НИКИМТ-АТОМСТРОЙ
РОСАТОМ

Акционерное общество
«Научно-исследовательский и конструкторский институт монтажной технологии – Атомстрой»
(АО «НИКИМТ-Атомстрой»)
Алтуфьевское шоссе, д.43, стр.2, Москва, 127410, Россия
Тел.: +7 (495) 411 65 50, Факс: +7 (495) 411 65 52
E-mail: post@atomrus.ru, nikimtatomstroy.ru
ОКПО 08621486, ОГРН 5087746235836
ИНН 7715719854, КПП 771501001

«УТВЕРЖДАЮ»
И.о. генерального директора
Панкратов Д.С.

«01» июня 2026 г.

ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертацию Дагаева Александра Евгеньевича
«РАЗРАБОТКА АЛГОРИТМОВ ИЗВЛЕЧЕНИЯ И ОБРАБОТКИ ИНФОРМАЦИИ ДЛЯ ГЕНЕРАЦИИ МНОГОФОРМАТНЫХ ТЕСТОВЫХ ЗАДАНИЙ»,
представленную на соискание учёной степени кандидата технических наук по научной специальности **2.3.8. – Информатика и информационные процессы**

Актуальность темы исследования обусловлена необходимостью разработки новых инструментальных средств, обеспечивающих объективную, масштабируемую и воспроизводимую оценку степени сформированности отдельных компонентов компетенций, включая когнитивные, аналитические и прикладные аспекты для полноценной реализации компетентностной модели образования. При этом, одним из ключевых средств такой оценки являются тестовые задания различных форматов, используемые для диагностики усвоения учебного материала и компонентов компетенций. Традиционные подходы к их разработке, основанные преимущественно на ручной подготовке, характеризуются высокой трудоёмкостью, субъективностью и ограниченной масштабируемостью, что существенно затрудняет их применение в современных образовательных системах, ориентированных как на массовое, так и персонализированное обучение. Разрабатываемые современные алгоритмы рассматриваются не как замена педагогической деятельности, а как вспомогательный инструментарий цифровой образовательной среды, предназначенный для снижения трудоёмкости рутинных операций преподавателя.

В последние годы в сфере образования всё большее внимание уделяется применению методов искусственного интеллекта и обработки естественного языка для автоматизированной генерации тестовых заданий на основе учебных текстов. Однако, несмотря на достигнутые значимые результаты, существующие решения обычно ориентированы на генерацию отдельных типов заданий или на формальное соответствие исходному тексту без учёта дидактических

требований, качества дистракторов и необходимости верификации корректности сформированных заданий, отсутствует интеграция этапов извлечения информации, генерации заданий и последующей оценки их качества.

Таким образом, актуальной научно-технической задачей является разработка алгоритмов извлечения и обработки информации, обеспечивающих автоматизированную генерацию многоформатных тестовых заданий и дистракторов с последующей интеллектуальной верификацией их качества. Решение этой задачи связано с интеграцией методов обработки естественного языка, алгоритмов генерации контента и формализованных процедур оценки характеристик формируемых заданий, ориентированных на требования компетентностного подхода и практику применения в современных образовательных системах.

В этой связи, тема диссертационного исследования Дагаева А.Е. является своевременной и актуальной, представляет как теоретический, так и практический интерес.

Степень обоснованности и достоверности научных положений, выводов и рекомендаций, сформулированных в диссертации. Работа посвящена исследованию и решению важной научно-технической задачи, связанной с разработкой набора алгоритмов извлечения и обработки информации для автоматизированной генерации многоформатных тестовых заданий и дистракторов с последующей верификацией их качества на основе применения технологий и методов искусственного интеллекта.

Целью работы является разработка и исследование интеллектуальных алгоритмов для извлечения и обработки информации, генерации и верификации многоформатных тестовых материалов из цифрового контента.

Цель и поставленные научные задачи определяют структуру диссертационной работы и логику изложения результатов исследования.

Достоверность положений, выводов и рекомендаций, сформулированных в диссертации Дагаева А.Е., обусловлена глубокой проработкой большого объема научно-практической, аналитической и статистической информации по тематике исследования. За время подготовки диссертационной работы соискателем выполнен анализ широкого спектра работ отечественных и зарубежных авторов, посвященных автоматизированной генерации тестовых заданий, алгоритмический подход, объединяющий извлечение и обработку информации, генерацию и интеллектуальную верификацию многоформатных тестовых материалов из цифрового контента, с учётом специфики русскоязычных образовательных текстов (список литературы включает 136 источников).

Обоснованность научных положений и выводов диссертационного исследования основана на использовании современных теоретических и практических методов исследования, верификации результатов за счет сопоставления с данными исследований других авторов и

материалами действующих нормативных документов, а также общенаучных методов: анализа, синтеза, обобщения, группировки, классификации, моделирования, прогнозирования и др. В работе используется сочетание методов компьютерной лингвистики, семантического извлечения информации и методов машинного обучения для разработки и валидации алгоритмов автоматической генерации тестовых заданий. При этом конкретные подходы включают глубокий семантический, синтаксический и морфологический анализ, а также применение больших языковых моделей для извлечения релевантных данных из образовательных текстов. Экспериментальная методология включает как качественные, так и количественные методы оценки, а также метрики качества сгенерированных тестовых заданий и сравнительный анализ с существующими алгоритмами.

Изложенный материал проиллюстрирован формулами и таблицами. В тексте диссертации имеются ссылки на источники информации, используемые соискателем при подготовке диссертации.

Апробация основных научных положений и практических результатов, полученных по итогам диссертационного исследования, проведена путём выполнения докладов и всестороннего обсуждения результатов на международных научных и научно-технических конференциях.

Результаты работы в виде разработанных алгоритмов в системах электронного обучения, платформах массовых открытых онлайн-курсов и корпоративных системах обучения внедрены в практическую деятельность профильных организаций и компаний, о чем свидетельствуют результаты и акты внедрения. Соискателем получены свидетельства о государственной регистрации программы для ЭВМ, выданные Федеральной службой по интеллектуальной собственности (Роспатент).

Оценка новизны научных результатов

В диссертационной работе разработан и на практике реализован набор эффективных алгоритмов, обеспечивающих извлечение и обработку информации, автоматизированную генерацию многоформатных тестовых заданий и дистракторов с последующей интеллектуальной верификацией их качества.

Научная новизна диссертационной работы заключается в разработке, исследовании свойств и имплементации набора алгоритмов, направленных на извлечение и обработку информации для генерации многоформатных тестовых заданий.

Среди существенных научных результатов диссертационного исследования можно выделить:

1. Впервые предложен алгоритм абстрактивного извлечения информации, отличающийся от существующих алгоритмов, основанных на статистических или общих семантических оценках качества тем, что вводит параметрическую интегральную метрику с доменно-

специфичными весовыми коэффициентами, что позволяет повысить релевантность извлекаемой информации.

2. Впервые разработан алгоритм автоматической генерации многоформатных тестовых заданий, отличающийся от существующих подходов на основе больших языковых моделей тем, что интегрирует модуль мультиагентной верификации, позволяющий существенно снизить долю дефектных заданий.

3. Впервые предложен алгоритм генерации дистракторов, отличающийся от алгоритмов, основанных на семантической близости к правильному ответу тем, что моделирует процесс построения корректной цепочки рассуждений с последующим внесением когнитивных ошибок для генерации неверных ответов, что позволяет повысить качество дистракторов относительно базовых алгоритмов.

Достоверность полученных результатов и обоснованность выводов подтверждается корректным использованием математических моделей при общепринятых допущениях, большим объёмом имитационного моделирования, проверкой основных теоретических положений в выполненных вычислительных экспериментах, связанных с проверкой адаптивности разработанных алгоритмов к различным образовательным контекстам.

Значимость полученных результатов для науки и практики. Теоретическая значимость исследования заключается в развитии соискателем научных направлений, связанных с обработкой естественного языка, компьютерной лингвистикой и педагогическими измерениями. При этом важной теоретической составляющей являются семантическая разметка ролей, теория когнитивной нагрузки в образовательной оценке и принципы психометрического конструирования тестов. Разработанные алгоритмы генерации дистракторов опираются на принципы когнитивного моделирования ошибок, а система верификации – на теорию многоагентных систем. Кроме того, выполненные исследования учитывают современные тренды и достижения в области информатики и информационных процессов, образовательных технологий, основанных на моделях искусственного интеллекта и машинном обучении.

Практическая значимость научных результатов заключается в разработанных прикладных программно-алгоритмических решениях для извлечения и обработки информации, обеспечивающих автоматизированную генерацию многоформатных тестовых заданий и дистракторов с последующей интеллектуальной верификацией их качества.

1. Разработан алгоритм семантического анализа образовательного контента и извлечения из него ключевой информации, релевантной для последующей генерации заданий.

2. Разработан алгоритм генерации многоформатных тестовых заданий, успешно формирующий задания с выбором одного или нескольких вариантов ответа, задания открытой формы, а также на установление соответствия и правильной последовательности.

3. Разработан алгоритм интеллектуального формирования дистракторов на основе когнитивного моделирования ошибок обучающихся.

4. Разработан программный комплекс для автоматизированной генерации контрольно-измерительных материалов с высокой эффективностью предложенных решений и дидактическим качеством генерируемых материалов, готовых к интеграции в современные информационно-образовательные системы.

Личный вклад соискателя состоит в проведении анализа существующих математических и программно-алгоритмических методов, используемых при обработке информации и информационных процессов для автоматизированной генерации многоформатных тестовых заданий и дистракторов с последующей интеллектуальной верификацией их качества на основе методов искусственного интеллекта; разработке алгоритмов автоматизированной генерации тестовых заданий для русскоязычных образовательных текстов; автоматизированном формировании тестовых заданий различных форматов и автоматической верификации их корректности и педагогической пригодности; разработке методики использования доменно-настроенных интегральных оценок, многоагентной верификации и генерации дистракторов на основе моделирования рассуждений для повышения диагностической ценности автоматически формируемых контрольно-измерительных материалов; экспериментальной апробации разработанных алгоритмов; подготовке рукописных и графических материалов для публикации результатов исследований и выступлении с докладами на международных научно-технических конференциях.

Рекомендации по использованию результатов и выводов диссертации.

Результаты диссертационного исследования Дагаева А.Е. внедрены и используются в Московском Политехе (г. Москва) и ООО «ОПТИК СМАРТ» (г. Тула). Кроме того, научно-практические результаты могут быть рекомендованы для применения в инженерно-педагогической практике при проектировании и эксплуатации компьютерных систем дистанционного обучения. Следует отметить, что решение исследуемой научной задачи особенно актуально для высших учебных заведений и учреждений дополнительного образования, включая организации машиностроительной и энергетической отрасли промышленности и, в частности – предприятия атомной энергетики.

Материалы и выводы диссертационной работы рекомендуются к использованию в учебном процессе высших учебных заведений Российской Федерации по направлениям подготовки бакалавриата: 09.03.01 – Информатика и вычислительная техника; 09.03.02 – Информационные системы и технологии; 09.03.04 – Программная инженерия; 15.03.04 – Автоматизация технологических процессов и производств; 44.03.01 – Педагогическое образование; 45.03.03 – Фундаментальная и прикладная лингвистика; 45.03.04 -

Интеллектуальные системы в гуманитарной сфере, а также – по направлениям подготовки магистратуры: 09.04.01 – Информатика и вычислительная техника; 09.04.02 – Информационные системы и технологии; 09.04.04 – Программная инженерия; 15.04.04 – Автоматизация технологических процессов и производств; 44.04.01 – Педагогическое образование; 45.04.03 – Фундаментальная и прикладная лингвистика; 45.04.04 – Интеллектуальные системы в гуманитарной сфере.

Замечания и рекомендации по диссертации и автореферату. Несмотря на выполненные объёмные исследования и вклад соискателя в разработку моделей и программно-алгоритмических решений для извлечения и обработки информации, обеспечивающих автоматизированную генерацию многоформатных тестовых заданий и дистракторов, по представленной работе необходимо высказать следующие замечания и рекомендации:

1. В тексте диссертационной работы описывается применение предобученных языковых моделей, например, на стр. 52. Однако в самой работе и, в частности, в приведенном случае, не прослеживается упоминание конкретных использованных языковых моделей с параметрами их размера и квантизации, условий их дообучения, методов обучения классификатора, параметров обучения, достигнутых при обучении метрик и диаграмм, иллюстрирующих процесс обучения.

2. На стр. 90, 91 диссертационной работы приводится объяснение полученных и приведенных в таблице 3.7 результатов: «…Высокие показатели снижения для предвзятости и копирования также связаны с относительной определённостью соответствующих критериев. Более низкие значения показателей тематической нерелевантности и несоответствия уровня сложности обусловлены тем, что границы между релевантным и нерелевантным содержанием, а также между соседними когнитивными уровнями, не поддаются столь же строгой формализации, как задачи фактической проверки…» Соискатель объясняет полученные результаты тем, что в отличии от агентов №2 для оценки «тематической релевантности» и №5 для оценки «соответствия уровню сложности», при определении «наличия предвзятости» с помощью агента №6 алгоритм использует более формализованные и однозначные критерии. Это входит в противоречие с описанием перечисленных агентов на стр. 77, 78, 80, 81, когда согласно тексту диссертации, «инструменты» для агентов №2 и №5 определены более точно и формализованы лучше, чем для агента №6, который содержит нестрого определенную LLM-модель и абстрактный «набор потенциально релевантных выражений». Среди «измерений» агента «наличия предвзятости» приводится «гендерный и культурный нейтралитет» и «социально-экономические допущения», но, при этом, не рассматривается вопрос утраты вовлеченности обучаемых и снижение прикладной практической ценности заданий из-за деградации контекста задачи в сторону чрезмерной абстрактности.

3. В диссертационной работе приведены алгоритмы с этапами, предполагающими применение больших языковых моделей (LLM). Упоминаются «LLM-декомпозиция», «LLM-рассуждение», «LLM-лингвистический анализ», «LLM-аудит», «LLM-декомпозиция». На стр. 87, 88, 93, 94 проводится анализ числа LLM-запросов. Приводятся библиографические ссылки на научные работы, в тексте которых упоминаются конкретные языковые модели и их версии. При этом из текста диссертации в явном виде не следует какие именно предобученные языковые модели используются в программных приложениях алгоритмов и какие обязательные требования предъявляются к таким языковым моделям.

4. В личный вклад соискателя (стр. 9 автореферата диссертации) включена программная реализация предложенных им алгоритмических решений. Приведены акты внедрения ряда зарегистрированных в Роспатенте программ для ЭВМ. На стр. 112 диссертации указано, что «…все программные компоненты комплекса реализованы на языке Python, что обеспечило единообразие средств разработки, повторное использование библиотек обработки естественного языка и машинного обучения, а также унификацию межмодульного взаимодействия…». При этом нигде в самой диссертационной работе или в приложениях к ней не приведены листинги программ реализации разработанных алгоритмов. Кроме того, не приведены диаграммы бизнес-процессов, например, в нотации BPMN 2.0, иллюстрирующие изменение бизнес-архитектуры прикладных программных решений с выделением вновь разработанных и внедренных в неё элементов, не указаны использованные соискателем бинарные библиотеки для обработки естественного языка и машинного обучения. Опущена важная с точки зрения программной архитектуры, практической реализации и цифровой безопасности информация об использовании для запуска языковых моделей локальных или удаленных специализированных фреймворков.

5. Для алгоритмов интеллектуального анализа текстовых данных и автоматической генерации многоформатных тестовых заданий, являющихся предметом исследования, следовало бы и в диссертации, и в автореферате привести стандартное описание в виде схем согласно ГОСТ 19.701-90/ ИСО 5807-85 «Единая система программной документации. Схемы алгоритмов, программ, данных и систем. Условные обозначения и правила выполнения».

Высказанные замечания по большей части носят некритичный и рекомендательный характер, не снижают общей положительной оценки уровня научной и практической значимости проделанной работы.

Соответствие диссертации требованиям, установленным Положением о присуждении ученых степеней.

Диссертационная работа Дагаева А.Е. на тему «Разработка алгоритмов извлечения и обработки информации для генерации многоформатных тестовых заданий» является завершенным научно-квалификационным исследованием, выполненным соискателем

самостоятельно на хорошем научном уровне, в которой содержится решение актуальной научной задачи разработки и исследования интеллектуальных алгоритмов для извлечения и обработки информации, генерации и верификации многоформатных тестовых материалов из цифрового контента, вносящей вклад в развитие информатики и информационных процессов.

Диссертационная работа и автореферат написаны ясным, грамотным языком, материалы изложены последовательно, оформлены в соответствии требованиям нормативных документов, выводы по проведённому исследованию представлены структурно и содержательно. Автореферат соответствует диссертационной работе и отражает её содержание.

Автореферат и опубликованные соискателем работы в полном объеме отражают основные положения и результаты диссертации, которые представлены в 20 публикациях, в том числе, в 4 работах в изданиях из Перечня рецензируемых научных изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание учёной степени кандидата наук, на соискание учёной степени доктора наук.

Содержание работы, её научная новизна, обоснованность и достоверность полученных результатов позволяет утверждать, что диссертация соответствует требованиям п.п. 9-11 Положения о присуждении ученых степеней, утверждённого постановлением Правительства РФ от 24.09.2013 №842 (с изменениями и дополнениями), предъявляемым к кандидатским диссертациям и соответствует научной специальности **2.3.8. – Информатика и информационные процессы**, в том числе, в части пунктов паспорта вышеуказанной специальности:

«4. Разработка методов и технологий цифровой обработки аудиовизуальной информации с целью обнаружения закономерностей в данных, включая обработку текстовых и иных изображений, видео контента. Разработка методов и моделей распознавания, понимания и синтеза речи, принципов и методов извлечения требуемой информации из текстов.».

«5. Лингвистическое обеспечение информационных систем и процессов. Методы и средства проектирования словарей данных, словарей индексирования и поиска информации, тезаурусов и иных лексических комплексов. Методы семантического, синтаксического и прагматического анализа текстовой информации для представления в базах данных и организации интерфейсов информационных систем с пользователями.».

Таким образом, согласно результатам всестороннего изучения и анализа материалов автореферата и диссертации на тему **«Разработка алгоритмов извлечения и обработки информации для генерации многоформатных тестовых заданий»**, выполненной по научной специальности **2.3.8. – Информатика и информационные процессы**, можно сделать вывод о том, что соискатель Дагаев Александр Евгеньевич заслуживает присуждения ему учёной степени

кандидата технических наук по вышеуказанной научной специальности на основании результатов публичной защиты.

Заключение в результате анализа материалов диссертационной работы и автореферата диссертации Дагаева А.Е., всестороннего обсуждения полученных соискателем научно-практических результатов принято на расширенном заседании Научно-инженерного центра цифровых компьютерных систем и автоматизации Научно-исследовательского и конструкторского института монтажной технологии (НИКИМТ).

Присутствовало на заседании 10 чел. Результаты голосования: «за» – 10 чел., «против» – 0 чел., «воздержалось» – 0 чел., протокол №3 от «22» мая 2026 г.

Председатель:
доктор технических наук, профессор,
начальник Научно-инженерного центра
цифровых компьютерных систем и
автоматизации НИКИМТ

Евгений Евгеньевич Ковшов

Секретарь:
кандидат технических наук,
старший научный сотрудник Научно-
инженерного центра цифровых
компьютерных систем и автоматизации
НИКИМТ

Владимир Сергеевич Кувшинников

Подписи Е.Е. Ковшова и В.С. Кувшинникова удостоверяю.

Главный эксперт /Ю.Л. Лаврова/

Служба управления персоналом

С отзывом ознакомлен 10.06.2026