

ВИ МЕЖДУНАРОДНАЯ НАУЧНО-ПРАКТИЧЕСКАЯ КОНФЕРЕНЦИЯ «ЦИФРОВИЗАЦИЯ ЯЗЫКОВ НАРОДОВ РОССИИ: МАСШТАБИРОВАНИЕ ОПЫТА И ПЕРСПЕКТИВЫ»

VII Международная научно-практическая конференция «Цифровизация языков народов России: масштабирование опыта и перспективы».

Спустя более чем год в Йошкар-Оле на базе Марийского государственного университета 16–17 апреля 2026 г. состоялась VII Международная научно-практическая конференция по цифровизации языков народов России. Очередное мероприятие было посвящено теме «Цифровизация языков народов России: масштабирование опыта и перспективы».

Участники обсудили актуальные вопросы цифровизации языков народов Российской Федерации как средства сохранения, трансляции и актуализации их культурного наследия. Конференция стала важной площадкой для обмена опытом и идеями между специалистами в области языкознания, образования, компьютерной лингвистики, цифровых технологий и сохранения языкового разнообразия.

В этом году гостей конференции встречала выставка цифровых решений для сохранения родных языков — демонстрационная площадка «Цифровизация языков: синергия в партнерстве».

Участниками конференции (117 человек) стали исследователи, программисты, преподаватели, аспиранты и студенты, учителя, специалисты в сфере языковой политики из Алматы (Казахстан), Астаны (Казахстан), Белграда (Сербия), Бишкека (Кыргызстан), Бирска (Республика Башкортостан), Владивостока, Ижевска (Удмуртская Республика), Йошкар-Олы (Республика Марий Эл), Казани (Республика Татарстан), Кирова, Москвы, Петрозаводска (Республика Карелия), Санкт-Петербурга, Саранска (Республика Мордовия), Сыктывкара (Республика Коми), Тараза (Казахстан), Ханты-Мансийска (Ханты-Мансийский автономный округ — Юг-

ра), Улан-Удэ (Республика Бурятия), Уфы (Республика Башкортостан), Чебоксар (Чувашская Республика), Якутска (Республики Саха (Якутия)), Китайской Народной Республики.

Доклад В. И. Токтаровой и А. В. Чемышева (Йошкар-Ола) «Цифровизация языков народов России: опыт Марийского государственного университета» был посвящен цифровизации марийского языка на современном этапе. «Из всех источников, которые содержат информацию о национальных костюмах, узорах, вышивке, антропологических особенностях и так далее, с помощью спецпрограммы мы аннотируем иллюстрации — назовем автора, место съемки. После этого мы с помощью компании Яндекс будем обучать нейросети. Так мы добьемся правильного распознавания особенностей марийского народа», — рассказал Андрей Валерьевич Чемышев. По словам проректора по стратегическому развитию МарГУ Веры Ивановны Токтаровой, марийский язык занимает четвертое место среди популярных языков сервиса «ЯндексПереводчик». «Люди хотят пользоваться родным языком в интернете, а мы помогаем обеспечивать эту возможность через цифровизацию: распознавание речи, электронные словари, системы проверки правописания, синтезатор речи, голосовые помощники. Однако одного голоса мало, и нужны правильные визуальные образы. За минуту можем сгенерировать иллюстрации к забытым сказкам, оживить старые фотографии, создать ИИ-ассистентов носителя языка. Мы в сотрудничестве с Яндексом начали работу над масштабным проектом — созданием уникальной базы данных изображений национальных костюмов и предметов быта марийского народа, — отметила докладчик. — Простыми словами, мы разрабатываем правильную обучающую выборку нейросетей».

На примере выбора нейросетей она показала, как сильно «буксует» ИИ на просьбу показать свадебные наряды в районах Марий Эл и смешивает всевозможные традиции, по итогу выдавая неверный результат. «В перспективе — оцифровка музыки, написанной марийскими композиторами. Когда будут изображения и звук, можно создавать полноценные видео. Кроме того, продолжается работа по синтезу и распознаванию марийской речи, улучшению машинных перевозчиков. В центре внимания — оптическое распознавание символов для марийских текстов, чтобы все сканы, фотографии из газет и журналов были правильно определены ИИ. Один из результатов — удалось обучить говорить нейросеть на марийском языке!».

Доклад А. П. Родионовой в соавторстве с И. П. Новак, Н. А. Пеллинен и Н. Б. Крижановской (Петрозаводск) «Новые инструменты и разделы корпуса ВепКар: к юбилею лингвистической платформы» был посвящен обновлениям в вепско-карельском корпусе ВепКар. «К юбилейным мероприятиям (корпус ведет историю с 2016 года, развиваясь из корпуса вепского языка с 2009 года) лингвистическая платформа ВепКар (Открытый корпус вепского и карельского языков) получила значительные обновления, направленные на расширение функционала и увеличение объема данных. Основные новые инструменты и разделы связаны с развитием речевого модуля, углублением семантической разметки и пополнением корпусов». Это коллекция аудиозаписей, снабженных транскрипцией, разметкой и переводом на русский язык. Инструментарий позволяет слушать речь, синхронизированную из текстов.

А. Н. Михеев (Москва) выступил с докладом «Языки народов России в цифровом пространстве: некоторые наблюдения от команды Яндекса» и призвал более активно сотрудничать с «Яндексом». В декабре 2025 года на IX Форуме «Языковая политика в Российской Федерации» в Москве МарГУ и компания «Яндекс» подписали меморандум о сотрудничестве в области развития

искусственного интеллекта. В рамках партнерства реализуется уникальный проект, направленный на обучение генеративных нейросетей Яндекса этнографически достоверным изображениям представителей марийского народа, его культуры и быта. МарГУ стал первой организацией в России, которая системно занимается исправлением этнических стереотипов в генеративном ИИ. Целью сотрудничества является организация научно-методического обмена данными: университет предоставляет Яндексу достоверные изображения, а также этнографические и антропологические данные о марийском этносе. Это позволит «обучить» нейросети создавать корректные и реалистичные визуальные образы, избегая ошибок и неточностей, которые сейчас часто встречаются при генерации. Данный проект имеет двойное стратегическое значение. Во-первых, он вносит вклад в развитие самих технологий искусственного интеллекта, повышая их точность, культурную чувствительность и способность работать с многообразием человеческих образов. Во-вторых, он является современным инструментом сохранения и защиты культурного богатства марийского народа, позволяя закрепить его визуальный код в глобальном цифровом пространстве.

В докладе М. В. Куцаевой (Москва) «Функционирование марийского языка за пределами Марий Эл (по данным экспедиций 2024–2025 гг.)» было отмечено на материале экспедиций в Республику Башкортостан, Республику Татарстан, Свердловскую область, что результаты исследования позволили сделать некоторые выводы о состоянии чтения художественной литературы и периодики на марийском языке представителями марийских диаспор и выявить трудности с чтением на марийском, обусловленные, помимо распространения современных информационных технологий, снижением общей потребности в чтении среди населения, закрытием школ с преподаванием марийского, уменьшением учащихся марийских классов, закрытием региональных профильных образовательных уч-

реждений по подготовке кадров для национальных школ.

Первая секция «Функционирование и развитие языков в условиях цифровых инноваций 1» была представлена следующими докладами: «Создание медийного образа первого ректора Сыктывкарского университета В. А. Витязевой: взгляд современных студентов» (Е. В. Остапова, В. В. Кутыкин), «Что мы можем сделать для малоресурсных языков?» (Т. Б. Агранат), «Чувашско-марийские языковые контакты в современных условиях» (Э. В. Фомин), «Технология подготовки выборок в среде Sугwin для составления словаря антонимов марийского языка» (А. А. Митрускова), «Памятники карельской письменности на цифровой лингвистической платформе LingvoDoc» (С. В. Нагурная), «Принцип мультимодальности и корпусные технологии в исследовании реальной и виртуальной коммуникации на финно-угорских языках. Подходы, решения, перспективы» (Н. А. Осьмак, И. А. Коваль), «Проблемы распознавания текстов на коми латинизированном алфавите 1930-х годов» (В. Н. Епанов), «Голоса Пармы в эпоху цифровизации» (Е. В. Остапова), «Особенности перевода сложноподчиненных предложений с русского на удмуртский язык с использованием русских союзов (на примере художественных текстов из русско-удмуртского параллельного корпуса)» (А. Ф. Уткина) и др.

На второй секции «Функционирование и развитие языков в условиях цифровых инноваций 2» прозвучали следующие доклады: «Корпус текстов и марийский религиозный культурный код: из опыта расшифровки» (Л. А. Абукаева), «Лингвосенсорика холода: особенности передачи холодových ощущений в машинных переводах с русского на удмуртский язык» (Н. В. Кондратьева), «О разработке электронного удмуртско-русского словаря для обучающихся 5–9 классов, не владеющих удмуртским языком» (Е. С. Рябина), «Адаптация обучающихся-билингвов в образовательной среде университета (на примере ФГБОУ ВО «СГУ имени Пятирима Сорокина)» (С. А. Асанова, Ба-

кытбек кызы Гулкайыр), «Медиапроекты Союза коми молодежи «МИ» как способ продвижения родного языка» (Э. Ю. Шлопова), «Эмодзи и графические символы как неотъемлемая часть современного языка» (Л. П. Водясова, А. И. Кокнаев), «Цифровизация фондов Республиканской детско-юношеской библиотеки им. В. Х. Колумба» (Л. Г. Иксанова), «Марийские устойчивые сравнения с зоонимами (на материалах корпуса марийского языка)» (Ю. В. Полякова), «К вопросу инструментального потенциала фольклорного подкорпуса Национального корпуса марийского языка (на примере частушечного пласта горных мари)» (М. В. Пенькова), «Марийские загадки против искусственного интеллекта: кто кого?» (М. А. Ключева) и др.

Третья секция «Цифровизация языка и культуры: новые форматы и смыслы в мире искусственного интеллекта» собрала, в основном, специалистов по компьютерной лингвистике, цифровым технологиям и сохранению языкового разнообразия. Были заслушаны следующие доклады: «Педагогические и социокультурные последствия искажений культурных кодов марийского этноса в AI-генерируемых образах при реализации цифровых этнокультурных проектов» (В. И. Токтарова, Н. В. Матросова), «Язык оцифровали, что дальше?» (А. Р. Кунафин), «Дизайн символов родных языков Российской Федерации — проблемы цифровизации» (С. Н. Бобрышев), «Lingtrain — сервис для создания параллельных корпусов и параллельных книг на любых языках» (С. А. Аверкиев), «Цифровая модель лексикографирования фразеологизмов с историко-этимологическим компонентом» (И. Р. Касимов), «Корпусы коми языка как эмпирическая база для лингвистических исследований» (Ф. Г. Исмагилова), «Видеоигры как форма популяризации языков и культур народов России» (Ю. М. Алюнина), «Масштабирование электронной картотеки топонимов Татарстана: от регионального ресурса к полиэтничной цифровой платформе тюркской топонимии» (О. Р. Хисамов), «Искусственный интеллект в ра-

боте над исследовательским проектом "Этнокультурный плакат"» (Г. И. Кальцев, В. А. Конюхова, М. Н. Кошкодан, Ю. И. Куратова, Т. А. Шатрова), «Моделирование языкового разнообразия России в цифровой среде: возможности и решения интернет-ресурса "Малые языки России"» (О. А. Казакевич, К. К. Поливанов, Д. Д. Мордашова), «Генерация визуального инструментария полевой лингвистики: применение ИИ для создания картинного опросника для сбора диалектного материала» (В. В. Елышев), «Цифровая образовательная среда университета как условие развития этнокультурной компетентности будущих учителей начальных классов» (В. Ф. Поберезкая), «Корпусные инструменты в поддержку языков народов России» (И. Н. Турабекова) и др.

Четвертая секция «Формирование цифровых компетенций и подготовка специалистов в новой цифровой реальности» собрала, в основном, работников образования и культуры. Были представлены следующие доклады: «Применение цифровых технологий в работе учителя родного языка» (Г. А. Натуральнова, У. В. Рузавина), «Лингвистические и культурные аспекты перевода марийских молитв: от языка оригинала к целевому языку» (Л. А. Абукаева), «Возможности применения технологий ИИ в выполнении выпускных квалификационных работ в форме общественных проектов на языковых направлениях обучения» (М. Ю. Андуганова), «Адаптивные электронные учебники в нейропедагогике: практические подходы к персонализированному обучению» (Ю. Ф. Рустамова), «Переводческая деятельность в эпоху цифровизации» (М. Т. Ипакова), «Югорно: дорога предков. Марийский эпос в цифре» (Е. А. Федоров), «Гибридные технологии в цифровизации родных языков: опыт создания образовательного контента с народными сказками на русском и марийском языках» (Ю. М. Ефремова), «Цифровизация марийского литературного наследия как фактор сохранения языкового разнообразия и укрепления культурного единства: методический аспект» (А. К. Иванова), «Исполь-

зование ИИ-технологий в деятельности преподавателя вуза: вызовы и перспективы» (М. С. Мамонтова), «Информационные технологии при обучении мокшанскому языку: преимущества и недостатки (на примере сервиса «Яндекс Переводчик»)» (М. З. Левина), «Конституционно-правовые гарантии сохранения этнокультурного многообразия в условиях цифровизации: опыт Республики Марий Эл и новая стратегия национальной политики РФ до 2036 года» (С. М. Смышляев), «Особенности создания учебного пособия: "Мастерская народа коми: живое наследие Пармы"» (В. В. Косолапова, Т. Ч. Алиева), «Интеграция технологий ИИ генерации изображений в преподавание иностранных и национальных языков» (С. А. Кременная), «Формирование цифровой грамотности через ИКТ на уроках литературы: интерактивные платформы и мультимедиа» (Т. Н. Беляева), «Формирование и развитие молодежных инициатив в области сохранения финно-угорских народов на примере работы Молодежного Совета Ассоциации финно-угорских народов Российской Федерации» (М. Г. Алгайкин).

В рамках проведения VII Международной научно-практической конференции состоялась творческая встреча «Роль современных медиатехнологий в сохранении культурного наследия» с участием В. В. Умановского, секретаря Союза журналистов России, директора Центрального Дома журналистов, профессора Академии Медиаиндустрии. Общаясь со студентами в неформальной обстановке, он поделился личным опытом работы в федеральных и региональных СМИ и дал советы начинающим авторам.

Заседание круглого стола «Цифровой фольклор и искусственный интеллект: от архива к среде интерпретации» было посвящено актуальным вопросам развития современного фольклорного архива. Круглый стол прошел под руководством Т. А. Золотовой и Л. Х. Давлетшиной. В обсуждении приняли участие профессиональные фольклористы, филологи, специалисты в области IT-технологий, а также студенты, магистры

ранты, аспиранты. Центральной темой стала проблема переосмысления фольклорного архива в условиях цифровой культуры. В ходе обсуждения была сформулирована ключевая идея, что фольклорный архив XXI века — это цифровая экосистема, в которой традиция не фиксируется раз и навсегда, а продолжает развиваться в диалоге исследователя, пользователя и технологий.

Научно-методический семинар «Цифровизация языков народов России: инновации в образовании и исследованиях» собрал государственных служащих, учителей родных языков и литераторов, директоров школ, исследователей, интересующих цифровизацией языков народов Российской Федерации.

На основе результатов фундаментальных и прикладных исследований, представленных на конференции, были предложены следующие актуальные подходы к цифровизации исследовательской практики, а также сохранению языков и культур Российской Федерации:

— разработка нейросетевых технологий;

— создание цифровых языковых архивов: формирование структурированных электронных корпусов текстов, баз данных и голосовых помощников для интеграции языков в цифровую среду;

— полилингвальное образование и цифровые ресурсы: разработка интерактивных образовательных платформ и

цифровых словарей, использующих технологии ИИ для обучения родным языкам;

— усиление прикладных лингвистических исследований: масштабирование опыта создания параллельных корпусов текстов для популярных сервисов.

Мероприятие объединило ведущих экспертов в области компьютерной лингвистики, цифровых технологий и сохранения языкового разнообразия для обсуждения актуальных вопросов, и выработки стратегий развития цифровизации языков народов Российской Федерации.

VII Международная научно-практическая конференция «Цифровизация языков народов России: масштабирование опыта и перспективы» Марийского государственного университета стала важным шагом в развитии цифровых технологий для сохранения языкового наследия и укрепления культурного разнообразия России.

Acknowledgements. The publication costs of this article were covered by the Estonian Academy of Sciences.

ВАЛЕРИЙ МАКСИМОВ (Йошкар-Ола)

Address

Valerij Maksimov
E-mail: sernur@rambler.ru