

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Государственный институт русского языка им. А.С. Пушкина»
Филологический факультет
Кафедра общего и русского языкознания

АННОТАЦИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ

Инструменты искусственного интеллекта для анализа языка и текста

Направление подготовки: 45.04.02 Лингвистика

Направленность (профиль): Общая и типологическая лингвистика и приложения
в области языкознания

Квалификация выпускника: магистр

Форма обучения: очная

Разработчик:

Р.Е. Тельпов, кандидат филологических наук, доцент кафедры общего и русского языкознания ФГБОУ ВО «Государственный институт русского языка им. А.С. Пушкина»

1. Цель освоения дисциплины

Целями освоения дисциплины является систематизация и расширение знаний в области новых информационных технологий, формирование у студента навыков использования методов статистической и автоматической обработки лингвистических явлений для решения прикладных задач в науке и осуществления грамотных лингвистических исследований.

2. Место дисциплины в структуре ОПОП ВО

Дисциплина «Инструменты искусственного интеллекта для анализа языка и текста» входит в состав базовой части основной образовательной программы магистратуры (Б1.О.13) и изучается во 2 семестре.

3. Планируемые результаты обучения по дисциплине, соотнесенные с планируемыми результатами освоения ОПОП

В процессе изучения дисциплины формируются следующие компетенции:

Таблица 1

Требования к результатам освоения дисциплины

Формируемые компетенции (код, содержание компетенции)	Планируемые результаты обучения по дисциплине, в соответствии с индикатором достижения компетенции			
	Индикатор достижения компетенции (код, содержание индикатора)	В результате обучения по дисциплине обучающиеся должны:		
		знать	уметь	владеть
1	2	3	4	5
ОПК-3 Способен применять в профессиональной деятельности общедидактические принципы обучения и воспитания, использовать современные методики и технологии организации образовательного процесса;	ОПК-3.2. Адекватно определяет методический потенциал современных электронных обучающих платформ (систем управления обучением) для разработки электронных учебных материалов по иностранным языкам	- основные понятия квантитативной лингвистики; - основные базы данных и информационно-поисковые системы, а также методы работы с этими системами и базами данных	- работать с основными информационно-поисковыми и экспертными системами, системами представления знаний, синтаксического и морфологического анализа, автоматического синтеза, распознавания и понимания речи	- навыками обработки лексикографической информации и автоматизированного перевода, автоматизированными системами идентификации и верификации личности
ОПК-7 Способен работать с основными	ОПК-7.1. Корректно использует профильные	о системных потенциях языка с опорой на	решать задачи определения авторского инварианта	навыками контент-анализа с использованием частотных

информационно-поисковыми и экспертными системами, системами представления знаний и обработки вербальной информации	информационные ресурсы информационно-телекоммуникационной сети «Интернет»	количественные методы исследования языковых единиц		характеристик языковых единиц
	ОПК-7.2. Владеет рациональными приемами поиска и применения программных продуктов лингвистического профиля	принципы создания и функционирования лингвистических компьютерных программ	пользоваться современным лексикографическим инструментарием	навыками использования современного лексикографического инструментария (электронные словари, сетевые словари, тезаурусы) для оптимизации лингвистических исследований

4. Содержание дисциплины

1. Методы статистической обработки данных в лингвистических исследованиях.

Критерий хи-квадрат. Расчет среднеквадратичного отклонения. Расчет стандартизированных остатков. Список Сводеша как объект статистического исследования. Закон Ципфа.

2. Стилеметрия и атрибуция текста.

Понятие стилеметрии и атрибуции. Область их применения. История развития автороведения и стилеметрии. Основные аспекты процесса атрибуции текста. Обзор программного обеспечения для идентификации авторства текстов.

3. Корпусная лингвистика. Национальные корпуса текстов.

Корпусная лингвистика как наука. Понятие «корпус текста». Классификация корпусов. Особенности применения корпусов текстов. Национальные корпуса текстов.

4. Информационные технологии в обработке текстов. Автоматическое реферирование и аннотирование текстов.

Обработка текстов на естественных языках и искусственный интеллект. Перевод, реферат и аннотирование текста. Общие понятия. Автоматический перевод, реферирование и аннотирование. Система автоматического перевода, реферирования и аннотирования текста.