

«УТВЕРЖДАЮ»



Проректор по науке и инновациям
ФГБОУ ВО «НИУ «МЭИ»
доктор технических наук, доцент
Комаров И.И.

« 30 » сентября 20 26 г.

ЗАКЛЮЧЕНИЕ

федерального государственного бюджетного образовательного учреждения
высшего образования

«Национальный исследовательский университет «МЭИ»

Диссертация «Методы и гетерогенная гибридная платформа многоуровневого ансамблирования объяснений при комплексировании моделей машинного обучения» выполнена соискателем Трофимовым Юрием Владиславовичем в федеральном государственном бюджетном образовательном учреждении высшего образования «Национальный исследовательский университет «МЭИ» на кафедре прикладной математики и искусственного интеллекта.

В период подготовки диссертации соискатель Трофимов Юрий Владиславович был прикреплен для подготовки диссертации на соискание ученой степени кандидата наук без освоения программы подготовки научных и научно-педагогических кадров в аспирантуре в федеральном государственном бюджетном образовательном учреждении высшего образования «Национальный исследовательский университет «МЭИ».

Соискатель в 2025 году окончил магистратуру федерального государственного бюджетного учреждения высшего образования Московской области «Университет «Дубна» по направлению/специальности 27.04.03 «Системный анализ и управление».

Сведения о сдаче кандидатских экзаменов приведены в справках:

1. Справка № 749 от 28.10.2025 г., выданная федеральным государственным бюджетным образовательным учреждением высшего образования Московской области «Государственный университет «Дубна».
2. Справка № 58/17 от 16.05.2026 г. выданная федеральным государственным бюджетным образовательным учреждением высшего образования «Национальный исследовательский университет «МЭИ» (на основании приказа № 515/А от 23.12.2025 г. был прикреплен для сдачи кандидатских экзаменов по специальной дисциплине 1.2.1 «Искусственный интеллект и машинное обучение (технические науки)»).

Научный руководитель – Аверкин Алексей Николаевич, кандидат физико-математических наук, доцент кафедры прикладной математики и искусственного интеллекта ФГБОУ ВО «НИУ «МЭИ», ведущий научный сотрудник отдела № 14 «Интеллектуальных систем» федерального исследовательского центра «Информатика и управление» Российской академии наук.

Доклад по теме диссертации был заслушан на заседании кафедры прикладной математики и искусственного интеллекта ФГБОУ ВО «НИУ «МЭИ».

Присутствовали:

д.т.н., проф. Борисов В.В., к.т.н., доц. Бартеньев О.В., д.т.н., проф. Гданский Н.И., д.т.н., проф. Еремеев А.П., д.ф.-м.н., проф. Ижуткин В.С., д.т.н., проф. Кутепов В.П., д.т.н., проф. Федулов А.С., к.ф.-м.н., доц. Алексиадис Н.Ф., к.т.н., доц. Варшавский П.Р., к.т.н.,

доц. Вежелис Т.М., к.т.н., доц. Воробьева И.А., к.ф.-м.н. ИONOва Т.В., к.т.н. Кружилов И.С., к.т.н., доц. Куриленко И.Е., к.т.н., доц. Маран М.М., к.т.н., доц. Меньшикова К.Г., к.т.н. Михайлов И.С., к.т.н. Ройзензон Г.В., к.т.н. Сергеев А.В., к.т.н., доц. Фомина М.В., к.т.н., доц. Хорев П.Б., к.т.н., доц. Чернецов А.М., к.т.н., доц. Чернов П.Л., к.т.н., доц. Шамаева О.Ю., к.т.н. Щирый А.О., к.пед.н., Паристова Е.А., к.т.н., Панявин Н.А., аспирант Задорин С.А., аспирант Сесин А.А.

Всего присутствовало на заседании 28 человек.

В ходе обсуждения работы соискателю были заданы вопросы:

Д.т.н., проф. Еремеев А.П.: Каким пунктам паспорта специальности 1.2.1 соответствуют полученные результаты и почему в формулировках научной новизны не упоминаются методы искусственного интеллекта и машинного обучения?

Д.т.н., проф. Еремеев А.П.: Что в работе понимается под гетерогенными объяснениями?

К.т.н., доц. Бартеньев О.В.: Как совместно обрабатываются разнородные объяснения, например, текстовые и графические, если в приведённых примерах объяснения представлены изображениями?

Д.т.н., проф. Еремеев А.П.: Есть ли акты от медицинских организаций и что в них сказано о пользе системы для врачей?

Д.т.н., проф. Еремеев А.П.: Как обеспечивается согласование объяснения с предметной областью, если признаки, важные для модели, не совпадают с признаками, важными для врача?

Д.т.н., проф. Еремеев А.П.: Проводилось ли сравнение с подходом на основе библиотеки прецедентов, накопленной медиками?

Д.т.н., проф. Федулов А.С.: Можно ли количественно оценить, насколько система помогает лицу, принимающему решение, например снижает риски или повышает достоверность диагноза?

К.т.н., доц. Ройзензон Г.В.: Почему для интегральной оценки объяснений выбрана аддитивная свёртка и откуда взяты веса критериев?

К.т.н., доц. Бартеньев О.В.: Как учитывается накопление ошибок отдельных объяснителей при их объединении и как выбираются пороги отсечения?

К.т.н., доц. Бартеньев О.В.: По каким метрикам оценивается качество объяснений?

К.т.н., доц. Бартеньев О.В.: Как агрегируются объяснения в задаче выявления пневмонии, где применяются четыре объяснителя, и чем используемый оператор отличается от обычного усреднения?

Д.т.н., проф. Еремеев А.П.: В чём научная новизна по сравнению с существующими мультимодальными подходами и где показаны количественные преимущества предложенных решений?

К.т.н., доц. Варшавский П.Р.: Какова роль пользователя в автоматизированном контуре и на каких этапах без его участия не обойтись?

На заданные вопросы соискатель дал обстоятельные и исчерпывающие ответы.

Свои замечания по итогам рецензирования диссертации высказали д.т.н., профессор Федулов А.С., д.т.н., профессор Борисов В.В. и к.т.н., доц. Ройзензон Г.В. Заключительное слово сказал научный руководитель к.ф.-м.н., доцент Аверкин А.Н.

Рецензент от кафедры Ройзензон Г.В. зачитала проект заключения кафедры, составленный на основе обсуждения работы и выступлений участников заседания.

По итогам обсуждения диссертации принято следующее заключение.

Актуальность темы исследования

Актуальность темы определяется возрастающим применением моделей искусственного интеллекта в задачах, где цена ошибки высока и наряду с точностью необходимы прослеживаемость хода получения решения, оценка устойчивости и меры достоверности. Объяснимость в таких условиях становится необходимым условием надежного, воспроизводимого и практически значимого применения искусственного интеллекта. Государственная значимость развития искусственного интеллекта закреплена в Указе Президента Российской Федерации от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» и Указе Президента Российской Федерации от 15.02.2024 № 124. Особую значимость проблема приобретает в здравоохранении, что отражено в распоряжении Правительства Российской Федерации от 17.04.2024 № 959-р, распоряжении Правительства Российской Федерации от 02.12.2021 № 3427-р и постановлении Правительства Российской Федерации от 18.07.2023 № 1164. Практическая востребованность результатов подтверждается их использованием в специализированном программном обеспечении системы поддержки принятия врачебных решений ООО «МАКАО ИТ», в ГБУЗ Московской области «Дубненская больница», Лаборатории информационных технологий им. М.Г. Мещерякова ОИЯИ, Научно-исследовательском центре искусственного интеллекта и учебном процессе Университета «Дубна», а также при выполнении НИР по государственному заданию Минобрнауки РФ № 124112200072-2 и грантам РНФ № 22-71-10112 и № 22-17-00045.

Личный вклад автора заключается в:

Основные научные результаты, представленные в диссертации, получены автором лично под научным руководством к.ф.-м.н., доцента Аверкина А.Н. Постановка научной задачи, формирование концептуальных подходов и определение направлений исследования осуществлялись совместно с научным руководителем. Автором лично разработаны методы, алгоритмы и модели, выполнено проектирование архитектуры и реализация программного обеспечения, проведены вычислительные эксперименты, анализ и интерпретация полученных результатов, а также подготовка научных публикаций по теме исследования.

Научная новизна полученных автором результатов заключается в следующем:

- метод многоуровневого формирования объяснений в иерархически организованных системах ИИ, отличающийся представлением объяснимого контура в виде ориентированного графа узлов и переходов и формированием объяснительных объектов на его узлах и рёбрах с последующим выделением признакового, модельного, интерпретируемого и результирующего уровней;
- метод согласования уровней объяснения, отличающийся использованием операторов межуровневого переноса, мер расхождения и показателей согласованности для установления формально проверяемой связи между признаковым, модельным, интерпретируемым и результирующим уровнями;
- формальный аппарат агрегирования, объединения и комплексирования гетерогенных объяснений, отличающийся разделением способов совместной обработки однородных и разнородных объяснительных объектов и сохранением их типа и происхождения в итоговом объяснении;
- метод квантификации неопределённости объединённого объяснения, учитывающий расхождения между различными методами объяснения, повторными запусками и моделями, а для мультимодальных систем также согласованность объяснений различных модальностей;
- протокол воспроизводимого выбора метода объяснения и метрический контур его оценки, отличающиеся совместным учётом качества, устойчивости,

неопределённости, вычислительной стоимости и контекста применения в фиксированной процедуре сопоставления кандидатов.

Практическая значимость работы состоит в:

- разработке архитектуры и программных компонентов платформы, представляющей собой программно-методический комплекс для обучения и применения моделей, автоматизированного подбора моделей и методов объяснения, построения конвейеров обработки данных и анализа;

- разработке решений, обеспечивающих формирование и оценку объяснений, учет неопределенности и повышение воспроизводимости вычислительных экспериментов; по результатам исследования разработаны три программы для ЭВМ, на которые получены свидетельства о государственной регистрации;

- внедрении разработанных методов, алгоритмов и программных средств в специализированное программное обеспечение системы поддержки принятия врачебных решений ООО «МАКАО ИТ», а также их использовании в ГБУЗ Московской области «Дубненская больница», Лаборатории информационных технологий им. М.Г. Мещерякова ОИЯИ (объединенный институт ядерных исследований), Научно-исследовательском центре искусственного интеллекта и учебном процессе Университета «Дубна»; результаты также использованы при выполнении НИР по государственному заданию Минобрнауки РФ № 124112200072-2 и грантам РФФИ № 22-71-10112 и № 22-17-00045.

Достоверность полученных результатов подтверждается корректным использованием методов, обоснованием постановки задач, а также экспериментальными исследованиями на открытых данных и данных промышленных партнеров, демонстрирующими эффективность предложенных методов и алгоритмов и корректность работы разработанных программных решений.

Специальность, которой соответствует диссертация

Тема диссертации «Методы и гетерогенная гибридная платформа многоуровневого ансамблирования объяснений при комплексировании моделей машинного обучения» и ее содержание соответствуют направлениям исследований по п. 2, 4, 7 и 12 паспорта научной специальности 1.2.1 «Искусственный интеллект и машинное обучение»:

- п. 2. Исследования в области оценки качества и эффективности алгоритмических и программных решений для систем искусственного интеллекта и машинного обучения. Методики сравнения и выбора алгоритмических и программных решений при многих критериях.

- п. 4. Разработка методов, алгоритмов и создание систем искусственного интеллекта и машинного обучения для обработки и анализа текстов на естественном языке, для изображений, речи, биомедицины и других специальных видов данных.

- п. 7. Разработка специализированного математического, алгоритмического и программного обеспечения систем искусственного интеллекта и машинного обучения. Методы и средства взаимодействия систем искусственного интеллекта с другими системами и человеком-оператором.

- п. 12. Исследования в области «доверенных» систем класса ИИ, включая проблемы формирования тестовых выборок прецедентов, надежности, устойчивости, переобучения и т.д.

Полнота изложения материалов диссертации:

Основные научные результаты диссертационной работы опубликованы в рецензируемых научных изданиях.

Всего по теме диссертации опубликовано 43 печатные работы, из них 2 статьи в изданиях, входящих в перечень, установленный Министерством науки и высшего образования Российской Федерации и 3 свидетельства о государственной регистрации программы для электронных вычислительных машин.

В изданиях, входящих в международную базу цитирования Scopus:

1. Trofimov Y., Averkin A. Enhancing Results with ANFIS and XAI for Soil Contamination Monitoring and Public Health Risk Assessment // 2024 IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (PIERE). Novosibirsk, Russian Federation, 2024. P. 1120–1123. DOI: 10.1109/PIERE62470.2024.10805065. (Scopus)

Лично соискателем построены и исследованы гибридные нечёткие модели.

2. Trofimov Y.V., Shevchenko A.V., Averkin A.N., Muravyov I.P., Kuznetsov E.M. Concept of Hierarchically Organized Explainable Intelligent Systems: Synthesis of Deep Neural Networks, Fuzzy Logic and Incremental Learning in Medical Diagnostics // 2025 VI International Conference on Neural Networks and Neurotechnologies (NeuroNT). Saint Petersburg, Russian Federation, 2025. P. 14–17. DOI: 10.1109/NeuroNT66873.2025.11049976. (Scopus)

Автором лично разработана концептуальная схема иерархически организованной объяснимой интеллектуальной системы, объединяющей глубокие нейронные сети, нечёткую логику и инкрементное обучение в медицинской диагностике.

3. Verified Explainability Core: A GD–ANFIS/SHAP Hybrid Architecture for XAI 2.0 / Y. Trofimov, A. Lebedev, A. Ilin, A. Averkin // Automatic Documentation and Mathematical Linguistics. – 2025. – Vol. 59, Suppl. 5. – P. S469–S478. – DOI 10.3103/S0005105525701420. (Scopus)

Трофимовым Ю. В. лично сформировано гибридное ядро верифицируемой объяснимости GD–ANFIS/SHAP и исследовано согласование нейронечётких выводов с признаковыми атрибуциями.

4. Averkin A.N., Trofimov Y.V., Lebedev A.D., Ilin A.S. Hybrid and Hierarchical Explainable AI Based on Kolmogorov-Arnold Networks // Proceedings of the Ninth International Scientific Conference “Intelligent Information Technologies for Industry” (IITI’25), Volume 2 / edited by S. Kovalev, I. Kotenko, A. Sukhanov. – Cham : Springer, 2026. – (Lecture Notes in Networks and Systems, vol. 1763). – DOI: 10.1007/978-3-032-13612-1_4. (Scopus)

Вклад соискателя состоит в личной разработке и исследовании принципов гибридной и иерархической объяснимости на основе сетей Колмогорова–Арнольда.

5. Trofimov Y., Shevchenko A., Trusov I. Causality-Guided Admission Control for Diffusion-Generated Synthetic Medical Image // Proceedings of the 2026 International Russian Smart Industry Conference (SmartIndustryCon). Sochi, Russian Federation, 2026. P. 1093–1098. DOI: 10.1109/SmartIndustryCon68821.2026.11492854. (Scopus)

Соискателем лично предложен и исследован подход к каузально управляемому отбору синтетических медицинских изображений, генерируемых диффузионными моделями.

6. Trofimov Y., Averkin A., Lebedev M., Lebedev A., Ilyin A., Alekseev A. RuFLEX: A Hybrid Deep Neuro-Fuzzy Intelligence Platform for Geospatial Data with Logical Rule Traceability and Counterfactual Explainability // Proceedings of the 2026 VII International Conference on Neural Networks and Neurotechnologies (NeuroNT). Saint Petersburg, Russian Federation, 2026. P. 11–14. DOI: 10.1109/NeuroNT71829.2026.11651423. (Scopus)

Лично автором разработаны и апробированы компоненты гибридной нейронечёткой геопространственной платформы RuFLEX, обеспечивающие трассируемость логических правил и контрфактуальную объяснимость.

7. Kuznetsov E., Averkin A., Trofimov Y., Lopatin M., Kondrashova E. Physio-JEPA for ICU Signals: Temporal Multimodal Explainability and the Temporal Range Coherence Metric for Ventricular Tachycardia Alarm Validation // Proceedings of the 2026 VII International Conference on Neural Networks and Neurotechnologies (NeuroNT). Saint Petersburg, Russian Federation, 2026. P. 28–31. DOI: 10.1109/NeuroNT71829.2026.11650778. (Scopus)

Трофимов Ю. В. лично провёл исследование временной мультимодальной объяснимости сигналов ОПИТ и получил результаты применения метрики временной когерентности диапазонов при проверке тревог желудочковой тахикардии.

8. Kim T., Averkin A., Alexeev A., Trusov I., Trofimov Y. A Method for Generating Synthetic Cellular Data Using Rules, Variational Modeling, and Explainable Artificial Intelligence // 2026 IEEE 27th International Conference of Young Professionals in Electron Devices and Materials (EDM). Novosibirsk, Russian Federation, 2026. P. 1–4. DOI: 10.1109/EDM69524.2026.11632177. (Scopus)

Соискателем лично выполнен XAI-анализ метода генерации синтетических клеточных данных, сочетающего правила и вариационное моделирование, и интерпретированы результаты вычислительных экспериментов.

9. Trofimov Y., Averkin A., Lebedev M., Lebedev A., Ilin A. Image Segmentation Algorithm Based on Anisotropic Fuzzy Local Convolution // 2026 IEEE 27th International Conference of Young Professionals in Electron Devices and Materials (EDM). Novosibirsk, Russian Federation, 2026. P. 1–5. DOI: 10.1109/EDM69524.2026.11632120. (Scopus)

Автор лично разработал алгоритмическую схему сегментации изображений на основе анизотропной нечёткой локальной свёртки и исследовал её работу вычислительным экспериментом.

10. Trofimov Y., Averkin A., Shevchenko A., Trusov I., Alekseev A., Kondrashova E. Localized SHAP Explainability for Object Detection Based on Parallel Attribution Analysis // Proceedings of the 2026 XXIX International Conference on Soft Computing and Measurements (SCM). Saint Petersburg, Russian Federation, 2026. P. 1–3. DOI: 10.1109/SCM71573.2026.11625907. (Scopus)

Лично соискателем разработан метод локализованной SHAP-объяснимости для детекции объектов на основе параллельного анализа атрибуций и проведена его экспериментальная проверка.

В изданиях, рекомендованных ВАК РФ по специальности 1.2.1:

11. Мультимодальная объяснимость для ОРИТ-сигналов (VTaC): метрические и асимптотические результаты / Ю. В. Трофимов, А. Н. Аверкин, Е. М. Кузнецов, А. П. Еремеев, А. В. Нечаевский // Вестник Тверского государственного университета. Серия: Прикладная математика. – 2025. – № 4. – С. 43–80. – DOI 10.26456/vtpmk757. (ВАК K2)

Трофимовым Ю. В. лично получены метрические и асимптотические результаты для мультимодальной объяснимости ОРИТ-сигналов VTaC и выполнен анализ их интерпретируемости.

12. Временная атрибуция в стохастических диффузионных процессах: математическая формализация, объяснимость и каузальная валидация / Ю. В. Трофимов, А. Н. Аверкин, М. А. Лопатин, И. А. Трусов, И. П. Муравьев, А. С. Ильин, А. В. Шевченко // Вестник Тверского государственного университета. Серия: Прикладная математика. – 2026. – № 1. – С. 89–137. – DOI 10.26456/vtpmk773. (ВАК K2)

Автором лично выполнена математическая формализация временной атрибуции в стохастических диффузионных процессах и проведена каузальная валидация предложенного объяснительного подхода.

13. Согласование глобальных и локальных объяснений моделей искусственного интеллекта с помощью регуляризации чувствительности признаков / Ю. В. Трофимов, А. В. Шевченко, А. Н. Аверкин, И. П. Муравьев // Программные продукты и системы. – 2026. – Т. 39, № 3. – С. 359–367. – DOI 10.15827/0236-235X.155. (ВАК K1)

Соискателем лично разработан способ согласования глобальных и локальных объяснений посредством регуляризации чувствительности признаков и получены результаты его вычислительной оценки.

В других, рекомендованных ВАК РФ:

14. Исследование сигнальных методов выявления состязательных атак на нейронные модели через объяснительную визуализацию / Д. А. Чеколаев, А. В. Шевченко, А. К. Алексеев, Ю. В. Трофимов, А. Н. Аверкин // Вопросы кибербезопасности. – 2025. – № 6(70). – С. 140–146. – DOI 10.21681/2311-3456-2025-6-140-146. – EDN JSORTT. (ВАК K-1)

Лично Трофимовым Ю. В. проведено экспериментальное исследование сигнальных признаков состязательных атак и их выявления средствами объяснительной визуализации нейронных моделей.

15. Алгебраические операции над многоуровневыми объяснениями и квантификация их неопределённости / Ю. В. Трофимов, А. Н. Аверкин, А. Д. Лебедев [и др.] // Мягкие измерения и вычисления. – 2026. Т. 2–2. – DOI: 10.36871/26189976.2026.02-2.008. (ВАК K-3)

Автором лично формализованы алгебраические операции над многоуровневыми объяснениями и разработан подход к количественной оценке неопределённости их объединённого результата.

16. Трофимов, Ю. В. Обзор и анализ методов объяснительного искусственного интеллекта для решения задач геоэкологического районирования и медицинской

профилактики населения / Ю. В. Трофимов, А. Н. Аверкин, Е. Н. Черемисина // Геоинформатика. – 2024. – № 4. – С. 93-118. – DOI 10.47148/1609-364X-2024-4-93-118. – EDN ZXOVIB. (BAK K-2) (RCSI Y3)

Трофимов Ю. В. лично систематизировал и сопоставил методы объяснимого искусственного интеллекта применительно к геоэкологическому районированию и задачам медицинской профилактики населения.

17. Аверкин, А. Н. Применение адаптивных нейро-нечетких систем вывода и объяснительного искусственного интеллекта для мониторинга загрязнения почв и оценки рисков для здоровья населения в городе Дубне / А. Н. Аверкин, Ю. В. Трофимов // Мягкие измерения и вычисления. – 2024. – Т. 84, № 11-1. – С. 70-90. – DOI 10.36871/2618-9976.2024.11.006. – EDN CSJTNK. (BAK K-3)

Соискателем лично разработана схема совместного применения адаптивной нейро-нечёткой системы вывода и ХАИ для мониторинга загрязнения почв и интерпретации рисков для здоровья населения.

18. Трофимов Ю. В. Связь доверенного искусственного интеллекта и ХАИ 2.0: теория и фреймворки / Ю. В. Трофимов, А. Н. Аверкин // Мягкие измерения и вычисления. – 2025. – Т. 90, № 5. – С. 68-84. – DOI 10.36871/2618-9976.2025.05.006. – EDN XUYADQ. (BAK K-3)

Лично автором обоснована взаимосвязь концепций доверенного искусственного интеллекта и ХАИ 2.0, а также структурированы соответствующие теоретические положения и фреймворки.

19. Генерация синтетических данных для обучения модели предсказания степени выживаемости раковых клеток с последующим ХАИ-анализом / Т. В. Ким, И. А. Булякова, Ю. В. Трофимов, А. Н. Аверкин // Мягкие измерения и вычисления. – 2025. – Т. 93, № 8. – С. 28-48. – DOI 10.36871/2618-9976.2025.08.004. – EDN JGDIUZ. (BAK K-3)

Трофимовым Ю. В. лично исследовано влияние синтетических данных на обучение модели прогноза выживаемости раковых клеток и выполнена интерпретация её решений средствами ХАИ.

20. Нечёткий трансформер для мультимодального ИИ: дифференцируемое нечёткое внимание и адаптивные объяснения / Ю. В. Трофимов, А. Н. Аверкин, А. С. Ильин [и др.] // Мягкие измерения и вычисления. – 2025. – Т. 96, № 11. – С. 34-55. – DOI 10.36871/2618-9976.2025.11.003. – EDN NBBCRM. (BAK K-3)

Соискателем лично разработаны элементы нечёткого трансформера, включающие дифференцируемый механизм нечёткого внимания и формирование адаптивных объяснений для мультимодальных данных.

21. Многоуровневая ГИС-архитектура и объяснимый ИИ для территориального районирования при размещении полигонов ТКО: обзор и концепция интеграции / Ю. В. Трофимов, И. З. Каманина, А. Н. Аверкин [и др.] // Мягкие измерения и вычисления. – 2025. – Т. 97, № 12. – С. 23-40. – DOI 10.36871/2618-9976.2025.12.003. – EDN ESPEKG. (BAK K-3)

Автором лично сформирована концепция многоуровневой ГИС-архитектуры с ХАИ для территориального районирования при размещении полигонов ТКО и обоснованы принципы её интеграции.

22. HYBRID-XIRIS: нейро-нечеткая архитектура для объяснительной биометрической идентификации по радужной оболочке глаза / Ю. В. Трофимов, И. А. Матвеев, А. Н. Аверкин [и др.] // Мягкие измерения и вычисления. – 2025. – Т. 97, № 12-2. – С. 119-132. – DOI 10.36871/2618-9976.2025.12-2.013. – EDN QBGESK. (БАК К-3)

Лично Трофимовым Ю. В. разработана нейро-нечёткая архитектура HYBRID-XIRIS и исследована возможность получения объяснений при биометрической идентификации по радужной оболочке глаза.

23. SCR-анализ сквозной объяснимости в AutoML и AutoXAI и синергетическое обоснование архитектурной интеграции / Ю. В. Трофимов, Е. С. Кондрашова, А. Н. Аверкин [и др.] // Мягкие измерения и вычисления. – 2026. – Т. 2. – С. 41-60. – DOI: 10.36871/26189976.2026.01-2.005. (БАК К-3)

Соискателем лично выполнен SCR-анализ сквозной объяснимости в AutoML и AutoXAI и получено архитектурное обоснование их интеграции в единый объяснительный контур.

24. Агентная модель сопоставимого выбора объяснителя для моделей машинного обучения / Ю. В. Трофимов, С. О. Косарев, А. Н. Аверкин [и др.] // Мягкие измерения и вычисления. – 2026. – Т. 1. – С. 56-78. – DOI: 10.36871/2618-9976.2026.03.005. (БАК К-3)

Автор лично разработал агентную модель сопоставимого выбора объяснителя и исследовал механизм принятия решения между альтернативными методами объяснения моделей машинного обучения.

25. Архитектурная модель платформы SynergiXAI для воспроизводимого управления жизненным циклом моделей искусственного интеллекта / Ю. В. Трофимов, Е. М. Кузнецов, А. Н. Аверкин [и др.] // Мягкие измерения и вычисления. – 2026. – Т. 2. – С. 101-119. – DOI: 10.36871/26189976.2026.03-2.011. (БАК К-3)

Трофимовым Ю. В. лично разработана архитектурная модель платформы SynergiXAI, ориентированная на воспроизводимое управление жизненным циклом моделей искусственного интеллекта и результатов их объяснения.

В других изданиях:

26. Автоматизация анализа рентгеновских снимков грудной клетки с использованием методов глубокого обучения и объяснимого искусственного интеллекта / М. И. Беляев, А. Н. Аверкин, Ю. В. Трофимов [и др.] // Системный анализ в науке и образовании. – 2025. – № 2. – С. 32-41. – EDN VHJCOS. (РИНЦ)

Соискателем лично проведено исследование применения методов объяснимого искусственного интеллекта к результатам глубокого анализа рентгеновских снимков грудной клетки и интерпретированы полученные предсказания.

27. Нечёткие продукционные правила и нейросети глубокого обучения: объяснимый искусственный интеллект 2.0 для диагностики коронарных стенозов / Ю.

В. Трофимов, В. С. Семашко, И. П. Муравьев [и др.] // Системный анализ в науке и образовании. – 2025. – № 2. – С. 73-82. – EDN ONRAKG. (РИНЦ)

Лично автором разработана схема объединения нечётких продукционных правил и глубоких нейронных сетей для объяснимой диагностики коронарных стенозов.

28. Новый плитогенно-нечеткий многослойный перцептрон (PN-MLP) на основе нечёткой, интуиционистской, нейтрософической и плитогенной логик с объяснительным искусственным интеллектом второго поколения (ХАИ 2.0) для платформы клинического анализа молочной железы / Ю. В. Трофимов, И. П. Муравьев, А. Н. Аверкин [и др.] // Нечеткие системы и мягкие вычисления. – 2025. – Т. 20, № 1. – С. 5-35. – DOI 10.26456/fssc132. – EDN SAKQYD. (РИНЦ)

Трофимов Ю. В. лично разработал архитектурные принципы PN-MLP, объединяющего нечёткую, интуиционистскую, нейтрософическую и плитогенную логики с ХАИ 2.0 для клинического анализа молочной железы.

29. Объяснительная интервальная ANFIS Типа 2 с интегралом Шоке и самонастраиваемыми параметрическими Т-нормами для прогнозирования и стратификации риска гестационного сахарного диабета / Ю. В. Трофимов, А. В. Шевченко, Е. М. Кузнецов [и др.] // Нечеткие системы и мягкие вычисления. – 2025. – Т. 20, № 2. – С. 43-67. – DOI 10.26456/fssc134. – EDN IVOAWY. (РИНЦ)

Соискателем лично сформирована и экспериментально исследована объяснительная интервальная ANFIS типа 2 с интегралом Шоке и самонастраиваемыми параметрическими Т-нормами для стратификации риска гестационного сахарного диабета.

30. Концепция иерархически организованных объяснимых интеллектуальных систем: синтез глубоких нейросетей, нечеткой логики и инкрементного обучения в медицинской диагностике / Ю. В. Трофимов, А. В. Шевченко, А. Н. Аверкин [и др.] // VI Международная конференция по нейронным сетям и нейротехнологиям (NeuroNT'2025): Материалы конференции, Санкт-Петербург, 04–05 июня 2025 года. – Санкт-Петербург: Санкт-Петербургский государственный электротехнический университет "ЛЭТИ" им. В.И. Ульянова (Ленина), 2025. – С. 3-6. – EDN CJPUMF. (РИНЦ)

Автором лично апробирована концепция иерархической объяснимой интеллектуальной системы и представлены результаты синтеза глубоких нейросетей, нечёткой логики и инкрементного обучения в диагностическом сценарии.

31. Использование нечетких множеств для построения гибридного объяснительного искусственного интеллекта нового поколения / Ю. В. Трофимов, А. Н. Аверкин, А. К. Алексеев, А. В. Шевченко // XXII национальная конференция по искусственному интеллекту с международным участием (КИИ-2025): Труды конференции, Санкт-Петербург, 06–10 октября 2025 года. – Санкт-Петербург: Санкт-Петербургский Федеральный исследовательский центр РАН, 2025. – С. 97-107. – DOI 10.15622/rcai.2025.040

Лично Трофимовым Ю. В. обоснованы принципы использования нечётких множеств при построении гибридного объяснимого искусственного интеллекта нового поколения.

32. Создание прототипа рекурсивного объяснимого искусственного интеллекта на основе поведенческо-логических моделей / Ю. В. Трофимов, А. Н. Аверкин, И. П.

Муравьев [и др.] // XXII национальная конференция по искусственному интеллекту с международным участием (КИИ-2025) : Труды конференции, Санкт-Петербург, 06–10 октября 2025 года. – Санкт-Петербург: Санкт-Петербургский Федеральный исследовательский центр РАН, 2025. – С. 108-119. – DOI 10.15622/rcai.2025.041

Соискателем лично создан прототип рекурсивного объяснимого искусственного интеллекта на основе поведенческо-логических моделей и исследован механизм последовательного формирования объяснений.

33. Трофимов Ю. В. Мягкие измерения нейронечётких составляющих ХАИ 2.0: мультиуровневая гибридная архитектура (CNN + ANFIS) / Ю. В. Трофимов, Е. М. Кузнецов, И. П. Муравьев // Наука настоящего и будущего. – 2025. – Т. 1. – С. 44-47. – EDN WTOSIF.

Автор лично выполнил мягкие измерения нейронечётких составляющих ХАИ 2.0 и оценил многоуровневую гибридную архитектуру, объединяющую CNN и ANFIS.

34. Трофимов Ю. В. Многоуровневый объяснительный ИИ для диагностики пневмонии в рентген-снимках: от Grad-CAM и LIME до генеративного автоописания с MLOPS- интеграцией / Ю. В. Трофимов, И. П. Муравьев, Е. М. Кузнецов // Наука настоящего и будущего. – 2025. – Т. 1. – С. 60-63. – EDN ORUBXV.

Трофимовым Ю. В. лично сформирован многоуровневый конвейер объяснения диагностики пневмонии, объединяющий Grad-CAM, LIME, генеративное автоописание и MLOps-интеграцию.

35. Интеграция объяснимого искусственного интеллекта и спутниковых данных в задаче районирования территории / Ю. В. Трофимов, И. З. Каманина, А. Н. Аверкин [и др.] // Актуальные проблемы экологии и природопользования: сборник научных трудов XXVI Международной научно-практической конференции. В 3 томах, Москва, 25–27 апреля 2025 года. – Москва: Российский университет дружбы народов (РУДН), 2025. – С. 390-394. – EDN KUYLYY.

Соискателем лично разработана схема интеграции ХАИ со спутниковыми данными и проведена её апробация в задаче территориального районирования.

36. Трофимов Ю. В. Объяснительный интеллект как основа доверия и адаптации знаний для устойчивого развития / Ю. В. Трофимов // Искусственный и естественный интеллект: алгоритмы, мышление и образовательные технологии: материалы XXI международного конгресса с элементами научной школы для молодых ученых: в 2-х томах, Москва, 27–28 марта 2025 года. – Москва: Московский университет им. С.Ю. Витте, 2025. – С. 134-151. – EDN AMOTXS. (РИНЦ)

Все изложенные в работе результаты получены Трофимовым Ю. В. лично: сформулировано представление об объяснительном интеллекте как механизме доверия и адаптации знаний для устойчивого развития.

37. Trofimov Yu. V. Hierarchical approach to multimodal Explainable Artificial Intelligence: Incremental integration of explanatory / Yu. V. Trofimov, A. N. Averkin // Scientific research of the SCO countries: synergy and integration: Proceedings of the International Conference, Beijing, 15 января 2025 года. – Beijing: Infiniti, 2025. – P. 130-135. – DOI 10.34660/INF.2025.87.97.124. – EDN CDNNDC.

Автором лично разработан иерархический подход к мультимодальному ХАИ и исследован механизм инкрементального включения объяснений в многоуровневую структуру.

38. Trofimov Yu. V. Zoning based on XAI and knowledge bases: a theoretical and applied approach to managed spatial systems / Yu. V. Trofimov, A. N. Averkin // Science. Education. practice : Proceedings of the International Science Conference, Delhi, 29 января 2025 года. – Delhi: ООО "Инфинити", 2025. – Р. 159-167. – DOI 10.34660/INF.2025.79.43.089. – EDN SONBOR.

Лично соискателем сформулирован теоретико-прикладной подход к районированию управляемых пространственных систем, объединяющий ХАИ и базы знаний.

39. Кузнецов, Е. М. Объяснимый инкрементный подход в диагностике пневмонии: от свёрточных нейросетей до генеративных текстовых заключений / Е. М. Кузнецов, И. П. Муравьев, Ю. В. Трофимов // XIV Конгресс молодых ученых : Тезисы научных работ, Санкт-Петербург, 07–11 апреля 2025 года. – Санкт-Петербург: Национальный исследовательский университет ИТМО, 2025. – С. 815-817. – EDN LPKXHE. (РИНЦ)

Трофимовым Ю. В. лично исследован инкрементный объяснимый подход к диагностике пневмонии, связывающий результаты свёрточной нейронной сети с генеративными текстовыми заключениями.

40. Трофимов Ю. В. Нечёткие продукционные правила и нейросети глубокого обучения: объяснимый искусственный интеллект 2.0 для диагностики коронарных стенозов / Ю. В. Трофимов, В. С. Семашко // XIV Конгресс молодых ученых : Тезисы научных работ, Санкт-Петербург, 07–11 апреля 2025 года. – Санкт-Петербург: Национальный исследовательский университет ИТМО, 2025. – С. 938-939. – EDN TLYRNK. (РИНЦ)

Соискателем лично представлены и апробированы результаты применения нечётких продукционных правил совместно с глубокими нейросетями для интерпретируемой диагностики коронарных стенозов.

Свидетельства о государственной регистрации программ для ЭВМ:

1. Ю. В. Трофимов [и др.] свидетельство о государственной регистрации программы для ЭВМ № 2025690301 Российская Федерация. Веб-приложение для интерпретируемой классификации пневмонии лёгких (версия 1.0): заявл. 25.10.2025; опубл. 06.11.2025

2. Ю. В. Трофимов [и др.] свидетельство о государственной регистрации программы для ЭВМ № 2025690503 Российская Федерация. Веб-приложение для мультиуровневого и интерпретируемого анализа ХОБЛ (версия 1.0): заявл. 25.10.2025; опубл. 07.11.2025

3. Д. А. Чеколаев, Ю. В. Трофимов [и др.] свидетельство о государственной регистрации программы для ЭВМ № 2025690674 Российская Федерация. Интерактивная демонстрационная платформа для изучения принципов объяснимого ИИ через моделирование состязательных атак: заявл. 25.10.2025; опубл. 10.11.2025

Результаты работы обсуждались на:

1. Основные положения диссертационной работы, разработанные модели, методы, алгоритмы и результаты практического использования докладывались и обсуждались на научных семинарах в Национальном исследовательском университете «МЭИ» (2026 год) и Государственном университете «Дубна» (2024-2026 годы).

2. Также результаты представлялись и докладывались автором на 26 международных и всероссийских научных конференциях: AI-ZAMAN 2025 Interdisciplinary Conference on Artificial Intelligence: from theory to practice (Казань, 2025), IEEE 3rd International Conference on Problems of Informatics, Electronics and Radio Engineering (Новосибирск, 2024), VI и VII Международной конференции по нейронным сетям и нейротехнологиям (NeuroNT'2025, 2026, Санкт-Петербург, ЛЭТИ), 11th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID'2025, Дубна), XXII Национальной конференции по искусственному интеллекту с международным участием (КИИ-2025, Санкт-Петербург), IX Международной научной конференции «Интеллектуальные информационные технологии для индустрии» (Федеральная территория «Сириус», 2025), XXXIII Международной конференции «Математика. Компьютер. Образование» (Дубна, 2026), XVII SRPSKI Simpozijum O Hidrogeologiji sa međunarodnim učešćem (Пирот, Сербия, 2024), International Science Conference "Science. Education. Practice" (Дели, Индия, 2025), Scientific Research of the SCO Countries: Synergy and Integration (Пекин, Китай, 2025), XXIX Международная конференция по мягким вычислениям и измерениям (SCM'2026 Санкт-Петербург, ЛЭТИ), 27th International Conference of Young Professionals in Electron Devices and Materials (EDM) (Республика Алтай, 2026), International Conference on Industrial Engineering (Сочи, 2026), International Russian Smart Industry Conference (Сочи, 2026), XIII Международная научно-практическая конференция «Интегрированные модели и мягкие вычисления в искусственном интеллекте» (ИММВ-2026, Коломна), Донецкий международный круглый стол искусственный интеллект: теоретические аспекты и практическое применение (ДМКС-ИИ-2026), а также на XXIX, XXX и XXXI Всероссийских научно-практических конференциях студентов, аспирантов и молодых специалистов (Дубна, 2024, 2025, 2026).

Диссертация соответствует требованиям, установленным разделом 2 Положения о присуждении ученых степеней в ФГБОУ ВО «НИУ «МЭИ», принятого Ученым советом ФГБОУ ВО «НИУ «МЭИ», протокол № 03/18 от 30.03.2018 (в действующей редакции).

Диссертация соответствует требованиям, установленным пунктами 9-14 Положения о присуждении ученых степеней, утвержденного Постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842 (с изменениями и дополнениями).

Общая оценка диссертации и рекомендации к защите

Диссертация Трофимова Юрия Владиславовича на соискание ученой степени кандидата технических наук является научно-квалификационной работой, в которой разработаны модели, методы, алгоритмы и программные решения для многоуровневого формирования, согласования, объединения и оценки объяснений в гетерогенных интеллектуальных системах, реализованные в составе воспроизводимой программной платформы анализа данных. Разработанные решения апробированы на биомедицинских задачах анализа маммографических изображений, диагностики стенозов коронарных артерий и заболеваний органов грудной клетки.

Диссертация «Методы и гетерогенная гибридная платформа многоуровневого ансамблирования объяснений при комплексировании моделей машинного обучения» Трофимова Юрия Владиславовича рекомендуется к защите на соискание ученой степени кандидата технических наук по научной специальности 1.2.1 «Искусственный интеллект и машинное обучение».

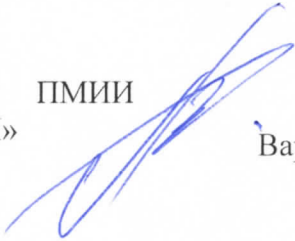
Заключение принято на заседании кафедры прикладной математики и искусственного интеллекта ФГБОУ ВО «НИУ «МЭИ» протокол № 1/4 от «29» сентября 2026 г.

Присутствовали на заседании 29 чел., из них с правом голоса 27 чел.

Результаты голосования:

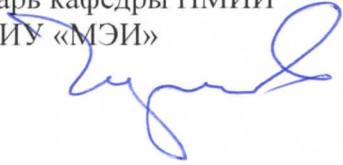
«за» – 27 чел., «против» – 0 чел., «воздержалось» – 0 чел.

Заведующий кафедрой ПМИИ
ФГБОУ ВО «НИУ «МЭИ»
к.т.н., доцент



Варшавский Павел Романович

Ученый секретарь кафедры ПМИИ
ФГБОУ ВО «НИУ «МЭИ»
к.т.н., доцент



Меньшикова Ксения Георгиевна