РЕЦЕНЗИЯ

официального рецензента на диссертационную работу Оспан Әсел Ғалымжанқызы
на тему «Разработка геоинформационных моделей с использованием онтологического подхода для извлечения знаний из гетерогенных источников», представленную на соискание степени доктора философии (PhD)
по специальности «8D06105 – Системная инженерия»

№ п/п	Критерии	Соответствие критериям (необходимо отметить один из вариантов ответа)	Обоснование позиции официального рецензента
1.	Тема диссертации (на дату ее утверждения) соответствует направлениям развития науки и/или государственным программам	1.1 Соответствие приоритетным направлениям развития науки или государственным программам:	Соответствует приоритетному направлению в области информационно-коммуникационных технологий.
		1) Диссертация выполнена в рамках проекта или целевой программы, финансируемого(ой) из государственного бюджета (указать название и номер проекта или программы); 2) Диссертация выполнена в рамках другой государственной программы (указать название программы); 3) Диссертация соответствует приоритетному направлению развития науки, утвержденному Высшей научно-технической комиссией при Правительстве Республики Казахстан (указать направление).	Диссертационная работа выполнена в рамках научно-исследовательских проектов МОН РК (источник финансирования Комитет науки МОН РК): - грантового финансирования (ГФ) AP09261344 «Разработка методов автоматического извлечения геопространственных объектов из гетерогенных источников для информационного обеспечения геоинформационных систем» в 2021–2023 годы; - программно-целевого финансирования (ПЦФ) BR24993001 «Создание большой языковой модели (LLM) для поддержки казахского языка и технологического прогресса» в 2024–2026 годы.
2.	Важность для науки	Работа вносит/не вносит существенный вклад в науку, а ее важность хорошо раскрыта/не раскрыта	Работа вносит значительный вклад в науку, ее важность раскрыта в основных разделах диссертации. Достоверность полученных результатов подтверждена проведёнными вычислительными экспериментами, сравнительным анализом с существующими зарубежными методами и апробацией на реальных данных. В рамках проведённых исследований разработаны методы и модели извлечения и интеграции знаний из гетерогенных источников на основе онтологического

			подхода для водных ресурсов Казахстана. Особое внимание уделено построению гидроонтологии, разработке итеративного метода STI для извлечения семантических триплетов, а также созданию геоинформационной модели, обеспечивающей интероперабельность данных и поддержку принятия решений в области управления водными ресурсами.
3.	Принцип самостоятельности	Уровень самостоятельности: 1) высокий; 2) средний; 3) низкий; 4) самостоятельности нет.	Знакомство с диссертационной работой и научными публикациями соискателя позволяет оценить высокий уровень самостоятельности при выполнении исследования, что особенно проявляется при разработке методологических и программных решений, а также при проведении вычислительных экспериментов. Результаты получены с использованием авторского итеративного метода STI, методов семантического анализа таблиц и онтологического моделирования, а также реализации сквозного конвейера обработки данных от гетерогенных источников до формирования графа знаний и выполнения SPARQL-запросов. Соискателем самостоятельно разработаны модули гидроонтологии и проведён сравнительный анализ с современными зарубежными методами. По теме диссертации опубликованы 6 научных работ и получено свидетельство о государственной регистрации прав на объекты авторского права (программное обеспечение).
4.	Принцип внутреннего единства	4.1 Обоснование актуальности диссертации: 1) обоснована; 2) частично обоснована; 3) не обоснована.	Диссертационная работа представляет собой значимый вклад в решение актуальной научной проблемы интеграции гетерогенных данных о водных ресурсах Республики Казахстан и формирования единой семантически согласованной базы знаний. Актуальность исследования обоснована отсутствием интегрированной системы, обеспечивающей комплексный анализ гидрологических, экологических и социально-экономических показателей, что затрудняет принятие научно обоснованных управленческих решений. Предложенный онтологический подход и разработка геоинформационной модели направлены на устранение фрагментарности данных и

			выявление семантических связей между объектами водной среды. Полученные результаты, подтверждённые вычислительными экспериментами и сравнительным анализом с современными методами, представляют собой существенный научный вклад.
		4.2 Содержание диссертации отражает тему диссертации: 1) отражает; 2) частично отражает; 3) не отражает.	Содержание диссертации полностью отражает тему диссертации.
		4.3. Цель и задачи соответствуют теме диссертации: 1) соответствуют; 2) частично соответствуют; 3) не соответствуют.	В диссертационной работе четко сформированы цели и задачи исследования, которые полностью соответствуют теме диссертации.
		4.4 Все разделы и положения диссертации логически взаимосвязаны: 1) полностью взаимосвязаны; 2) взаимосвязь частичная; 3) взаимосвязь отсутствует.	Все разделы диссертационной работы логически полностью взаимосвязаны. Каждый раздел диссертации подробно описан. Полученные в ходе исследований результаты служат основой для выводов и положений, представленных на защиту.
		4.5 Предложенные автором новые решения (принципы, методы) аргументированы и оценены по сравнению с известными решениями: 1) критический анализ есть; 2) анализ частичный; 3) анализ представляет собой не собственные мнения, а цитаты других авторов.	Автором проведён критический анализ существующих методов (TableMiner++, QTLTableMiner++, MantisTable), выявлены их ограничения и предложены собственные решения, включая итеративный подход STI и гидро-онтологическую модель. Предложенные методы аргументированы и экспериментально подтверждают превосходство по сравнению с известными подходами, что свидетельствует о наличии самостоятельного научного вклада.

5.	Принцип научной новизны	5.1 Научные результаты и положения являются новыми? 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%).	Научные результаты диссертационной работы являются полностью новыми и заключаются в разработке улучшенного итеративного подхода Semantic Table Interpretation (STI) с расширенными эвристиками и механизмом автоматического пополнения графов знаний, а также в создании гидро-онтологической модели водных ресурсов Казахстана, выровненной с международными стандартами и дополненной авторскими модулями. Предложенные положения, выносимые на защиту, обеспечивают более точный семантический анализ табличных данных и автоматизированный вывод знаний, что подтверждает их оригинальность и научную значимость.
		5.2 Выводы диссертации являются новыми? 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%).	Выводы диссертационной работы являются полностью новыми, так как основаны на разработке оригинального итеративного подхода STI, онтологически-ориентированной геоинформационной модели и воспроизводимого конвейера интеграции гетерогенных данных. Полученные результаты и выводы подтверждены экспериментальной валидацией, сравнительным анализом с существующими методами и апробацией на реальных данных Иле-Балхашского бассейна.
		5.3 Технические, технологические, экономические или управленческие решения являются новыми и обоснованными: 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%).	Предложенные в диссертационной работе технические и технологические решения, включая разработку онтологически-ориентированной геоинформационной модели, итеративного метода STI и сквозного конвейера обработки данных, являются полностью новыми и обоснованными. Их эффективность подтверждена экспериментальной апробацией, сравнительным анализом с существующими подходами и применением на реальных данных водных ресурсов Казахстана.
6.	Обоснованность основных выводов	Все основные выводы основаны/не основаны на весомых с научной точки зрения доказательствах либо достаточно хорошо обоснованы (для qualitative research и направлений подготовки по искусству и гуманитарным наукам)	Все основные выводы диссертационной работы основаны на весомых с научной точки зрения доказательствах и являются достаточно хорошо обоснованными. Обоснованность подтверждается результатами вычислительных экспериментов, формированием эталонного набора данных, сравнительным анализом с современными методами (TableMiner++, QTLTableMiner++, MantisTable), а также практической

			апробацией разработанного подхода на реальных данных Иле-Балхашского бассейна. Высокие значения метрик качества (Triple F1, Precision, Recall), а также результаты применения правило-ориентированного вывода и построения масштабируемого графа знаний свидетельствуют о достоверности и научной обоснованности полученных выводов.
7.	Основные положения, выносимые на защиту	Необходимо ответить на следующие каждому положению в отдельности: 7.1 Доказано ли положение? 1) доказано; 2) скорее доказано; 3) скорее не доказано; 4) не доказано. 7.2 Является ли тривиальным? 1) да; 2) нет. 7.3 Является ли новым? 1) да; 2) нет. 7.4 Уровень для применения: 1) узкий; 2) средний; 3) широкий. 7.5 Доказано ли в статье? 1) да; 2) нет.	Положение 1. Итеративный подход Semantic Table Interpretation (STI). 7.1 1) доказано. Подтверждено результатами экспериментов и сравнительным анализом с существующими методами. 7.2 2) нет. Содержит новые эвристики и механизм автоматического пополнения графа знаний. 7.3 1) да. Предложен оригинальный подход STI с улучшенной точностью извлечения триплетов. 7.4 3) широкий. Может применяться для различных доменов и типов гетерогенных табличных данных. 7.5 1) да. Результаты метода опубликованы в рецензируемых научных изданиях. Положение 2. Гидро-онтологическая модель Казахстана с модулями и SWRL-выводом. 7.1 1) доказано. Подтверждено реализацией модели и её применением на реальных данных. 7.2 2) нет. Включает оригинальные модули и интеграцию нормативных правил. 7.3 1) да.

			Разработана новая модель, адаптированная к национальному контексту. 7.4 3) широкий. Может использоваться в системах мониторинга, анализа и управления водными ресурсами. 7.5 1) да. Основные результаты опубликованы и прошли научную апробацию.
8.	Принцип достоверности. Достоверность источников и предоставляемой информации.	8.1 Выбор методологии – обоснован методология достаточно подробно описана 1) да; 2) нет	Выбор методологии обоснован и подробно описан, включая методы онтологического моделирования, итеративной семантической интерпретации таблиц (STI), правило-ориентированного вывода (SWRL) и экспериментальной валидации на реальных данных, что обеспечивает достоверность полученных результатов.
		8.2 Результаты диссертационной работы получены с использованием современных методов научных исследований и методик обработки и интерпретации данных с применением компьютерных технологий: 1) да; 2) нет.	Результаты диссертационной работы получены с использованием современных методов онтологического моделирования, семантического анализа данных, итеративной интерпретации таблиц (STI), а также технологий обработки и интеграции данных (RDF, GraphDB, SPARQL, SWRL) с применением компьютерных средств.
		8.3 Теоретические выводы, модели, выявленные взаимосвязи и закономерности доказаны и подтверждены экспериментальным исследованием (для направлений подготовки по педагогическим наукам результаты доказаны на основе педагогического эксперимента): 1) да; 2) нет.	Теоретические положения и разработанные модели подтверждены результатами вычислительных экспериментов, сравнительным анализом с существующими методами и апробацией на реальных данных водных ресурсов Иле-Балхашского бассейна, что свидетельствует о достоверности выявленных взаимосвязей и закономерностей.

		8.4 Важные утверждения подтверждены/частично подтверждены/не подтверждены ссылками на актуальную и достоверную научную литературу	Важные утверждения диссертационной работы подтверждены ссылками на актуальную и достоверную научную литературу, включая фундаментальные и современные исследования в области Semantic Web, онтологического моделирования и семантической интерпретации табличных данных. Использованы как классические работы (Gruber, Berners-Lee), так и современные публикации и стандарты, что свидетельствует о корректности теоретической базы исследования.
		8.5 Использованные источники литературы достаточны/не достаточны для литературного обзора	Использованные источники литературы являются достаточными для проведения полноценного литературного обзора и охватывают широкий спектр направлений: онтологии, графы знаний, гидрологические модели, а также современные методы семантического анализа табличных данных. Представленный список включает как фундаментальные, так и новейшие научные работы, что обеспечивает полноту и глубину анализа исследуемой проблемы.
9	Принцип практической ценности	9.1 Диссертация имеет теоретическое значение: 1) да; 2) нет.	Диссертация имеет теоретическое значение, поскольку в работе формализован итеративный процесс STI, включающий выявление ролей столбцов, определение предикатных и атрибутных связей, а также онтологическое выравнивание и правило-ориентированный вывод знаний.
		9.2 Диссертация имеет практическое значение и существует высокая вероятность применения полученных результатов на практике: 1) да; 2) нет.	Диссертационная работа имеет практическую значимость, результаты исследования могут быть широко применены на практике, включая управление водными ресурсами, гидротехническими сооружениями и планирование водной инфраструктуры.
		9.3 Предложения для практики являются новыми? 1) полностью новые; 2) частично новые (новыми являются 25-75%); 3) не новые (новыми являются менее 25%).	Предложения для практики являются полностью новыми, поскольку основаны на разработке оригинального онтологически-ориентированного конвейера интеграции гетерогенных данных и автоматического формирования графа знаний, ранее не реализованного для водных ресурсов Казахстана.

10.	Качество написания и оформления	Качество академического письма: 1) высокое; 2) среднее; 3)ниже среднего; 4) низкое.	Диссертационная работа отличается высоким уровнем академического письма, логичной структурой изложения, корректным использованием научной терминологии и соответствием требованиям к оформлению научных исследований.
11.	Замечания к диссертации	Диссертационная работа выполнена на высоком научном уровне, представленные замечания носят рекомендательный характер и не снижают общей положительной оценки работы. 1. В диссертации отсутствует информация о доступности используемых данных и разработанных ресурсов, включая аннотированный корпус, исходный код метода STI и файлы онтологии (OWL/TTL), что затрудняет независимую проверку и расширение результатов другими исследователями. 2. Сравнительный анализ проведён на данных одного предметного домена (водные ресурсы Казахстана), что может создавать эффект «home advantage» и ограничивает оценку универсальности предложенного метода при его применении к другим типам табличных данных.	
12.	Научный уровень публикаций докторанта по теме исследования (в случае защиты диссертации в виде серии статей официальные рецензенты оценивают научный уровень каждой статьи докторанта по теме исследования)	Публикации докторанта Оспан Әсел Ғалымжанқызы по теме диссертационного исследования являются достаточными. Статья "Ontology-Driven Semantic Analysis of Tabular Data: An Iterative Approach with Advanced Entity Recognition" опубликована в журнале Applied Sciences (MDPI, Scopus, Q1), а также статья "The Development of a Water Resource Monitoring Ontology as a Research Tool for Sustainable Regional Development" — в журнале Data (MDPI, Scopus, Q2). Статьи, опубликованные в научных изданиях, рекомендованных Комитетом по обеспечению качества в сфере науки и высшего образования Министерства науки и высшего образования Республики Казахстан: 1. Nugumanova A., Apayev K., Baiburin Y., Mansurova M., Ospan A. QURMA: A Table Extraction Pipeline for Knowledge Base Population. Journal of Mathematics, Mechanics and Computer Science, 2022;114(2). 2. Ospan A., Mansurova M., Kakimzhanov E., Ixanov S., Barakhnin V. Development of a Program for the Integration of Socio-Economic Indicators with Spatial Data to Analyze the Standard of Living of the Population of Kazakhstan. Вестник Национальной инженерной академии Республики Казахстан, 2022;(85):67–78. 3. Оспан Ә., Мансурова М. Разработка гибридной модели для эффективного распределения водных ресурсов на основе модели прогнозирования. Физика-математика ғылымдары, 2021 №1(69). В результате исследования разработан итеративный алгоритм для семантического анализа таблиц из гетерогенных источников, который защищён авторским правом и внесён в государственный реестр Республики Казахстан (свидетельство № 39296 от 28 сентября 2023 года), что подтверждает практическую значимость работы. Представленные публикации раскрывают основные направления исследования, отражают научную новизну и в полной мере соответствуют содержанию диссертационной работы.	

13.	Решение официального рецензента (в соответствии с пунктом 28 настоящих Правил)	Диссертационная работа Оспан Әсел Галымжанқызы на тему «Разработка геоинформационных моделей с использованием онтологического подхода для извлечения знаний из гетерогенных источников», представленная на соискание степени доктора философии (PhD) по специальности «8D06105 – Системная инженерия», соответствует требованиям Комитета по обеспечению качества в сфере науки и высшего образования Министерства науки и высшего образования Республики Казахстан и заслуживает присуждения степени доктора философии (PhD) по указанной специальности.

Официальный рецензент:
Евразийский национальный университет им. Л.Н. Гумилева
PhD, ассоциированный профессор,
доцент кафедры «Информационные системы» **М.А. Самбетбаева**