информации в интеллектуальных системах. Причиной их использования большим количеством исследователей служит то, что графы знаний обладают богатой семантической информацией и четкой логической структурой. В международном научном сообществе графы знаний рассматривают как новый носитель характеристики человека и познания мира [1]. В целом графы знаний позволяют повысить эффективность обработки данных, улучшить понимание контекста и отношений, а также обеспечить более точные и персонализированные рекомендации и выводы.

Исходя из вышесказанного, целью данной работы являются обзор и анализ определенных (представлений) графа знаний, приводимых разными авторами в их оригинальной трактовке, и формирование представления о графах знаний для использования в дальнейшем исследовании. Как описано в работе [2], опасность многоязычия была замечена еще в древневавилонское время. В отличие от гуманитарного подхода, техническое и компьютерное направление требует избавления от многозначности, то есть приведения к единым стандартам, так как это может привести к шуму и искажениям при обмене информацией. Следовательно, избавление от многозначности является важным направлением в разработке формальных методов для представления информации и знаний. Для определения сути понятия графа знаний необходимо рассмотреть историю его зарождения и эволюцию.

Краткая история графов знаний

Кратко проследивая историю становления графов знаний, началом можно считать 60-е годы XX века, когда в Стэнфорде началась работа с экспертными системами и встал вопрос о переходе от традиционных алгоритмов рассуждений к алгоритмам, основанных на знаниях. Данный переход вызвал большую волну исследований в последующие десятилетия. В 2004 году Консорциумом World Wide Web Consortium (W3C) была разработана модель среды описания ресурса Resource Description Framework (RDF) для представления данных и метаданных. RDF представляет утверждения о ресурсах в виде, пригодном для машинной обработки. Основной моделью данных RDF является тройное выражение-тройка «субъект-предикат-объект», которое указывает, что две сущности (субъект и объект) связаны через отношение (предикат). RDF предоставляет средства для построения информационных моделей, но не касается семантики описы-

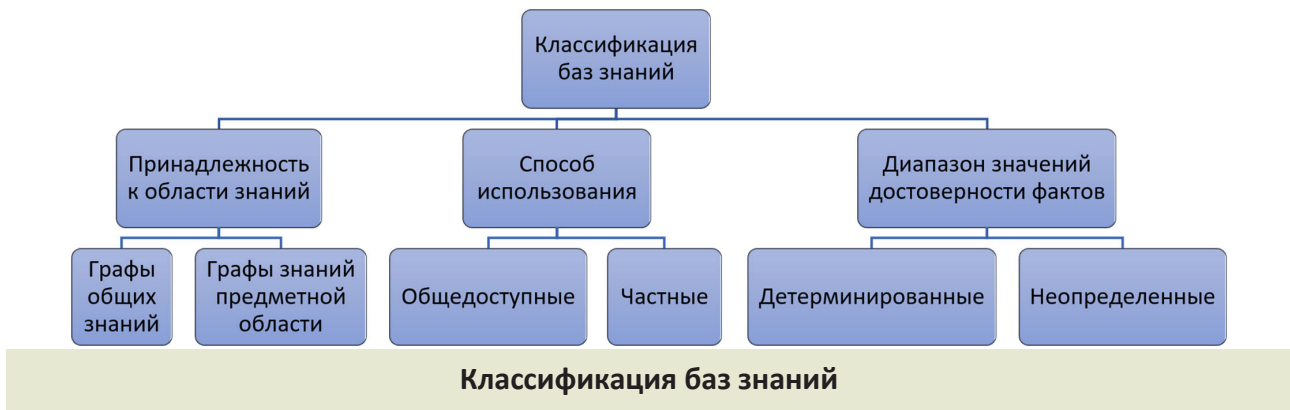
ваемого. Взятый в отдельности граф RDF можно понимать только как граф, поэтому для решения этой проблемы W3C предложил схему среды описания ресурсов Resource Description Framework Schema (RDFS) и язык веб-онтологий Web Ontology Language (OWL). RDFS и OWL расширяют RDF, добавляя общие предопределенные словари на уровне схемы, чтобы они могли представлять абстрактные отношения, такие как классы и свойства.

Основные принципы графов знаний широко применялись при создании баз знаний, таких как: DBPedia (2007, Германия) – предназначалась для извлечения структурированной информации из Wikipedia; FreeBase (2007, США) – онлайн-коллекция структурированных метаданных из разных wiki источников; YAGO (2008, Германия) – мультязычная база знаний из Wikipedia, Wordnet и Geonames; NELL (2010, США) – разработка средств ответов на вопросы, заданные пользователями на естественном языке. В 2012 году компания Google предложила современную технологию для улучшения семантического поиска и сбора информации, дав ей название Knowledge Graph («граф знаний»). Эта технология в качестве основного источника информации использовала приобретенную Google базу знаний FreeBase. Вдохновленные технологией Google, ведущие мировые компании разработали большое количество различных баз знаний, которые в настоящее время объединены в облако связанных данных – Linked Open Data Cloud (LOD Cloud). Так как графы знаний образуют семантическую сеть, которая хранится в базах знаний, и основные принципы графов знаний использовались при создании многих баз знаний, мы приходим к выводу, что база знаний, использующая графовую структуру для организации и представления информации, является практической реализацией графа знаний для решения определенных практических задач в конкретной предметной области, таких как создание рекомендательных систем, семантический поиск и анализ текста, поиск похожих элементов и другие задачи.

Классификация баз знаний

Ввиду большого разнообразия графов знаний, выделяют несколько классификаций, которые приведены на рисунке.

Во-первых, в зависимости от принадлежности к области знаний графы знаний делят на два типа: 1) графы общих знаний, которые относятся к подходам на основе словарей, включая WordNet, Yago, DBpedia, Zhishi.me; 2) графы знаний предметной области, включающие в себя графы из различных предметных областей, например, финан-



совый граф знаний Ant Affair, граф знаний предприятий Qichacha, граф биологии и медицинских знаний Bio2RDF и UMLS и так далее [3].

Во-вторых, в зависимости от способа использования, графы знаний делятся на два типа: общедоступные; частные, которые обычно создаются для внутреннего использования внутри команды. В-третьих, в соответствии с диапазоном значений достоверности фактов базы знаний делят на две категории: детерминированные графы знаний, такие как Freebase и YAGO, достоверность факта которых равна 1 или 0 (1 означает, что факт встречается в графе знаний, 0 – нет); неопределенные графы знаний или недетерминированные графы знаний, такие как NELL и ConceptNet, в которых каждому факту соответствует оценка достоверности от 0 до 1, представляющая собой степень доверия факта к истине.

Обзор и анализ определений (представлений) графов знаний

Замечание. В литературе, в оригинальных текстах статей приводится большое количество определений графов знаний, но в данной статье в определении графов знаний основной акцент мы будем делать на представлении графов знаний.

Исследователи изучают два основных способа представления графов знаний: символьная форма и векторная форма. Символьное представление использует символы для записи фактов в виде онтологических представлений, основанных на формальной логике и семантике, в то время как векторное представление использует статистические методы для проецирования сущностей и отношений в векторное пространство для минимизации расстояния между соседними сущностями в этом пространстве и упрощения обработки при сохранении внутренней структуры. Нельзя сказать, что одно представление лучше другого: символьные представления очень полезны в задачах интеграции данных, тогда как векторные

представления часто используются для обработки данных [4].

Рассмотрим определения графов знаний в статьях, приводимых авторами за разные годы. Обзор определений приведем после 2012 года, когда графы знаний получили общее признание.

В 2013 году Дж. Пуджара [5] в своей работе при определении графов знаний акцентирует свое внимание на извлечении информации: «...которые используют различные методы для извлечения новых знаний в виде фактов из Интернета. Эти факты взаимосвязаны, ..., недавно это извлеченное знание стало называться графом знаний». В своей работе автор выполняет преобразование сущностей, коллективную классификацию и прогнозирование связей, а также применяет онтологические правила к графу знаний, называя этот процесс идентификацией графа знаний.

Х. Полхейм [6] в 2016 году формулирует признаки графов знаний: «Граф знаний: 1) в основном описывает сущности реального мира и их взаимосвязи, организованные в виде графа; 2) определяет возможные классы и отношения сущностей в схеме; 3) позволяет потенциально связывать произвольные объекты друг с другом; 4) охватывает различные тематические области». Его работа опирается на онтологическое представление графов знаний, рассмотрены методы для улучшения существующих графов знаний.

В том же году М. Фарбер [7] и другие формально определяют граф знаний с точки зрения модели данных RDF (Resource Description Framework): «Мы определяем граф знаний как граф RDF. Граф RDF состоит из набора троек RDF, где каждая тройка RDF (s, p, o) представляет собой упорядоченный набор следующих терминов RDF: субъект $s \in U \cup B$, предикат $p \in U$, и объект $o \in B \cup L$. Термин RDF является либо URI $u \in U$, пустая вершина $b \in B$ или литерал $l \in L$. U, B , и L попарно не пересекаются».

Они анализируют и сравнивают пять известных баз знаний DBPedia, Freebase, OpenCyc, Wikidata, и YAGO и оценивают их по нескольким критериям качества данных. Его определение графа знаний фокусируется на обработке структуры данных.

В 2016 году Л. Эрлингер [8] проводит анализ определений и предлагает свое определение графа знаний, которое строится на понятиях базы знаний и онтологии: «Граф знаний собирает и интегрирует информацию в онтологии и применяет различные механизмы вывода для получения новых знаний». Они считают, что графы знаний являются не новой технологией, а всего лишь модным словом, и скорее похожи на абстрактную, нежели математическую структуру.

В 2017 году К. Ванг [9] и другие рассматривают понятие графов знаний в контексте векторного представления: «Граф знаний – это мультиреляционный граф, состоящий из сущностей (узлов) и отношений (разные типы ребер). Каждое ребро представлено в виде тройки в форме (головная сущность, отношение, хвостовая сущность), также называемой фактом, указывающим на то, что две сущности связаны определенным отношением... символьная природа таких троек обычно затрудняет их обработку». В данном определении акцент сделан на разных видах отношений, которые формируют разные факты, формальная обработка которых вызывает трудности при символьном представлении.

А. Хоган [10] и другие в 2021 году в своей публикации дают всестороннее представление о графах знаний: «Мы называем граф знаний графом данных, потенциально дополненным представлениями схемы, идентичности, контекста, онтологий и/или правил». Помимо общего понятия, они выделяют четыре вида моделей графов знаний: 1. ориентированный граф с помеченными ребрами; 2. гетерогенный граф с помеченными ребрами; 3. граф свойств; 4. комбинированные графы знаний; и дают им формальные определения. Такая формализация определений в зависимости от модели данных, вызвана необходимостью повысить эффективность процессов обработки, передачи и хранения графов знаний как большого массива данных.

В статье [11] от 2021 года Ш. Цзи с соавторами рассматривает графы знаний и базы знаний как синонимы, когда речь идет о формальной семантике. Они проводят обзор графов знаний и выводят следующее определение: «Мы определяем граф знаний как $G = \{E, R, F\}$, где E , R и F являются множеством сущностей, отношений и фактов, со-

ответственно. Факт обозначается как тройка $(h, r, t) \in F$. Эта тройка состоит из головы, отношения и хвоста». В работе проводится всесторонний обзор векторного представления графов знаний с полномасштабным систематическим обзором.

Дж. Барраса [12] в том же 2021 году исследует графы знаний с фокусом на модель данных: «Графы знаний – это особый тип графов с акцентом на контекстуальное понимание. Графы знаний связаны наборами фактов, которые описывают сущности, события или предметы реального мира и их взаимосвязи в понятном человеку и машине формате». Он утверждает, что модель графа свойств является наиболее популярной моделью для современных графовых баз данных и, как следствие, популярным методом обработки графов знаний.

В. Чаудхри [13] и другие в 2022 году рассматривают синтез различных точек зрения на графы знаний: «Граф знаний – это ориентированный помеченный граф, в котором понятия, специфичные для предметной области, ассоциируются с узлами и ребрами. Узел может представлять любую сущность реального мира, например людей, компании и компьютеры. Метка ребра фиксирует интересные отношения между двумя узлами». Указывают, что в ранних системах искусственного интеллекта с использованием графов знаний, акцент делался на сложных логических рассуждениях, а современные графы знаний сфокусированы на сборе фактов для поиска и аналитики. Также представляют несколько контрастирующих точек зрения для лучшего понимания графов знаний, таких как символьное представление в сравнении с векторным представлением, автоматизированное построение графов знаний в сравнении с ручным построением, сравнение большой и малой степени семантического смысла. В качестве вывода они заключают, что нет необходимости выделять среди всех рассмотренных подходов лучший, а необходимо предоставить возможность для параллельного развития каждого из них, что послужит идеальной платформой для полного раскрытия потенциала графов знаний и оказания максимального общественного и научного воздействия.

Статья Л. Тиан [14] от 2022 года оценивает будущее развитие графов знаний: «Граф знаний описывает понятия, сущности и отношения объективного мира в виде графов. Он может организовывать, управлять и понимать массивную информацию способом, близким к человеческому когнитивному мышлению». Они рассматривают архитектуру графов знаний как: «Граф знаний может быть выражен в форме троек $G = \{E, R, F\}$.

Е представляет множество сущностей $\{e_1, e_2, \dots, e_i\}$. Сущность e является простейшим элементом в графе знаний. R представляет множество отношений $\{r_1, r_2, \dots, r_j\}$. Отношение R представляет собой ребро определенной связи между двумя разными сущностями. F представляет множество фактов $\{f_1, f_2, \dots, f_3\}$ и каждый факт описывается как тройка $(h, r, t) \in F$, где h, r и t представляет собой головную сущность, отношение и хвостовую сущность, соответственно». Его определение близко к определению Shaoxiong Ji [11]. В работе представлен обзор графов знаний, включая их представление, хранение и основные технологии. В частности, обобщается несколько типов подходов к осмыслению графов знаний: методы, основанные на логических правилах, на основе векторных представлений и на основе нейронных сетей. Кроме того, в этой статье анализируются методы представления гиперграфов знаний.

В заключение обзора представлений о графах знаний укажем, что в январе 2023 года был опубликован стандарт «IEEE P2807 Структура графов знаний» [15], в котором были учтены исследования, посвященные графам знаний. Стандарт определяет граф знаний как: «Совокупность элементов знаний и их отношений, описанных в структурированной форме... Элемент знаний – это общее название для терминов сущности, концепции (типа сущности), атрибута, отношения, типа отношения, события или правила». Определение из стандарта является более общим, чем определения, приводимые разными исследователями, что означает, что все приведенные выше определения являются подмножеством стандартизированного.

Сведем представления о графах знаний в таблицу с указанием акцента на подход, применяемый для описания графов знаний.

Как мы видим из таблицы, разные авторы описывают графы знаний с разных точек зрения, а именно с акцентом на онтологии для их построения и интеграции, с акцентом на модели данных для более удобной обработки данных, например поиска, и с уклоном на векторное представление, например, для применения в задачах классификации или кластеризации.

Наконец, новый стандарт позволяет на концептуальном уровне унифицировать онтологическое представление знаний и структурную обработку данных, что в свою очередь позволяет различным системам и приложениям взаимодействовать и обмениваться данными более эффективно.

Определение графа знаний

Пользуясь теоретическими методами анализа и синтеза для нашего дальнейшего исследования, примем следующее определение:

Граф знаний – это мультисущностный и мультиреляционный ориентированный помеченный граф, состоящий из разного типа сущностей и отношений, дополненный представлениями схемы, идентичности, контекста, онтологий и/или правил.

Сущностью может быть человек, предмет, дата, концепция или любой другой материальный объект или абстрактное понятие.

Схема представляет собой семантическую схему, которая позволяет определить значение высокоуровневых терминов (словарь или терминологию), используемых в графе, что облегчает рассмотрение графов с использованием этих терминов, например RDFS.

Идентичность – это устранение неоднозначности узла по отношению к внешнему источнику, например, наличие идентификатора Internationalised Resource Identifiers (IRI).

Подходы, применяемые для описания графов знаний

Автор	Символьное представление		Векторное представление
	Онтологии	Модель данных	
Дж. Пуджара, 2013	+		
Х. Полхейм, 2016	+		
М. Фарбер, 2016		+	
Л. Эрлингер, 2016	+		
К. Ванг, 2017			+
А. Хоган, 2021	+	+	
Ш. Цзи, 2021			+
Дж. Барраса, 2021		+	
В. Чаудхри, 2022	+	+	+
Л. Тиан, 2022		+	+
IEEE P2807, 2023	+	+	

Контекст – окружение, в котором некоторые данные считаются истинными, например, временной контекст.

Онтология – формальное представление концепций и терминов в используемой области, например, с помощью языка OWL.

Правило – это правила вывода для обеспечения автоматизированного доступа к знаниям в виде запросов, например на языке SPARQL Inferencing Notation (SPIN).

Заключение

Таким образом, среди определений графов знаний наблюдается большое разнообразие. База знаний, использующая графовую структуру для организации и представления информации, является практической реализацией графа знаний для решения практических задач в конкретной предметной области, например, создание рекомендательных систем. Многие авторы считают понятия графов знаний и баз знаний синонимами, но общим является то, что графы знаний рассматриваются как особый

тип графов с акцентом на контекстуальное понимание. Наиболее распространенное определение графа знаний состоит в том, что это ориентированный граф, в котором и ребра, и узлы имеют метки. К современным графам знаний относят и графы свойств, которые облегчают хранение и обработку графов знаний в графовых базах данных. Публикация первого стандарта делает попытку унифицировать представление графов знаний в приложениях для более эффективного обмена данными и взаимодействия между всеми заинтересованными сторонами. В новом стандарте пришли к выводу, что для каждого конкретного случая используется определенное представление графа знаний, подходящее для решения определенного класса задач, поэтому стандартизованное определение очень обобщено. Пользуясь обзором и анализом представлений графов знаний у разных авторов и в стандарте, мы описываем свое представление графа знаний, на которое будем опираться в дальнейшем исследовании.

СПИСОК ЛИТЕРАТУРЫ

1. Pu, L., Xin Wang, Hui Liang, Suzhi Zhang, Yazhou Zhang, Yuncheng Jiang, & Yong Tang. (2022). A fuzzy semantic representation and reasoning model for multiple associative predicates in knowledge graph. *Information Sciences*, 599, 208-230. <https://doi.org/10.1016/j.ins.2022.03.076>
2. Федотов А.М., Федотова О.А., Самбетбаева М.А. Информационные системы: модели и технологии. Отв. ред. В.Б. Барахнин. Новосибирск: ИПЦ НГУ. 2019.
3. Chen, L., Liu, D., Yang, J., Jiang, M., Liu, S., & Wang, Y. (2022). Construction and application of COVID-19 infectors activity information knowledge graph. *Computers in Biology and Medicine*, 148, 105908. <https://doi.org/10.1016/j.compbimed.2022.105908>
4. Галкин М., Сафронов В., Иванов С., Муромцев Д. Курс лекций по графам знаний. <https://migalkin.github.io/kgcourse2021/>
5. Pujara, J., Miao, H., Getoor, L., Cohen, W. (2013). Knowledge Graph Identification. In: et al. *The Semantic Web – ISWC 2013*. ISWC 2013. Lecture Notes in Computer Science, vol 8218. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-41335-3_34
6. Paulheim, H. (2016). Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web*, 8, 489-508. <https://doi.org/10.3233/SW-160218>
7. Färber, M., Bartscherer, F., Menne, C., & Rettinger, A. (2018). Linked data quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. *Semantic Web Journal*, 9 (1), 77-129. <https://doi.org/10.3233/SW-170275>
8. Ehrlinger, L., & Wöß, W. (2016). Towards a Definition of Knowledge Graphs. In *Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems – SEMANTiCS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCESS'16) co-located with the 12th International Conference on Semantic Systems (SEMANTiCS 2016)*, Leipzig, Germany, September 12-15, 2016 (CEUR Workshop Proceedings), Michael Martin, Martí Cuquet, and Erwin Folmer (Eds.), Vol. 1695. Sun SITE Central Europe (CEUR), 4. <http://ceur-ws.org/Vol-1695/paper4.pdf>
9. Wang, Q., Mao, Z., Wang, B., & Guo, L. (2017). Knowledge Graph Embedding: A Survey of Approaches and Applications. *IEEE Transactions on Knowledge and Data Engineering*, 29 (12), 2724-2743. <https://doi.org/10.1109/TKDE.2017.2754499>
10. Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutierrez, C., Kirrane, S., Labra Gayo, J.E., Navigli, R., Neumaier, S., Ngonga Ngomo, A.-C., Polleres, A., Rashid, S.M., Rula, A., Schmelzeisen, L.,

- Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge Graphs. ACM Computing Surveys, 54 (4), Article 71. <https://doi.org/10.1145/3447772>
11. Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P.S. (2022). A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. IEEE Transactions on Neural Networks and Learning Systems, 33 (2), 494-514. <https://doi.org/10.1109/TNNLS.2021.3070843>
 12. Barrasa, J., Hodler, A.E., & Webber, J. (2021). Knowledge Graphs. O'Reilly Media.
 13. Chaudhri, V., Baru, C., Chittar, N., Dong, X., Genesereth, M., Hendler, J., Kalyanpur, A., Lenat, D., Sequeda, J., Vrandečić, D., & Wang, K. (2022). Knowledge Graphs: Introduction, History and, Perspectives. AI Magazine, 43 (1), 17-29. <https://doi.org/10.1002/aaai.12033>
 14. Tian, L., Zhou, X., Wu, Y.P., Zhou, W.T., Zhang, J.H., & Zhang, T.S. (2022). Knowledge graph and knowledge reasoning: A systematic review. Journal of Electronic Science and Technology, 20 (2), 100159. <https://doi.org/10.1016/j.jnlest.2022.100159>
 15. IEEE. (2023). IEEE Standard for Framework of Knowledge Graphs. IEEE Std 2807-2022, (1), 1-52. <https://doi.org/10.1109/IEEESTD.2022.10017167>

Білім графтарының көріністеріне шолу және талдау

¹***РАМАЗАНОВА Валия Сапкиреевна**, докторант, ravasa@mail.ru,

¹**САМБЕТБАЕВА Мадина Аралбаевна**, PhD, қауымдастырылған профессор, madina_jgtu@mail.ru,

²**ЗАГОРУЛЬКО Юрий Алексеевич**, т.ғ.к., зертхана меңгерушісі, zagor@iis.nsk.su,

¹«Л.Н. Гумилев атындағы Еуразия ұлттық университеті» КеАҚ, Сәтпаев көшесі, 2, Астана, Қазақстан,

²А.П. Ершов атындағы жүйелік информатика институты, РФА СБ, академик Лаврентьев даңғылы, 6, Новосибирск, Ресей,

*автор-корреспондент.

Аңдатпа. Жасанды интеллекттің дамуымен ақпараттық жүйелерде білім графтарының маңыздылығы артып келеді. Мақалада әртүрлі авторлардың білім графтарының анықтамалары мен ұсыныстарына шолу жасалған. Зерттеудің мақсаты – кейінгі зерттеулер үшін білім графтары туралы түсінік қалыптастыру. Білім графтары ұғымының мәнін анықтау үшін оның тарихы мен эволюциясы қарастырылады, сонымен қатар индукция мен дедукция әдістерін қолдана отырып, соңғы 10 жылдағы қолданыстағы анықтамаларға шолу және салыстырмалы талдау жасалады. Зерттеушілер өз жұмыстарында сүйенген негізгі акценттері де ерекшеленеді. Анализ және синтез әдістерінің нәтижесінде білім графтары туралы авторлық түсінік алынады.

Кілт сөздер: білім графтары, білім базалары, білім базаларының жіктелуі, білім графтарын ұсыну, білім графтарының анықтамаларына шолу, білім графтарының көріністерін талдау, онтология, білім графтарының құрылымы, білім графтарының стандарттары.

Overview and Analysis of Knowledge Graph Representations

¹***RAMAZANOVA Valiya**, Doctoral Student, ravasa@mail.ru,

¹**SAMBETBAEVA Madina**, PhD, Associate Professor, madina_jgtu@mail.ru,

²**ZAGORULKO Yuriy**, Cand. of Tech. Sci., Head of Laboratory, zagor@iis.nsk.su

¹NPJSC «L.N. Gumilyov Eurasian National University», Satpayev Street, 2, Astana, Kazakhstan,

²A.P. Ershov Institute of Informatics Systems, SB RAS, Academic Lavrentiev Avenue, 6, Novosibirsk, Russia,

*corresponding author.

Abstract. With the development of artificial intelligence, knowledge graphs are becoming increasingly important in information systems. The article reviews the definitions and representations of knowledge graphs by different authors. The purpose of the study is to form an idea of knowledge graphs for subsequent research. To determine the essence of the concept of the knowledge graph, its history and evolution are considered, as well as a review and com-

parative analysis of existing definitions over the past 10 years using induction and deduction methods. The main accents on which the researchers relied in their works are also highlighted. As a result, the author's idea of knowledge graphs is derived by the analysis and synthesis methods.

Keywords: *knowledge graphs, knowledge bases, classification of knowledge bases, representation of knowledge graphs, overview of knowledge graph definitions, analysis of knowledge graph representations, ontologies, knowledge graph structure, knowledge graph standards.*

REFERENCES

1. Pu, L., Xin Wang, Hui Liang, Suzhi Zhang, Yazhou Zhang, Yuncheng Jiang, & Yong Tang. (2022). A fuzzy semantic representation and reasoning model for multiple associative predicates in knowledge graph. *Information Sciences*, 599, 208-230. <https://doi.org/10.1016/j.ins.2022.03.076>
2. Fedotov, A.M., Fedotova, O.A., Sambetbaeva, M.A. (2019). *Informatsionnye sistemy: modeli i tekhnologii*. Otv. red. V.B. Barakhnin. Novosibirsk: IPC NGU.
3. Chen, L., Liu, D., Yang, J., Jiang, M., Liu, S., & Wang, Y. (2022). Construction and application of COVID-19 infectors activity information knowledge graph. *Computers in Biology and Medicine*, 148, 105908. <https://doi.org/10.1016/j.compbiomed.2022.105908>
4. Galkin M., Safronov V., Ivanov S., Muromtsev D. Kurs lektzii po grafam znanii. <https://migalkin.github.io/kgcourse2021/>
5. Pujara, J., Miao, H., Getoor, L., Cohen, W. (2013). Knowledge Graph Identification. In: et al. *The Semantic Web – ISWC 2013*. ISWC 2013. Lecture Notes in Computer Science, vol 8218. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-41335-3_34
6. Paulheim, H. (2016). Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web*, 8, 489-508. <https://doi.org/10.3233/SW-160218>
7. Färber, M., Bartscherer, F., Menne, C., & Rettinger, A. (2018). Linked data quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO. *Semantic Web Journal*, 9 (1), 77-129. <https://doi.org/10.3233/SW-170275>
8. Ehrlinger, L., & Wöß, W. (2016). Towards a Definition of Knowledge Graphs. In *Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems – SEMANTiCS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCESS'16) co-located with the 12th International Conference on Semantic Systems (SEMANTiCS 2016)*, Leipzig, Germany, September 12-15, 2016 (CEUR Workshop Proceedings), Michael Martin, Martí Cuquet, and Erwin Folmer (Eds.), Vol. 1695. Sun SITE Central Europe (CEUR), 4. <http://ceur-ws.org/Vol-1695/paper4.pdf>
9. Wang, Q., Mao, Z., Wang, B., & Guo, L. (2017). Knowledge Graph Embedding: A Survey of Approaches and Applications. *IEEE Transactions on Knowledge and Data Engineering*, 29 (12), 2724-2743. <https://doi.org/10.1109/TKDE.2017.2754499>
10. Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutierrez, C., Kirrane, S., Labra Gayo, J.E., Navigli, R., Neumaier, S., Ngonga Ngomo, A.-C., Polleres, A., Rashid, S.M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge Graphs. *ACM Computing Surveys*, 54 (4), Article 71. <https://doi.org/10.1145/3447772>
11. Ji, S., Pan, S., Cambria, E., Marttinen, P., & Yu, P.S. (2022). A Survey on Knowledge Graphs: Representation, Acquisition, and Applications. *IEEE Transactions on Neural Networks and Learning Systems*, 33 (2), 494-514. <https://doi.org/10.1109/TNNLS.2021.3070843>
12. Barrasa, J., Hodler, A.E., & Webber, J. (2021). *Knowledge Graphs*. O'Reilly Media.
13. Chaudhri, V., Baru, C., Chittar, N., Dong, X., Genesereth, M., Hendler, J., Kalyanpur, A., Lenat, D., Sequeda, J., Vrandečić, D., & Wang, K. (2022). Knowledge Graphs: Introduction, History and, Perspectives. *AI Magazine*, 43 (1), 17-29. <https://doi.org/10.1002/aaai.12033>
14. Tian, L., Zhou, X., Wu, Y.P., Zhou, W.T., Zhang, J.H., & Zhang, T.S. (2022). Knowledge graph and knowledge reasoning: A systematic review. *Journal of Electronic Science and Technology*, 20 (2), 100159. <https://doi.org/10.1016/j.jnlest.2022.100159>
15. IEEE. (2023). IEEE Standard for Framework of Knowledge Graphs. *IEEE Std 2807-2022*, (1), 1-52. <https://doi.org/10.1109/IEEESTD.2022.10017167>