



Национальный исследовательский университет «Высшая школа экономики»
Программа дисциплины «Практикум по проектированию и разработке лингвистических систем и
компонентов» для направления 45.03.03
«Фундаментальная и прикладная лингвистика» подготовки бакалавра

**Федеральное государственное автономное образовательное учреждение
высшего образования
"Национальный исследовательский университет
"Высшая школа экономики"**

Факультет гуманитарных наук
Школа лингвистики

**Рабочая программа дисциплины (научно-исследовательского
семинара)
«Практикум по проектированию и разработке лингвистических систем и
компонентов»**

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

Авторы программы:

Ляшевская Ольга Николаевна, к.ф.н., olesar@yandex.ru

Одобрена на заседании Школы лингвистики ФГН «30» мая 2016 г.
Руководитель Школы лингвистики Е.В. Рахилина

Рекомендована Академическим советом образовательной программы
«01» июня 2016 г., № протокола 10

Утверждена «01» июня 2016 г.
Академический руководитель образовательной программы
Ю.А. Ландер

Москва, 2016

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



1 Область применения и нормативные ссылки

Научно-исследовательский семинар «Практикум по проектированию и разработке лингвистических систем и компонентов» являются частью учебной программы третьего и четвертого года подготовки бакалавров направления «Фундаментальная и компьютерная лингвистика».

Программа разработана в соответствии с:

- Образовательным стандартом НИУ Высшей школы экономики, в отношении которого установлена категория «национальный исследовательский университет»;
- Образовательной программой направления «Фундаментальная и прикладная лингвистика» подготовки бакалавра;
- Рабочим учебным планом НИУ ВШЭ по направлению подготовки «Фундаментальная и компьютерная лингвистика».

2 Цели освоения дисциплины

Целями освоения дисциплины являются формирование навыков по осуществлению научных исследований, в том числе навыков работы в команде, формулирования и решения поставленных задач, планирования проекта, методологически корректного сбора данных, проведения экспериментов и экспертизы их результатов, публичного представления результатов проекта.

3 Компетенции обучающегося, формируемые в результате освоения дисциплины

В результате освоения дисциплины студент должен:

Знать

- основные принципы разработки и создания лингвистических корпусов и ресурсов;
- принципы разметки лингвистических данных;
- преимущества и недостатки формальных моделей, лежащих в основе различных модулей автоматической обработки текста;
- необходимые этапы анализа данных и проблемы, возникающие при моделировании каждого из этапов;
- основные алгоритмы, используемые для построения автоматического синтаксического анализа;
- наиболее известные доступные для свободного использования компоненты автоматического анализа, в том числе синтаксические и морфологические парсеры, системы распознавания именованных сущностей (NER);
- принципы оценки качества таких систем.

Уметь

- проводить оценку качества систем автоматического морфологического, синтаксического и семантического анализа;
- использовать соответствующие модули в различных приложениях;
- участвовать в разработке отдельных модулей в системах автоматического анализа текстов.

Иметь навыки (приобрести опыт)

- самостоятельной разработки лингвистических ресурсов



В результате освоения дисциплины студент осваивает следующие компетенции:

Компетенция	Код по ФГОС/ НИУ	Дескрипторы – основные признаки освоения (показатели достижения результата)	Формы и методы обучения, способствующие формированию и развитию компетенции
Способен планировать научно-исследовательскую деятельность, проводить самостоятельные исследования и получать новые научные результаты в области профессиональной деятельности	ПК-1	понимает постановку задачи в области компьютерной лингвистики, использует современные методы тестирования качества, применяет современные подходы к решению задач в области компьютерной лингвистики	проектная работа
Умеет пользоваться лингвистически ориентированными программными продуктами, способен разрабатывать и совершенствовать электронные языковые ресурсы, лингвистические компоненты информационных систем, лингвистические модули компьютерных систем обучения	ПК-17	подбирает материал для решения исследовательской задачи, знает основные лингвистические ресурсы, владеет цепочкой лингвистической обработки текста, умеет использовать среды разработки лингвистических систем	проектные решения, использование сред разработки лингвистических технологий

4 Место дисциплины в структуре образовательной программы

В ходе изучения дисциплины студенты учатся применять на практике понятия, усвоенные в рамках теоретических курсов. Кроме того, в ходе курса студенты могут использовать знания, полученные при изучении программирования и математики на школьном и начальном университетском уровне (в области обработки текста, теории вероятности и статистики).

Полученные знания и навыки могут быть эффективно использованы при написании курсовых и выпускных квалификационных работ, при создании электронных ресурсов (баз данных, словарей и корпусов) для самостоятельной работы по лингвистическим дисциплинам. Студенты получают подготовку для дальнейшей профильной профессиональной деятельности в следующих областях:

- компьютерная лингвистика;
- корпусная лингвистика;
- компьютерная лексикография;
- лексикология; русистика; типология.



5 Тематический план учебной дисциплины

Семинар рассчитан на два модуля, общее количество аудиторных часов — 32 (8 пар лекций в каждом модуле).

№	Название раздела	Всего часов	Аудиторные часы			Самостоятельная работа
			Лекции	Семинары	Практические занятия	
	3 курс:					
1	Постановка задачи, планирование	16		6		10
2	Сбор данных	46		6		40
3	Программирование по проекту / Разметка данных / Анализ данных	54		4		50
	Всего	114		14		100
	4 курс:					
1	Программирование по проекту / Разметка данных / Анализ данных	44		8		36
2	Экспертиза результатов, публикация результатов	28		18		10
	ИТОГО:	72		26		46

6 Формы контроля знаний студентов

Тип контроля	Форма контроля	3 курс				4 курс			
		-	2	3	4	1	-	-	-
Текущий (неделя)	Контрольная работа								
	Эссе								
	Реферат								
	Коллоквиум			*					
	Домашнее задание					3			
Промежуточный	Экзамен				*				Предварительная презентация проекта
Итоговый	Экзамен						*		Открытая презентация проекта

6.1 Критерии оценки знаний, навыков

По самостоятельной работе и презентации работы над проектом выставляется интегрированная оценка по десятибалльной шкале. Финальная оценка проекта выставляется по 10 бальной шкале

Материалы проектных работ публикуются в открытом доступе в Интернет на одной из общедоступных платформ. Презентации результатов работы над проектом проходят в регулярном режиме, не реже, чем раз в две недели (определяется куратором проекта).

7 Содержание дисциплины

Содержание работы над конкретным проектом определяется куратором, ведущим проект.

8 Образовательные технологии

Обсуждение проектов происходит в интерактивном режиме и сопровождаются проведением дискуссий и решением кейсов на семинарах и в ходе самостоятельной работы

9 Оценочные средства для текущего контроля и аттестации студента

9.1 Примеры заданий промежуточного/итогового контроля

1. Изложите суть и задачу проекта
2. Напишите план реализации проекта
3. Подготовьте презентацию плана
4. Подготовьте презентацию данных, проблем, связанных с данными, и методов их решения
5. Подготовьте исходную презентацию проекта
6. Подготовьте коллективный обзор на одну из тем, предложенных куратором

10 Порядок формирования оценок по дисциплине

Преподаватель оценивает работу студентов на семинарских и практических занятиях. Оценки за работу на семинарских и практических занятиях преподаватель выставляет в рабочую ведомость. Накопленная оценка по 10-ти балльной шкале за работу на семинарских и практических занятиях определяется перед промежуточным или итоговым контролем - $O_{\text{аудиторная}}$.

Преподаватель оценивает самостоятельную работу студентов. Оценки за самостоятельную работу студента преподаватель выставляет в рабочую ведомость. Накопленная оценка по 10-ти балльной шкале за самостоятельную работу определяется перед промежуточным или итоговым контролем – $O_{\text{сам. работа}}$.

Преподаватель оценивает проектную работу студентов, проекты выполняются группой студентов. Накопленная оценка по 10-ти балльной шкале за работу на аудиторных занятиях определяется перед промежуточным или итоговым контролем – $O_{\text{проектная}}$

Результирующая оценка за промежуточный (итоговый) контроль в форме экзамена выставляется по следующей формуле, где $O_{\text{экзамен}}$ – оценка за работу непосредственно на экзамене:

$$O_{\text{промежуточный/итоговый}} = k_1 \cdot O_{\text{экзамен}} + k_2 \cdot O_{\text{проектная}} + k_3 \cdot O_{\text{сам. работа}}$$

При этом удельный вес форм контроля распределяется следующим образом:

$$k_1 = 0,2 \quad k_2 = 0,4 \quad k_3 = 0,4$$

Способ округления накопленной оценки промежуточного (итогового) контроля в форме экзамена: в пользу студента

На передаче студенту не предоставляется возможности получить дополнительный балл для компенсации оценки за текущий контроль.

11 Учебно-методическое и информационное обеспечение дисциплины

11.1 Базовый учебник

не предполагается

11.2 Основная литература

Определяется индивидуально для каждого проекта куратором проекта.

12 Материально-техническое обеспечение дисциплины

Занятия должны проходить в компьютерных классах или аудиториях, необходим проектор.