

Аннотация

Дисциплина «Корпусная лингвистика» входит в образовательную программу высшего образования – программу бакалавриата по направлению подготовки/специальности 45.03.02 «Лингвистика» направленности «Теоретическая и прикладная лингвистика». Дисциплина реализуется кафедрой «№63».

Дисциплина нацелена на формирование у выпускника следующих компетенций:

ПК-3 «Владение методами проведения лингвистических экспертиз»

ПК-5 «Способность решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-лингвистических технологий и с учетом основных требований информационной безопасности»

ПК-7 «Владение основными математико-статистическими методами обработки лингвистической информации с учетом элементов программирования и автоматической обработки лингвистических корпусов»

Содержание дисциплины охватывает круг вопросов, связанных с анализом языка на базе корпусных данных. Корпусная лингвистика как направление прикладной и компьютерной лингвистики. Основные понятия корпусной лингвистики. История создания лингвистических корпусов. Корпусная лингвистика в России. Корпусная русистика.

Типология корпусов. Фундаментальные корпусы - Британский Национальный корпус, Национальный корпус русского языка. Параллельные корпусы. Учебные (Learner) корпусы. Обзор иных существующих корпусов различных типов. Зарубежные национальные корпусы. Корпусы русского языка. Специальные корпусы.

Создание корпусов. Проблемы репрезентативности и выборки текстов.

Разметка. Средства создания и разметки корпусов. Типы разметки. Металингвистическая разметка. Автоматический морфологический и синтаксический анализ в корпусах на примере национальных корпусов. Стандартизация в корпусной лингвистике. Международные стандарты и проекты (TEI, EAGLES, CDIF, XCES и др.).

Использование корпусов – корпусный подход.

Корпусные исследования на примерах. Лексические исследования, базирующиеся на корпусах. Грамматические исследования. Семантические исследования, базирующиеся на корпусах. Использование корпусов в социологии, политологии, методике иностранных языков, переводе и др.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточная аттестация в форме зачета.

Общая трудоемкость освоения дисциплины составляет 2 зачетных единицы, 72 часа.

Язык обучения по дисциплине «русский»