

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

20–21 марта 2025 г. в г. Йошкар-Оле на базе Марийского государственного университета состоялась VI Международная научно-практическая конференция по цифровизации языков народов России. Очередная мероприятие было посвящено теме «Цифровизация языков народов России: масштабирование опыта и перспективы».

Свое начало этой конференции исходит к истокам 2011 года:

- 1) 22–27 апреля 2011 г., г. Йошкар-Ола — Международная конференция «Языки меньшинств в компьютерных технологиях: опыт, задачи и перспективы»;
- 2) 16–17 марта 2017 г., г. Сыктывкар — Международная научная конференция «Электронная письменность народов Российской Федерации: опыт, проблемы и перспективы»;
- 3) 11–12 декабря 2019 г., г. Уфа — Международная научная конференция «Электронная письменность народов Российской Федерации: опыт, проблемы и перспективы»;
- 4) 23–24 сентября 2021 г., г. Сыктывкар — Международная научно-практическая конференция «Электронная письменность народов Российской Федерации — 2021 & IWCLUL 2021»;
- 5) 21–22 февраля 2024 г., г. Йошкар-Ола — V Международная научно-практическая конференция «Электронная письменность в России: опыт, задачи, векторы развития».

Организаторами конференции выступили Федеральное агентство по делам национальностей России, Правительство Республики Марий Эл, Марийский государственный университет, Общество с ограниченной ответственностью «Яндекс», Марийский научно-исследовательский институт языка, литературы и истории им. В. М. Васильева.

Участники обсудили актуальные проблемы и поделились успешными

практиками в области цифровизации. Конференция стала важной площадкой для обмена опытом и идеями между специалистами в области языкознания, образования и ИТ-технологий.

Участниками конференции (более 80) стали исследователи, программисты, преподаватели, аспиранты и студенты, специалисты в сфере языковой политики (очной и заочной формы) из Андиджана (Узбекистан), Алматы (Казахстан), Белграда (Сербия), Бирска (Республика Башкортостан), Ижевска (Удмуртская Республика), Йошкар-Олы (Республика Марий Эл), Казани (Республика Татарстан), Кирова, Лабытнанги (Ямало-Ненецкий автономный округ), Москвы, Петрозаводска (Республика Карелия), Санкт-Петербурга, Саранска (Республика Мордовия), Сыктывкара (Коми Республика), Ханты-Мансийска (Ханты-Мансийский автономный округ-Югра), Уфы (Республика Башкортостан), Чебоксар (Чувашская Республика).

На пленарном заседании были заслушаны следующие доклады: «Родной язык в цифровом пространстве» П. Н. Тултаева (г. Саранск), который предложил активизировать цифровизацию родных языков. Вопрос о необходимости активной деятельности по сохранению языков коренных народов занимает важное место в общей повестке мирового сообщества. Об этом свидетельствуют глобальные инициативы Организации Объединённых Наций (ООН) и других международных организаций. С помощью языка человек добывает нужную себе информацию, хранит и передает долголетний общественный опыт из поколения в поколение. Это важный элемент, характеризующий образ жизни, саму жизнь народа.

Второй доклад Л. А. Абукаевой, Р. И. Чузаева, А. В. Чемышева (г. Йош-

кар-Ола). «Цифровизация марийского языка: от инициативы к технологическому стандарту», был посвящён цифровизации марийского языка на современном этапе. Часть языков не дождется «спасения» в цифре, — ведь каждый из них очень сложен, а для создания полноценных словарей и переводчиков нужны ресурсы. Необходимы записи живых носителей, оцифровка литературы, исследовательская работа.

Следующий доклад А. Н. Михеева, (заместитель руководителя проекта «Языки народов России», Международная компания публичное акционерное общество «Яндекс», г. Москва) «Сервисы и технологии для языков народов России: опыт компании Яндекс», призвал более активно сотрудничать с «Яндексом». В прошлом году «Яндекс» подписал соглашение с Республикой Марий Эл и Марийским государственным университетом, благодаря которому совместная работа ведется по нескольким направлениям. Это и внедрение марийского языка в сервисы компании, и оцифровка архивных документов, и разработка образовательных программ и интересных проектов.

В докладе В. В. Умановского (директор по развитию Новой интернет-энциклопедии РУВИКИ, г. Москва; его доклад прочитал А. В. Чемышев) «Роль РУВИКИ в цифровизации языков народов России» было отмечено, что за полгода интерфейс площадки был создан для 20 языков: алтайского, башкирского, бурятского, вепского, ингушского, калмыцкого, коми, коми-пермяцкого, ливвиковского, марийского, мокшанского, русского, саха, татарского, тувинского, удмуртского, хакасского, чеченского, чувашского, эрзянского языков. Общий объем составляет около 100 тыс. сообщений, которые синхронизированы между различными языками.

В конце ноября 2023 года для публичного доступа открылись пять разделов: башкирский, марийский, саха, татарский, чеченский разделы. В ближайшее время будут доступны еще шесть: алтайский, мокшанский, тувин-

ский, удмуртский, чувашский, эрзянский. Общее количество энциклопедических статей в этих разделах составит более 1,1 млн единиц.

Интернет-энциклопедия РУВИКИ предлагает нам одну из таких возможностей: площадку творчества, где каждый может поделиться частичкой своих знаний на родном языке.

По окончании пленарного заседания почётные гости и эксперты приняли участие в работе научных секций, каждая из которых была посвящена важным аспектам развития языков и современных технологий.

Первая секция «Информационные технологии и языки: опыт и расширение перспектив» были заслушаны следующие доклады: «Итоги первого мониторинга состояния и развития языков народов Российской Федерации в сфере информационных технологий» (Ф. С. Арония); «Граф знаний — инновационное ИТ-решение для развития ИИ и создания интеллектуальных технологий нового поколения» (Р. А. Гильмуллин, А. Р. Гатиатуллин); «Дизайн национальных символов в цифровых шрифтах» (С. Н. Бобрышев); «Подкорпус памятников письменности корпуса ВенКар» (И. П. Новак, Н. Б. Крижановская); «Перспективы использования интернет-ресурсов в сохранении татарского языка и культурного наследия татарского народа (на примере интерактивных проектов Центра письменного наследия ИЯЛИ АН РТ)» (И. И. Халиков); «Мордовские языки в цифровом пространстве» (И. Н. Рябов); «Электронные проекты ИЯЛИ АН РТ в области лингвистики» (О. Р. Хисамов); «Сравнительный анализ возможностей нейросетей ChatGPT, Yandex GPT и Microsoft Copilot в обработке и генерации текста» (А. К. Кшнякина); «Графико-синтаксические средства как предмет кодирования и декодирования финно-угорского фольклорного текста» (М. Ю. Андуганова); «Фонетические варианты русских заимствований в чувашском языке: опыт предварительного анализа по корпусу чувашского языка» (Д. М. Леонтьев); «Тексты марийской этнической религии на

русском и английском языке: переводчик-исследователь vs ИИ» (Л. А. Абукаева) и др.

На второй секции «Функционирование языков в цифровом пространстве: исследование, сохранение, развитие» были заслушаны такие доклады: «Маркеры диглоссии в языковом ландшафте: киберпространство и реальность» (С. А. Москвичева, Лука Меденица); «Функционирование многокомпонентных сложноподчиненных предложений в современном удмуртском языке (на материале художественных текстов из Национального корпуса удмуртского языка)» (А. Ф. Уткина); «Фольклорный подкорпус Национального корпуса марийского языка: к вопросу создания» (М. В. Пенькова); «О перспективах интеграции науки и образования в региональных центрах развития и сохранения родных языков» (С. А. Ткачев); «Рейтинги частотности марийских глаголов, по корпусным данным» (М. А. Ключева); «Потенциал национального корпуса в составлении «Словаря антонимов марийского языка» (А. А. Митрускова); «Язык в системе этноидентификационных факторов (сравнительный анализ социологических данных 2015 и 2021 годов)» (Р. А. Кудрявцева, В. И. Шабыков); «Перевод текстов официально-делового стиля марийского языка и возможности Яндекс Переводчика» (М. Т. Ипакова); «Сравнительный анализ машинных переводчиков коми языка» (М. С. Федина, Д. А. Левченко); «Возможности использования национального корпуса марийского языка в лингвистических исследованиях» (А. А. Фокина, И. Иванова) и др.

Третья секция «Цифровизация и цифровые технологии в образовании и культуре», собрала, в основном, работников образования и культуры. Были заслушаны следующие доклады: «Филологическое образование в вузах в условиях цифровой трансформации языков народов России (на примере УдГУ)» (Г. А. Глухова, Л. П. Федорова); «Электронные образовательные ресурсы как эффективное средство в преподавании родных языков и литературы» (Н. В. Гусева); «Как извлечь всю пользу из данных? Опыт по развитию языков народов России в Яндекс Переводчике» (П. С. Зеленская); «Лингвоэкология в цифровом формате (о справочно-информационной платформе Белемле.ру)» (Г. Н. Ягафарова); «Возможности медиаплатформы «Мэгърифэт» («Просвещение») в обучении родному (татарскому) языку и литературе» (Г. Н. Мухарлямова, Л. М. Гиниятуллина); «Электронные ресурсы как инструмент повышения мотивации обучающихся к изучению родного (марийского) языка» (Л. А. Сергеева); «Цифровые технологии в сохранении нематериального этнокультурного достояния» (Г. Н. Кадыкова); «Роль интернет-энциклопедии в цифровизации удмуртского языка» (С. Ю. Поторочина); «Курс «Фольклор марийского народа» в электронном пространстве» (Е. Г. Ласточкина); «Возможности сквозных технологий как инструментария реализации погружения в культурную среду при изучении марийского фольклора» (А. Э. Сидорова, Н. В. Матросова); «Применение цифровых источников в преподавании финского языка» (Т. И. Исламаева) и др.

Мастер-классы собрали студентов, ученых и исследователей, интересующих цифровизацией языков.

На основе результатов фундаментальных и прикладных исследований, представленных на пленарных, секционных заседаниях и мастер-классах конференции, были предложены актуальные подходы, связанные с необходимостью усиления разработок в области внедрения цифровых технологий в исследовательскую практику, а также в сохранение и развитие языков и культуры народов Российской Федерации.

Участниками конференции определены высоко значимые перспективные направления, модели развития, основные цели и задачи в области цифровизации языков народов России. Продвижение языков в цифровом и информационном пространстве является приоритетом международного десятилетия языков коренных народов России. На современном этапе разрозненности

ненные усилия исследователей, программистов, языковых активистов доведены до системной работы, имеются значительные достижения в сфере современных ИТ-разработок, создания и применения лингвистических ресурсов (языковых корпусов, электронных словарей, машинных переводчиков, анализаторов и синтезаторов речи).

С учетом высокой потребности и значимости необходимо:

- поддержать проведение исследований, на основании которых могут быть разработаны рекомендации по масштабированию опыта цифровизации языков Российской Федерации;
- продолжить практику проведения научно-практических конференций, посвященных развитию электронной письменности в России, расширив состав экспертов и участников научного форума.

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.

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