

Электронные базы данных и мультимедийные технологии
в филологических исследованиях
кураторы: чл.-к. А.М. Молдован, д.ф.н. В.Л. Кляус

Основными целями и задачами направления «Электронные базы данных и мультимедийные технологии в филологических исследованиях» является внедрение в филологические исследования методик, основанных на использовании информационных технологий, создание информационных систем, баз данных и других электронных ресурсов, в том числе открытого доступа в сети Интернет, посвященных различным аспектам и проблемам филологии.

☐ Проект «**Эпиграфическая письменность Древней Руси (XI–XV вв.): электронная база данных**» (рук. **А.А. Зализняк**, ИС РАН) направлен на создание базы данных, содержащей надписи-граффити XI-XIII вв. Новгородского Софийского собора, церкви Спаса Преображения на Нередице, Георгиевского собора Юрьева монастыря, церкви Николая на Липне, а также древнерусские надписи XI-XV вв. Константинопольского Софийского собора. В рамках проекта осуществляется исследовательская работа по фиксации, прочтению, стратиграфической и внестратиграфической датировке, морфологической разметке древнерусской эпиграфики. Предполагается создание специализированного веб-ресурса и подготовка печатных публикаций.

☐ Проект «Создание геоинформационного ресурса «**Топонимическая карта Олонецкой Карелии**» (рук. **И.И. Муллонен**, ИЯЛИ КарНЦ РАН) направлен на развитие системы геоинформационных баз данных по топонимике Карелии, основанной на Научной картотеке топонимов Карелии и сопредельных областей ИЯЛИ КарНЦ РАН, содержащей более 100 000 топонимов. Олонецкая равнина традиционно характеризуется высокой плотностью населения и именно поэтому хорошо разработана топонимически: экспедиционные исследования за последние 40 лет позволили зафиксировать в этом регионе Карелии не только названия поселений, водных объектов и их частей, но также полей, полян, покосов, бывших подсек, наименований отдельных камней, деревьев, колодцев и т.д.

Информационно-аналитическая система будет включать в себя базу данных топонимов этого региона, электронную картографическую основу для отображения и анализа исследуемых объектов и инструментальные средства, позволяющие выполнять все операции по вводу, сопровождению и анализу данных, в том числе и картографическому. Топонимы Олонецкой Карелии (около 7000) будут расклассифицированы по 20 параметрам таким образом, чтобы обеспечить возможность поиска, выборки и ареализации по языку происхождения, структурным моделям и компонентам.

☉ Проект **«Разработка компьютерной модели «Смысл – Звучащая речь» и электронной базы данных для ее поддержки»** (рук. Ю.Д. Апресян, ИРЯ РАН) направлен на построение компьютерной модели озвучивания русского связного текста и апробация этой модели на экспериментальном массиве текстов. В основе проекта лежит теория интонационных стратегий русской речи (Т.Е. Янко). Экспериментальные наработки по синтезу речи с использованием некоторых морфо-синтаксических сведений о предложении, предоставляемых системой ЭТАП-3 (ИППИ РАН), позволяет успешно реализовать данный проект, решив ряд задач, в том числе – разработку алгоритма выбора словоформы-акцентоносителя в предложении; проведение информационно-коммуникативной разметки, в которой коммуникативно-информационные атрибуты должны приписываться не отдельным словам, а целым группам слов; подготовка фрагментов лексической базы данных и диагностического массива предложений и текстов.

☉ Проект **«Электронная база данных «Русская народная топонимия»** (рук. Е.Л. Березович, ИРЯ РАН), основанный на топонимической полевой картотеке кафедры русского языка и общего языкознания Уральского федерального университета, содержащей около 3 млн. единиц хранения, направлен на развитие электронной базы данных топонимики Русского Севера (Архангельская и Вологодская области) и прилегающей к нему Костромской области. Актуальность обращения к топонимии этих территорий, представляющих собой самостоятельные лингвистически и культурно значимые регионы, связано с тем, что она демонстрирует как собственно русскую архаику, так и результаты активного этнического взаимодействия. В рамках проекта будут разрабатываться способы систематизации топонимических данных, направленных на решение проблемы комплексного электронного описания топонимического материала. Будет проведена первичная электронная обработка полевого топонимического

материала, в результате которой в электронную базу данных войдет более 35 тыс. новых записей топонимов, содержащих информацию по исходной грамматической форме; фонетические и грамматические характеристики; социолингвистические и стилистические пометы; вид географического объекта и его привязка на местности; системно-языковая информация; сведения о включенности объекта в топонимическую микросистему; мотивационное значение и др. характеристики. Участниками проекта также будут проводиться полевые исследования в Костромской области, в ходе которых предполагается фронтальный сбор топонимического материала с его одновременным включением в создаваемую электронную базу.

☐ Проект **«Создание электронной базы данных по «Новому объяснительному словарю синонимов русского языка»** (рук. **Б.Л. Иомдин**, ИРЯ РАН) направлен на перевод в электронную базу данных материалов «Нового объяснительного словарь синонимов русского языка» (НОСС). Принципиальное отличие НОСС от всех других синонимических словарей русского языка состоит в том, что синонимы характеризуются в нем по гораздо большему числу параметров описания, выполненных с учетом всех достижений современной теоретической лингвистики. Актуальность проекта обусловлена востребованностью НОСС в образовательной и научной сфере. Его реализация в виде открытого веб-ресурса даст к материалам словаря доступ массовому пользователю и позволит проводить научные исследования в области лексико-семантической системы современного русского языка.

☐ Проект **«Полная электронная этимологическая база данных по тюркским языкам»** (рук. **А.В. Шеймович**, ИЯз РАН) направлен на создание новой электронной базы данных по тюркским языкам, содержащей пратюркскую, общетюркскую и межтюркскую лексику, которая будет доступна благодаря специализированному веб-интерфейсу в сети Интернет. Реализация проекта позволит реконструировать лексику тюркского праязыка, создавать сравнительные грамматики и этимологические словари по отдельным тюркским группам и языкам, решать другие проблемы тюркологии и алтаистики, а также общетеоретические вопросы исторической типологии лексической семантики. Использование в филологии современных компьютерных технологий соответствует общемировой тенденции развития гуманитарного знания. В целом традиционным исследовательским направлениям это дает возможность

разрабатывать и осуществлять новые методики анализа языка и текста, делая результаты филологических исследований более доступными для общества, что, в свою очередь, позволяет активно внедрять их в современный образовательный процесс и научно-исследовательскую практику.