

Правительство Российской Федерации

**Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
«Национальный исследовательский университет
„Высшая школа экономики“»**

Школа лингвистики Гуманитарного факультета

**Рабочая программа дисциплины «Практикум по проектированию и разработке
лингвистических систем и компонентов»**

для образовательной программы «Фундаментальная и компьютерная лингвистика»
направления 45.03.03 «Фундаментальная и прикладная лингвистика»
подготовки бакалавра
3 курс

Разработчики программы:

Т. А. Архангельский, к. ф. н., старший преподаватель (tarkhangelskiy@hse.ru)

Одобрена на заседании школы лингвистики «12» мая 2015 г.

Руководитель школы Е.В. Рахилина _____ [подпись]

Рекомендована Академическим советом образовательной программы
«19» мая 2015 г., Протокол № 4

Утверждена «21» мая 2015 г.

Академический руководитель образовательной программы

Ю.А. Ландер _____ [подпись]

Москва, 2015

*Настоящая программа не может быть использована другими подразделениями
университета и другими вузами без разрешения кафедры-разработчика программы.*



1 Область применения и нормативные ссылки

Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей, ведущих данную дисциплину, учебных ассистентов и студентов направления 45.03.03 «Фундаментальная и компьютерная лингвистика» подготовки бакалавра, изучающих дисциплину «Практикум по проектированию и разработке лингвистических систем и компонентов». Настоящая программа учебной дисциплины устанавливает минимальные требования к знаниям и умениям студента и определяет содержание и виды учебных занятий и отчетности.

Программа разработана в соответствии с:

- образовательным стандартом Федерального государственного автономного образовательного учреждения высшего профессионального образования «Высшая школа экономики», в отношении которого установлена категория «Национальный исследовательский университет» (ФГАОУ ВПО НИУ ВШЭ), протокол от 24.06.2011
- образовательной программой «Фундаментальная и компьютерная лингвистика» подготовки бакалавра;
- рабочим учебным планом ФГАОУ ВПО НИУ ВШЭ на 2015/2016 по направлению 45.03.03 «Фундаментальная и прикладная лингвистика» подготовки бакалавра, утвержденным в 2015 году.

2 Цели освоения дисциплины

Целями освоения дисциплины «Практикум по проектированию и разработке лингвистических систем и компонентов» является развитие и применение на практике навыков, полученных в результате освоения дисциплин «Программирование (язык Python)», «Компьютерные инструменты лингвистического исследования», «Теория языка» и др. Изучение дисциплины предполагает участие студента в учебной группе из 2-6 человек, которая работает над одним из нескольких предложенных проектов. Каждый проект направлен на решение реальной задачи в области теоретической или компьютерной лингвистики, связанную с разработкой компьютерной лингвистической системы или компонента такой системы. Работа над проектами требует синтеза знаний и навыков, полученных в ходе освоения как технических, так и теоретических дисциплин.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

- уметь применять на практике знания и навыки, полученные в ходе освоения ряда дисциплин;
- иметь представление о ходе работы над проектами в области компьютерной лингвистики (разработка технического задания, сбор данных, работа над программным продуктом, тестирование);
- иметь навыки командной работы и уметь пользоваться необходимыми для этого техническими средствами (например, репозиториями)

В результате освоения дисциплины студент осваивает следующие компетенции:

Способен выявлять научную сущность проблем в	СК-Б3
--	-------



профессиональной области.	
Способен решать проблемы в профессиональной деятельности на основе анализа и синтеза	СК-Б4
Способен оценивать потребность в ресурсах и планировать их использование при решении задач в профессиональной деятельности	СК-Б5
Способен работать с информацией: находить, оценивать и использовать информацию из различных источников, необходимую для решения научных и профессиональных задач (в том числе на основе системного подхода)	СК-Б6
Способен работать в команде	СК-Б8
Способен критически оценивать и переосмысливать накопленный опыт (собственный и чужой), рефлексировать профессиональную и социальную деятельность	СК-Б10

Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
Демонстрирует владение основными приемами обработки языковых данных с помощью компьютерных инструментов, осознает задачи моделирования языковых явлений	Практические занятия, самостоятельное чтение, самостоятельные работы (домашние задания)
Демонстрирует способность выстраивать и излагать аргументацию	Практические занятия, домашние работы
Владеет навыками самостоятельного чтения научной литературы	Самостоятельная работа, обсуждения на практических занятиях
Демонстрирует знание возможных подходов к задачам моделирования и обработки языковых данных и проектирования лингвистических систем и компонентов и обосновывает свое решение	Дискуссии на практических занятиях
внятно и грамотно представляет решения задач в письменном виде	Экзамен, работа на практических занятиях
Соотносит простейшие задачи по проектированию лингвистических систем с оптимальными способами решения	Практические занятия, самостоятельная работа

4 Место дисциплины в структуре образовательной программы

Настоящая дисциплина относится к циклу гуманитарных дисциплин, обеспечивающих в конечном итоге профессионально-языковую подготовку бакалавра. Изучение данной дисциплины базируется на следующих дисциплинах: «Программирование (язык Python)», «Компьютерные инструменты лингвистического исследования», «Теория языка». Основные положения дисциплины должны быть использованы в дальнейшем при изучении следующих дисциплин: «Программирование (язык Python)» (4 курс), «Автоматическая обработка естественного языка» (4 курс).



5 Тематический план учебной дисциплины

Название темы	Всего часов	Аудиторные часы	Самостоятельная работа
		Практические занятия	
Этапы разработки (pipeline) лингвистических систем и компонентов		4	8
Совместная работа над проектом: распределение задач, взаимодействие с помощью технических средств (репозитории и т. п.)		4	8
Работа над проектами		18	30
Итого:	72	26	46

6 Формы контроля знаний студентов

Тип контроля	Форма контроля	1 год				Параметры
		1	2	3	4	
Текущий (неделя)	Домашнее задание		*	*	*	
Итоговый	Экзамен				*	

6.1 Критерии оценки знаний, навыков

Своевременное выполнение самостоятельных работ, внятное изложение решений.

Знакомство с основными методами и инструментами, обсуждавшимися на занятиях, владение основными методами анализа.

Знакомство с содержанием материалов, заданных для самостоятельного изучения.

Оценки по всем формам текущего контроля выставляются по десятибалльной шкале

6.2 Порядок формирования оценок по дисциплине

Итоговая оценка складывается из следующих факторов:

- домашнее задание (75%)
- экзамен (25%)

Экзамен состоит в презентации лингвистической системы или компонента, разработанных в ходе освоения дисциплины.

7 Содержание дисциплины

1. Моделирование языковых явлений с помощью формальных моделей и технических средств. Обзор существующих лингвистических систем.



2. Этапы разработки лингвистических систем и их компонентов. Создание технического задания. Сбор и первичная обработка данных. Разработка инструментов в соответствии с техническим заданием; технологии, используемые при разработке лингвистических систем. Основные понятия машинного обучения; обучающая и тестовая выборка. Тестирование лингвистических систем и параметры оценки качества (точность, полнота, F-мера).
3. Совместная работа над проектом: распределение обязанностей, использование репозиторий, баг-трекеров и других технических средств для совместной работы.
4. Выбор одной из предложенных задач (связанных с разработкой корпусов, словарей, справочных систем, систем обработки языковых данных и т. п.) и работа над проектом.

8 Образовательные технологии

Рекомендуемые образовательные технологии включают обсуждения, практические занятия, самостоятельную работу студентов, интерактивные формы занятий (не менее 60% от аудиторной нагрузки – работа в группах, командные обсуждения). Обучение предполагает большую долю самостоятельной внеаудиторной работы, направляемой и контролируемой преподавателем. Студенты имеют возможность получать консультации во внеаудиторное время через систему LMS, а также по электронной почте.

9 Оценочные средства для текущего контроля и аттестации студента

9.1 Тематика заданий текущего контроля

Экзамен состоит в презентации лингвистической системы или компонента, разработанных в ходе освоения дисциплины.

9.2 Вопросы для оценки качества освоения дисциплины

- назвать основные этапы разработки лингвистической системы;
- описать технические средства для взаимодействия во время совместной работы над проектом;
- описать приёмы и технологии, используемые для разработки лингвистических систем;
- назвать типы лингвистических систем и компонентов.

9.3 Примеры заданий промежуточного /итогового контроля

Экзамен состоит в презентации лингвистической системы или компонента, разработанных в ходе освоения дисциплины.

10 Учебно-методическое и информационное обеспечение дисциплины

10.1 Базовый учебник

курс лекций

10.2 Основная литература

курс лекций



10.3 Дополнительная литература

-

11 Материально-техническое обеспечение дисциплины

Компьютеры с установленным интерпретатором языка Python и другими необходимыми программами; проецирующее оборудование; доступ в интернет.