

Паспорт образовательной программы «Компьютерная лингвистика»

Обучение ведется по направлению

45.04.03 Фундаментальная и прикладная лингвистика

Утверждение программы

Приказ №6.18.1.-01/1507-07 от 15.07.2014

Образовательный стандарт НИУ ВШЭ

[Образовательный стандарт НИУ ВШЭ уровень высшего образования – магистратура \(с 2022 года набора\) \(электронная подпись\)](#) (PDF, 1006 Кб)

Дата обновления паспорта

01.10.2020 решением академического руководителя

Сетевая форма реализации

Нет

Срок, форма обучения и объем

2 года

Очная форма обучения, 120 з.е.

Язык реализации



Обучение ведется на русском и частично на английском языке

Квалификация выпускника

Магистр

Программа двух дипломов

Нет

Применение электронного обучения, дистанционных образовательных технологий

С применением

Траектории

2025/2026 УЧЕБНЫЙ ГОД

Прикладная компьютерная лингвистика

Вид: Прикладная

Наставник: [Клышинский Эдуард Станиславович](#)

Язык реализации: Русский

Применение электронного обучения и дистанционных технологий: С применением

Ключевые образовательные результаты:

KOP-1	Знает основные принципы работы с данными, владеет современными инструментами анализа данных на базовом уровне, в т.ч. навыками программирования, алгоритмизации и математическими методами при решении задач анализа данных
KOP-2	Знает историю и состояние компьютерной лингвистики, существующие подходы, направления, понимает связи с другими гуманитарными и естественно-научными дисциплинами, понимает возможности и ограничения методов компьютерной лингвистики
KOP-3	Умеет формулировать исследовательскую задачу в области компьютерной лингвистики, выбирает и применяет необходимые методы и инструменты ее решения
KOP-4	Владеет навыками проектного менеджмента применительно к проектам в области компьютерной лингвистики, способен работать в междисциплинарной проектной команде и организовать ее работу
KOP-5	Владеет методами и инструментами компьютерной лингвистики на уровне, достаточном для решения профессиональных задач

Характеристика профессиональной деятельности:

Выпускник магистратуры по компьютерной лингвистике факультета гуманитарных наук НИУ ВШЭ:

- имеет хороший теоретический бэкграунд в теоретической лингвистике, математических методах обработки данных;
- знает, как устроена современная электронная языковая среда, уметь увидеть в ней новые проектные задачи; уметь их формулировать и формализовать;
- умеет работать с данными: понимать, какие именно данные нужны для решения поставленной задачи, как взаимодействуют лингвистические и иные данные, откуда их можно получить;
- понимает, какие языковые структуры лежат за той или иной локальной задачей, каковы лингвистические и экстралингвистические свойства текстов, используемых для решения задач;
- знает, какие лингвистические технологии следует применять в зависимости от условий задачи, понимать на какие лингвистические данные и ресурсы может опираться технология;

- умеет программировать прототипы и модели решений, знает алгоритмы машинного и глубокого обучения, умеет подготавливать необходимые лингвистические ресурсы;
- умеет проектировать цепочку процесса обработки языковых данных и интерпретировать результаты автоматической обработки лингвистических данных;
- умеет оценивать трудоемкость разных решений и пороги допустимых ошибок и выбирать оптимальную методологию тестирования.

Возможности труда выпускника специализации по компьютерной лингвистике, кроме собственно академической карьеры в России или за рубежом включают в себя работу в качестве компьютерного лингвиста в коммерческих компаниях. Среди российских работодателей интерес к выпускникам программы проявляют такие «технологические гиганты», как компании Сбербанк, Яндекс, ABBYY, МТС, Mail.ru и др. В то же время, компьютерные лингвисты востребованы и в стартапах, небольших компаниях, чей бизнес связан с научными и информационными технологиями.

Перечень профессиональных компетенций:

ПК-1	Способен планировать научно-исследовательскую деятельность, проводить самостоятельные исследования и получать новые научные результаты в области профессиональной деятельности
ПК-10	Способен осуществлять письменный перевод различных типов текстов (в том числе художественных произведений) и сопровождать их необходимым редакторским комментарием
ПК-11	Способен проводить квалифицированное переводческое сопровождение международных мероприятий разного типа (в том числе научных) с использованием современных информационных технологий и программного обеспечения
ПК-12	Способен осуществлять редактирование научной и научно-популярной литературы по лингвистике и смежным областям, в том числе переводы с русского языка на иностранный и с иностранного на русский
ПК-13	Способен осуществлять лингвистическую обработку текстов в производственно-практических целях
ПК-14	Способен разрабатывать и совершенствовать электронные языковые ресурсы, лингвистические компоненты информационных систем, лингвистические модули компьютерных систем обучения
ПК-15	Способен проводить квалифицированный анализ различных типов устного и письменного дискурса в производственно-практических целях
ПК-16	Способен разрабатывать методики и алгоритмы дигитализации гуманитарного наследия, в том числе специальные языковые разметки корпусов
ПК-17	Способен разработать концепцию, осуществить подготовку, проведение и организационное сопровождение научного мероприятия, в том числе международного характера
ПК-18	Способен определять, транслировать общие цели в профессиональной и социальной деятельности
ПК-19	Способен порождать принципиально новые идеи и продукты, обладает креативностью, инициативностью

ПК-2	Способен проводить теоретические и практические занятия по дисциплинам предметного поля, создавать методические пособия и подготавливать учебно-методические и научно-методические материалы, в т.ч. междисциплинарного характера
ПК-20	Способен создавать, описывать и ответственно контролировать выполнение технологических требований и нормативов в профессиональной деятельности
ПК-3	Способен произвести углубленный анализ проблем, постановку и обоснование проектных задач, связанных с лингвистической проблематикой, моделированием и разработкой языковых компонент в междисциплинарных проектах
ПК-4	Способен самостоятельно разработать методику и инструментарий для осуществления исследовательской и проектной деятельности в области фундаментальной и прикладной лингвистики
ПК-5	Способен разрабатывать и управлять проектами (подпроектами), анализировать риски
ПК-6	Способен планировать и реализовать проекты в сфере межкультурной коммуникации, в том числе направленные на изучение языка как объекта культуры
ПК-7	Способен проводить анализ качества языковых данных, корпусов, систем, использующихся для автоматической обработки естественного языка
ПК-8	Способен обеспечивать аналитическую поддержку выработки и принятия решений в сферах, касающихся языковой политики, мониторинга текстовых данных, сохранения языкового наследия
ПК-9	Способен разрабатывать структуру и содержание образовательного процесса в области лингвистики, создавать учебные материалы и прикладные программы языкового обучения в системе общего и среднего профессионального образования, дополнительного образования

Характеристика образовательных модулей:

АДАПТАЦИОННЫЙ КУРС:

Студенты, не имеющие образования по направлению “Фундаментальная и прикладная лингвистика”, могут прослушать курс “Введение в лингвистику”

ОБЩАЯ БАЗОВАЯ ЧАСТЬ

В общей базовой части представлены две ключевых дисциплины, обеспечивающих высокий уровень лингвистической подготовки. Задача этих дисциплин состоит в том, чтобы познакомить студентов с современными подходами к анализу языковых структур и научить их пользоваться современными методами и инструментами для анализа данных. Эти задачи обеспечиваются курсами

- Формальные модели в естественном языке
- Функциональные и когнитивные модели в естественном языке

ДИСЦИПЛИНЫ СПЕЦИАЛИЗАЦИИ КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА

Пул обязательных дисциплин по компьютерной лингвистике включает в себя собственно курс по компьютерной лингвистике, а также тематически скординированные с ним курсы программирования на языке Python и математических оснований компьютерной лингвистики. На втором году в обязательный

цикл входит курс Веб-девелопмента. Практические компетенции, полученные в рамках этого курса используются студентами для создания веб-приложений проектов, которые они разрабатывают на проектном семинаре.

Дисциплины вариативной части направлены на расширение практических компетенций и теоретических знаний студентов в области наук о данных (курсы Анализ лингвистических данных: квантизативные методы и визуализации, Машинное обучение).

На втором году обучения студентам предлагаются дисциплины по выбору, предусматривающие углубленное погружение в одно из направлений компьютерной лингвистики. Это курс по нейросетям - введение в глубокое обучение, курс по анализу социальных сетей, курс по онтологиям и семантическим технологиям.

ПРАКТИКИ И НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА

Научно-исследовательская работа формируется из трех компонентов: проектно-исследовательского семинара, в рамках которого студенты выполняют проекты в группах, индивидуальной научно-исследовательской работы, в результате которой студент пишет и защищает курсовую работу, и мастер-классов, на которых студенты знакомятся с профессионалами и экспертами в изучаемых областях.

Студенческая практика организована на базе ведущих компаний и организаций, занимающихся компьютерной лингвистикой.

ИТОГОВАЯ ГОСУДАРСТВЕННАЯ АТТЕСТАЦИЯ

Итоговая государственная аттестация проходит в формате защиты дипломной работы с привлечением внешних оппонентов, в том числе (по возможности) из зарубежных вузов или институтов.

Теоретические основы компьютерной лингвистики

Вид: Исследовательская

Наставник: [Клышинский Эдуард Станиславович](#)

Язык реализации: Русский и английский

Применение электронного обучения и дистанционных технологий: С применением

Ключевые образовательные результаты:

КОР-1	Знает основные принципы работы с данными, владеет современными инструментами анализа данных на базовом уровне, в т.ч. навыками программирования, алгоритмизации и математическими методами при решении задач анализа данных
КОР-2	Знает историю и состояние компьютерной лингвистики, существующие подходы, направления, понимает связи с другими гуманитарными и естественно-научными дисциплинами, понимает возможности и ограничения методов компьютерной лингвистики
КОР-3	Умеет формулировать исследовательскую задачу в области компьютерной лингвистики, выбирает и применяет необходимые методы и инструменты ее решения

КОР-4	Владеет навыками проектного менеджмента применительно к проектам в области компьютерной лингвистики, способен работать в междисциплинарной проектной команде и организовать ее работу
КОР-5	Владеет методами и инструментами компьютерной лингвистики на уровне, достаточном для решения профессиональных задач

Характеристика профессиональной деятельности:

Выпускник магистратуры по компьютерной лингвистике факультета гуманитарных наук НИУ ВШЭ:

- имеет хороший теоретический бэкграунд в теоретической лингвистике, математических методах обработки данных;
- знает, как устроена современная электронная языковая среда, уметь увидеть в ней новые проектные задачи; уметь их формулировать и формализовать;
- умеет работать с данными: понимать, какие именно данные нужны для решения поставленной задачи, как взаимодействуют лингвистические и иные данные, откуда их можно получить;
- понимает, какие языковые структуры лежат за той или иной локальной задачей, каковы лингвистические и экстралингвистические свойства текстов, используемых для решения задач;
- знает, какие лингвистические технологии следует применять в зависимости от условий задачи, понимать на какие лингвистические данные и ресурсы может опираться технология;
- умеет программировать прототипы и модели решений, знает алгоритмы машинного и глубокого обучения, умеет подготавливать необходимые лингвистические ресурсы;
- умеет проектировать цепочку процесса обработки языковых данных и интерпретировать результаты автоматической обработки лингвистических данных;
- умеет оценивать трудоемкость разных решений и пороги допустимых ошибок и выбирать оптимальную методологию тестирования.

Возможности трудоустройства выпускника специализации по компьютерной лингвистике, кроме собственно академической карьеры в России или за рубежом включают в себя работу в качестве компьютерного лингвиста в коммерческих компаниях. Среди российских работодателей интерес к выпускникам программы проявляют такие «технологические гиганты», как компании Сбербанк, Яндекс, АБВYY, МТС, Mail.ru и др. В то же время, компьютерные лингвисты востребованы и в стартапах, небольших компаниях, чей бизнес связан с наукоемкими и информационными технологиями.

Перечень профессиональных компетенций:

ПК-1	Способен планировать научно-исследовательскую деятельность, проводить самостоятельные исследования и получать новые научные результаты в области профессиональной деятельности
ПК-10	Способен осуществлять письменный перевод различных типов текстов (в том числе художественных произведений) и сопровождать их необходимым редакторским комментарием
ПК-11	Способен проводить квалифицированное переводческое сопровождение международных мероприятий разного типа (в том числе научных) с использованием современных информационных технологий и программного обеспечения

ПК-12	Способен осуществлять редактирование научной и научно-популярной литературы по лингвистике и смежным областям, в том числе переводы с русского языка на иностранный и с иностранного на русский
ПК-13	Способен осуществлять лингвистическую обработку текстов в производственно-практических целях
ПК-14	Способен разрабатывать и совершенствовать электронные языковые ресурсы, лингвистические компоненты информационных систем, лингвистические модули компьютерных систем обучения
ПК-15	Способен проводить квалифицированный анализ различных типов устного и письменного дискурса в производственно-практических целях
ПК-16	Способен разрабатывать методики и алгоритмы дигитализации гуманитарного наследия, в том числе специальные языковые разметки корпусов
ПК-17	Способен разработать концепцию, осуществить подготовку, проведение и организационное сопровождение научного мероприятия, в том числе международного характера
ПК-18	Способен определять, транслировать общие цели в профессиональной и социальной деятельности
ПК-19	Способен порождать принципиально новые идеи и продукты, обладает креативностью, инициативностью
ПК-2	Способен проводить теоретические и практические занятия по дисциплинам предметного поля, создавать методические пособия и подготавливать учебно-методические и научно-методические материалы, в т.ч. междисциплинарного характера
ПК-20	Способен создавать, описывать и ответственно контролировать выполнение технологических требований и нормативов в профессиональной деятельности
ПК-3	Способен произвести углубленный анализ проблем, постановку и обоснование проектных задач, связанных с лингвистической проблематикой, моделированием и разработкой языковых компонент в междисциплинарных проектах
ПК-4	Способен самостоятельно разработать методику и инструментарий для осуществления исследовательской и проектной деятельности в области фундаментальной и прикладной лингвистики
ПК-5	Способен разрабатывать и управлять проектами (подпроектами), анализировать риски
ПК-6	Способен планировать и реализовать проекты в сфере межкультурной коммуникации, в том числе направленные на изучение языка как объекта культуры
ПК-7	Способен проводить анализ качества языковых данных, корпусов, систем, использующихся для автоматической обработки естественного языка
ПК-8	Способен обеспечивать аналитическую поддержку выработки и принятия решений в сферах, касающихся языковой политики, мониторинга текстовых данных, сохранения языкового наследия

ПК-9	Способен разрабатывать структуру и содержание образовательного процесса в области лингвистики, создавать учебные материалы и прикладные программы языкового обучения в системе общего и среднего профессионального образования, дополнительного образования
------	---

Характеристика образовательных модулей:

АДАПТАЦИОННЫЙ КУРС:

Студенты, не имеющие образования по направлению “Фундаментальная и прикладная лингвистика”, могут прослушать курс “Введение в лингвистику”

ОБЩАЯ БАЗОВАЯ ЧАСТЬ

В общей базовой части представлены две ключевых дисциплины, обеспечивающих высокий уровень лингвистической подготовки. Задача этих дисциплин состоит в том, чтобы познакомить студентов с современными подходами к анализу языковых структур и научить их пользоваться современными методами и инструментами для анализа данных. Эти задачи обеспечиваются курсами

- Формальные модели в естественном языке
- Функциональные и когнитивные модели в естественном языке

ДИСЦИПЛИНЫ СПЕЦИАЛИЗАЦИИ КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА

Пул обязательных дисциплин по компьютерной лингвистике включает в себя собственно курс по компьютерной лингвистике, а также тематически скоординированные с ним курсы программирования на языке Python и математических оснований компьютерной лингвистики. На втором году в обязательный цикл входит курс Веб-девелопмента. Практические компетенции, полученные в рамках этого курса используются студентами для создания веб-приложений проектов, которые они разрабатывают на проектном семинаре.

Дисциплины вариативной части направлены на расширение практических компетенций и теоретических знаний студентов в области наук о данных (курсы Анализ лингвистических данных: квантизативные методы и визуализации, Машинное обучение).

На втором году обучения студентам предлагаются дисциплины по выбору, предусматривающие углубленное погружение в одно из направлений компьютерной лингвистики. Это курс по нейросетям - введение в глубокое обучение, курс по анализу социальных сетей, курс по онтологиям и семантическим технологиям.

ПРАКТИКИ И НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА

Научно-исследовательская работа формируется из трех компонентов: проектно-исследовательского семинара, в рамках которого студенты выполняют проекты в группах, индивидуальной научно-исследовательской работы, в результате которой студент пишет и защищает курсовую работу, и мастер-классов, на которых студенты знакомятся с профессионалами и экспертами в изучаемых областях.

Студенческая практика организована на базе ведущих компаний и организаций, занимающихся компьютерной лингвистикой.

ИТОГОВАЯ ГОСУДАРСТВЕННАЯ АТТЕСТАЦИЯ

Итоговая государственная аттестация проходит в формате защиты дипломной работы с привлечением внешних оппонентов, в том числе (по возможности) из зарубежных вузов или институтов.

2024/2025 УЧЕБНЫЙ ГОД

Прикладная компьютерная лингвистика

Вид: Прикладная

Наставник: [Клышинский Эдуард Станиславович](#)

Язык реализации: Русский

Применение электронного обучения и дистанционных технологий: С применением

Ключевые образовательные результаты:

KOP-1	Знает основные принципы работы с данными, владеет современными инструментами анализа данных на базовом уровне, в т.ч. навыками программирования, алгоритмизации и математическими методами при решении задач анализа данных
KOP-2	Знает историю и состояние компьютерной лингвистики, существующие подходы, направления, понимает связи с другими гуманитарными и естественно-научными дисциплинами, понимает возможности и ограничения методов компьютерной лингвистики
KOP-3	Умеет формулировать исследовательскую задачу в области компьютерной лингвистики, выбирает и применяет необходимые методы и инструменты ее решения
KOP-4	Владеет навыками проектного менеджмента применительно к проектам в области компьютерной лингвистики, способен работать в междисциплинарной проектной команде и организовать ее работу
KOP-5	Владеет методами и инструментами компьютерной лингвистики на уровне, достаточном для решения профессиональных задач

Характеристика профессиональной деятельности:

Выпускник магистратуры по компьютерной лингвистике факультета гуманитарных наук НИУ ВШЭ:

- имеет хороший теоретический бэкграунд в теоретической лингвистике, математических методах обработки данных;
- знает, как устроена современная электронная языковая среда, уметь увидеть в ней новые проектные задачи; уметь их формулировать и формализовать;
- умеет работать с данными: понимать, какие именно данные нужны для решения поставленной задачи, как взаимодействуют лингвистические и иные данные, откуда их можно получить;
- понимает, какие языковые структуры лежат за той или иной локальной задачей, каковы лингвистические и экстралингвистические свойства текстов, используемых для решения задач;
- знает, какие лингвистические технологии следует применять в зависимости от условий задачи, понимать на какие лингвистические данные и ресурсы может опираться технология;
- умеет программировать прототипы и модели решений, знает алгоритмы машинного и глубокого

обучения, умеет подготавливать необходимые лингвистические ресурсы;

- умеет проектировать цепочку процесса обработки языковых данных и интерпретировать результаты автоматической обработки лингвистических данных;
- умеет оценивать трудоемкость разных решений и пороги допустимых ошибок и выбирать оптимальную методологию тестирования.

Возможности трудоустройства выпускника специализации по компьютерной лингвистике, кроме собственно академической карьеры в России или за рубежом включают в себя работу в качестве компьютерного лингвиста в коммерческих компаниях. Среди российских работодателей интерес к выпускникам программы проявляют такие «технологические гиганты», как компании Сбербанк, Яндекс, ABBYY, МТС, Mail.ru и др. В то же время, компьютерные лингвисты востребованы и в стартапах, небольших компаниях, чей бизнес связан с наукоемкими и информационными технологиями.

Перечень профессиональных компетенций:

ПК-1	Способен планировать научно-исследовательскую деятельность, проводить самостоятельные исследования и получать новые научные результаты в области профессиональной деятельности
ПК-10	Способен осуществлять письменный перевод различных типов текстов (в том числе художественных произведений) и сопровождать их необходимым редакторским комментарием
ПК-11	Способен проводить квалифицированное переводческое сопровождение международных мероприятий разного типа (в том числе научных) с использованием современных информационных технологий и программного обеспечения
ПК-12	Способен осуществлять редактирование научной и научно-популярной литературы по лингвистике и смежным областям, в том числе переводы с русского языка на иностранный и с иностранного на русский
ПК-13	Способен осуществлять лингвистическую обработку текстов в производственно-практических целях
ПК-14	Способен разрабатывать и совершенствовать электронные языковые ресурсы, лингвистические компоненты информационных систем, лингвистические модули компьютерных систем обучения
ПК-15	Способен проводить квалифицированный анализ различных типов устного и письменного дискурса в производственно-практических целях
ПК-16	Способен разрабатывать методики и алгоритмы дигитализации гуманитарного наследия, в том числе специальные языковые разметки корпусов
ПК-17	Способен разработать концепцию, осуществить подготовку, проведение и организационное сопровождение научного мероприятия, в том числе международного характера
ПК-18	Способен определять, транслировать общие цели в профессиональной и социальной деятельности
ПК-19	Способен порождать принципиально новые идеи и продукты, обладает креативностью, инициативностью

ПК-2	Способен проводить теоретические и практические занятия по дисциплинам предметного поля, создавать методические пособия и подготавливать учебно-методические и научно-методические материалы, в т.ч. междисциплинарного характера
ПК-20	Способен создавать, описывать и ответственно контролировать выполнение технологических требований и нормативов в профессиональной деятельности
ПК-3	Способен произвести углубленный анализ проблем, постановку и обоснование проектных задач, связанных с лингвистической проблематикой, моделированием и разработкой языковых компонент в междисциплинарных проектах
ПК-4	Способен самостоятельно разработать методику и инструментарий для осуществления исследовательской и проектной деятельности в области фундаментальной и прикладной лингвистики
ПК-5	Способен разрабатывать и управлять проектами (подпроектами), анализировать риски
ПК-6	Способен планировать и реализовать проекты в сфере межкультурной коммуникации, в том числе направленные на изучение языка как объекта культуры
ПК-7	Способен проводить анализ качества языковых данных, корпусов, систем, использующихся для автоматической обработки естественного языка
ПК-8	Способен обеспечивать аналитическую поддержку выработки и принятия решений в сферах, касающихся языковой политики, мониторинга текстовых данных, сохранения языкового наследия
ПК-9	Способен разрабатывать структуру и содержание образовательного процесса в области лингвистики, создавать учебные материалы и прикладные программы языкового обучения в системе общего и среднего профессионального образования, дополнительного образования

Характеристика образовательных модулей:

АДАПТАЦИОННЫЙ КУРС:

Студенты, не имеющие образования по направлению “Фундаментальная и прикладная лингвистика”, могут пролушать курс “Введение в лингвистику”

ОБЩАЯ БАЗОВАЯ ЧАСТЬ

В общей базовой части представлены две ключевых дисциплины, обеспечивающих высокий уровень лингвистической подготовки. Задача этих дисциплин состоит в том, чтобы познакомить студентов с современными подходами к анализу языковых структур и научить их пользоваться современными методами и инструментами для анализа данных. Эти задачи обеспечиваются курсами

- Формальные модели в естественном языке
- Функциональные и когнитивные модели в естественном языке

ДИСЦИПЛИНЫ СПЕЦИАЛИЗАЦИИ КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА

Пул обязательных дисциплин по компьютерной лингвистике включает в себя собственно курс по компьютерной лингвистике, а также тематически скординированные с ним курсы программирования на языке Python и математических оснований компьютерной лингвистики. На втором году в обязательный цикл входит курс Веб-девелопмента. Практические компетенции, полученные в рамках этого курса используются студентами для создания веб-приложений проектов, которые они разрабатывают на проектном семинаре.

Дисциплины вариативной части направлены на расширение практических компетенций и теоретических знаний студентов в области наук о данных (курсы Анализ лингвистических данных: квантовые методы и визуализации, Машинное обучение).

На втором году обучения студентам предлагаются дисциплины по выбору, предусматривающие углубленное погружение в одно из направлений компьютерной лингвистики. Это курс по нейросетям - введение в глубокое обучение, курс по анализу социальных сетей, курс по онтологиям и семантическим технологиям.

ПРАКТИКИ И НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА

Научно-исследовательская работа формируется из трех компонентов: проектно-исследовательского семинара, в рамках которого студенты выполняют проекты в группах, индивидуальной научно-исследовательской работы, в результате которой студент пишет и защищает курсовую работу, и мастер-классов, на которых студенты знакомятся с профессионалами и экспертами в изучаемых областях.

Студенческая практика организована на базе ведущих компаний и организаций, занимающихся компьютерной лингвистикой.

ИТОГОВАЯ ГОСУДАРСТВЕННАЯ АТТЕСТАЦИЯ

Итоговая государственная аттестация проходит в формате защиты дипломной работы с привлечением внешних оппонентов, в том числе (по возможности) из зарубежных вузов или институтов.

Теоретические основы компьютерной лингвистики

Вид: Исследовательская

Наставник: [Клышинский Эдуард Станиславович](#)

Язык реализации: Русский и английский

Применение электронного обучения и дистанционных технологий: С применением

Ключевые образовательные результаты:

KOP-1	Знает основные принципы работы с данными, владеет современными инструментами анализа данных на базовом уровне, в т.ч. навыками программирования, алгоритмизации и математическими методами при решении задач анализа данных
KOP-2	Знает историю и состояние компьютерной лингвистики, существующие подходы, направления, понимает связи с другими гуманитарными и естественно-научными дисциплинами, понимает возможности и ограничения методов компьютерной лингвистики
KOP-3	Умеет формулировать исследовательскую задачу в области компьютерной лингвистики, выбирает и применяет необходимые методы и инструменты ее решения
KOP-4	Владеет навыками проектного менеджмента применительно к проектам в области компьютерной лингвистики, способен работать в междисциплинарной проектной команде и организовать ее работу

КОР-5	Владеет методами и инструментами компьютерной лингвистики на уровне, достаточном для решения профессиональных задач
-------	---

Характеристика профессиональной деятельности:

Выпускник магистратуры по компьютерной лингвистике факультета гуманитарных наук НИУ ВШЭ:

- имеет хороший теоретический бэкграунд в теоретической лингвистике, математических методах обработки данных;
- знает, как устроена современная электронная языковая среда, уметь увидеть в ней новые проектные задачи; уметь их формулировать и формализовать;
- умеет работать с данными: понимать, какие именно данные нужны для решения поставленной задачи, как взаимодействуют лингвистические и иные данные, откуда их можно получить;
- понимает, какие языковые структуры лежат за той или иной локальной задачей, каковы лингвистические и экстралингвистические свойства текстов, используемых для решения задач;
- знает, какие лингвистические технологии следует применять в зависимости от условий задачи, понимать на какие лингвистические данные и ресурсы может опираться технология;
- умеет программировать прототипы и модели решений, знает алгоритмы машинного и глубокого обучения, умеет подготавливать необходимые лингвистические ресурсы;
- умеет проектировать цепочку процесса обработки языковых данных и интерпретировать результаты автоматической обработки лингвистических данных;
- умеет оценивать трудоемкость разных решений и пороги допустимых ошибок и выбирать оптимальную методологию тестирования.

Возможности трудоустройства выпускника специализации по компьютерной лингвистике, кроме собственно академической карьеры в России или за рубежом включают в себя работу в качестве компьютерного лингвиста в коммерческих компаниях. Среди российских работодателей интерес к выпускникам программы проявляют такие «технологические гиганты», как компании Сбербанк, Яндекс, АБВЬЮ, МТС, Mail.ru и др. В то же время, компьютерные лингвисты востребованы и в стартапах, небольших компаниях, чей бизнес связан с наукоемкими и информационными технологиями.

Перечень профессиональных компетенций:

ПК-1	Способен планировать научно-исследовательскую деятельность, проводить самостоятельные исследования и получать новые научные результаты в области профессиональной деятельности
ПК-10	Способен осуществлять письменный перевод различных типов текстов (в том числе художественных произведений) и сопровождать их необходимым редакторским комментарием
ПК-11	Способен проводить квалифицированное переводческое сопровождение международных мероприятий разного типа (в том числе научных) с использованием современных информационных технологий и программного обеспечения
ПК-12	Способен осуществлять редактирование научной и научно-популярной литературы по лингвистике и смежным областям, в том числе переводы с русского языка на иностранный и с иностранного на русский
ПК-13	Способен осуществлять лингвистическую обработку текстов в производственно-практических целях

ПК-14	Способен разрабатывать и совершенствовать электронные языковые ресурсы, лингвистические компоненты информационных систем, лингвистические модули компьютерных систем обучения
ПК-15	Способен проводить квалифицированный анализ различных типов устного и письменного дискурса в производственно-практических целях
ПК-16	Способен разрабатывать методики и алгоритмы дигитализации гуманитарного наследия, в том числе специальные языковые разметки корпусов
ПК-17	Способен разработать концепцию, осуществить подготовку, проведение и организационное сопровождение научного мероприятия, в том числе международного характера
ПК-18	Способен определять, транслировать общие цели в профессиональной и социальной деятельности
ПК-19	Способен порождать принципиально новые идеи и продукты, обладает креативностью, инициативностью
ПК-2	Способен проводить теоретические и практические занятия по дисциплинам предметного поля, создавать методические пособия и подготавливать учебно-методические и научно-методические материалы, в т.ч. междисциплинарного характера
ПК-20	Способен создавать, описывать и ответственно контролировать выполнение технологических требований и нормативов в профессиональной деятельности
ПК-3	Способен произвести углубленный анализ проблем, постановку и обоснование проектных задач, связанных с лингвистической проблематикой, моделированием и разработкой языковых компонент в междисциплинарных проектах
ПК-4	Способен самостоятельно разработать методику и инструментарий для осуществления исследовательской и проектной деятельности в области фундаментальной и прикладной лингвистики
ПК-5	Способен разрабатывать и управлять проектами (подпроектами), анализировать риски
ПК-6	Способен планировать и реализовать проекты в сфере межкультурной коммуникации, в том числе направленные на изучение языка как объекта культуры
ПК-7	Способен проводить анализ качества языковых данных, корпусов, систем, использующихся для автоматической обработки естественного языка
ПК-8	Способен обеспечивать аналитическую поддержку выработки и принятия решений в сферах, касающихся языковой политики, мониторинга текстовых данных, сохранения языкового наследия
ПК-9	Способен разрабатывать структуру и содержание образовательного процесса в области лингвистики, создавать учебные материалы и прикладные программы языкового обучения в системе общего и среднего профессионального образования, дополнительного образования

Характеристика образовательных модулей:

АДАПТАЦИОННЫЙ КУРС:

Студенты, не имеющие образования по направлению “Фундаментальная и прикладная лингвистика”, могут прослушать курс “Введение в лингвистику”

ОБЩАЯ БАЗОВАЯ ЧАСТЬ

В общей базовой части представлены две ключевых дисциплины, обеспечивающих высокий уровень лингвистической подготовки. Задача этих дисциплин состоит в том, чтобы познакомить студентов с современными подходами к анализу языковых структур и научить их пользоваться современными методами и инструментами для анализа данных. Эти задачи обеспечиваются курсами

- Формальные модели в естественном языке
- Функциональные и когнитивные модели в естественном языке

ДИСЦИПЛИНЫ СПЕЦИАЛИЗАЦИИ КОМПЬЮТЕРНАЯ ЛИНГВИСТИКА

Пул обязательных дисциплин по компьютерной лингвистике включает в себя собственно курс по компьютерной лингвистике, а также тематически скординированные с ним курсы программирования на языке Python и математических оснований компьютерной лингвистики. На втором году в обязательный цикл входит курс Веб-девелопмента. Практические компетенции, полученные в рамках этого курса используются студентами для создания веб-приложений проектов, которые они разрабатывают на проектном семинаре.

Дисциплины вариативной части направлены на расширение практических компетенций и теоретических знаний студентов в области наук о данных (курсы Анализ лингвистических данных: квантизативные методы и визуализации, Машинное обучение).

На втором году обучения студентам предлагаются дисциплины по выбору, предусматривающие углубленное погружение в одно из направлений компьютерной лингвистики. Это курс по нейросетям - введение в глубокое обучение, курс по анализу социальных сетей, курс по онтологиям и семантическим технологиям.

ПРАКТИКИ И НАУЧНО-ИССЛЕДОВАТЕЛЬСКАЯ РАБОТА

Научно-исследовательская работа формируется из трех компонентов: проектно-исследовательского семинара, в рамках которого студенты выполняют проекты в группах, индивидуальной научно-исследовательской работы, в результате которой студент пишет и защищает курсовую работу, и мастер-классов, на которых студенты знакомятся с профессионалами и экспертами в изучаемых областях.

Студенческая практика организована на базе ведущих компаний и организаций, занимающихся компьютерной лингвистикой.

ИТОГОВАЯ ГОСУДАРСТВЕННАЯ АТТЕСТАЦИЯ

Итоговая государственная аттестация проходит в формате защиты дипломной работы с привлечением внешних оппонентов, в том числе (по возможности) из зарубежных вузов или институтов.

Конкурентные преимущества программы

Актуальность магистерской программы по компьютерной лингвистике состоит в том, что она позволяет адаптировать и расширять профессиональные компетенции выпускников, традиционных лингвистических бакалавриатов, ориентированных в первую очередь на свободное владение иностранными языками, к потребностям современного информационного сообщества. Одновременно структура программы дает возможность профессиональной лингвистической подготовки для выпускников нелингвистических специальностей, в первую очередь тех, кто закончил бакалавриат или специалитет по прикладной математике или информатике.

Ориентированность программы на широкую аудиторию абитуриентов с разным бекграундом и ее гибкая структура дает возможность ей занять особую нишу относительно программ, не связанных с лингвистикой: программа расширяет компетенции студентов, дает возможность овладеть новой специальностью или же дополнить уже имеющиеся профессиональные навыки совершенно новыми умениями. Проекты программы ведут преподаватели НИУ ВШЭ, европейских и американских университетов, сотрудники ведущих российских компаний. Студенты проходят практику в компаниях, занимающихся разработками в области автоматической обработки естественного языка

Характеристика профессиональной деятельности и перечень профессиональных компетенций выпускника

Характеристики образовательных модулей программы

Адаптация программы для обучения лиц с ограниченными возможностями здоровья и инвалидов

Образовательная программа высшего образования НИУ ВШЭ адаптирована для обучения на ней инвалидов и лиц с ограниченными возможностями здоровья. В учебном процессе используются специальные технические средства обучения коллективного и индивидуального пользования для инвалидов и лиц с ограниченными возможностями здоровья. Особенности адаптации программ учебных дисциплин содержатся в полной версии каждой программы учебной дисциплины и доступны студентам через электронную образовательную среду.

Комплект документов образовательной программы

Все документы образовательной программы хранятся в электронном виде на настоящем сайте образовательной программы. Учебные планы, календарные учебные графики, программы учебных дисциплин разрабатываются и проходят электронные процедуры утверждения в корпоративных информационных системах. Их актуальные версии автоматически публикуются на сайте ОП. Методические материалы, оценочные средства и иные материалы образовательной программы в актуальном виде хранятся на сайте образовательной программы в соответствии с локальными нормативными актами университета.

Подтверждаю актуальность комплекта документов образовательной программы, размещенных на настоящем сайте образовательной программы.

Проректор С.Ю. Рощин

[Паспорт образовательной программы «Компьютерная лингвистика»](#)

