

## **Собеседование на программу компьютерная лингвистика**

### **Цель и основные задачи вступительного испытания**

Цель вступительного испытания – определить готовность и мотивацию абитуриента, поступающего в магистратуру освоить выбранную программу магистратуры, определить у поступающих базовый уровень подготовки в области компьютерной лингвистики.

Основные задачи вступительных испытаний:

- проверить уровень знаний претендента;
- определить склонность к научно-исследовательской деятельности;
- выяснить мотивы поступления в магистратуру;
- определить область научных интересов;
- определить уровень научной эрудиции претендента.

### **Формат вступительного испытания**

Вступительное испытание проводится в форме устной беседы и свободного обсуждения с абитуриентом материалов его профессионального портфолио.

Устная беседа проводится экзаменационной комиссией с каждым абитуриентом индивидуально. Абитуриенту задаются вопросы, которые позволяют оценить его профессиональный и личностный потенциал. На каждого абитуриента отводится не более 30 мин. Предварительно составляется список, в котором за каждым абитуриентом закрепляется время проведения собеседования. Список формируется с учетом пожеланий абитуриентов, публикуется в интернете и является доступным для всех участников собеседования. Собеседование проводится после объявления итогов конкурса портфолио согласно графику вступительных испытаний.

### **Перечень примерных вопросов**

1. **Блок вопросов на мотивацию** – рассказать о причинах поступления в магистратуру, об опыте участия в проектах или исследованиях, связанных с компьютерной лингвистикой, опыте прохождения он-лайн курсов, чтения специальной литературы
2. **Блок вопросов на эрудицию** – рассказать о лингвистических ресурсах и их использовании, о ключевых конференциях по компьютерной лингвистике, о лингвистических парсерах, о соревнованиях лингвистических систем
3. **Блок вопросов на подготовленность:**

вопрос типа кейс – выстроить цепочку лингвистических модулей для решения заданной задачи

*например: для задачи оценки тональности будут последовательно применены модули обработки текста (токенизация, лемматизация), модуль обращения к словарю тонально окрашенных слов и выражений, модуль с правилами поиска тональных слов и правилами вычисления итоговой тональности.*

вопрос типа рецензия – ознакомиться с текстом на профессиональную тему, изложить основные тезисы, дать собственную оценку.

#### **4. Блок вопросов на знакомство со современными задачами компьютерной лингвистики:**

Рассказать о последних достижениях в одной из областей компьютерной лингвистики (дается несколько областей на выбор), ресурсах, связанных с этой областью, основных технологиях которые применяются.

Примерный список тем:

Новостные агрегаторы

Чатботы,

Bionlp,

Удобочитаемость текста,

Анализ тональности отзывов

Машинный перевод

Электронные словари

Соревнования типа shared tasks в разных областях

Извлечение именованных сущностей

Морфологические анализаторы

Синтаксические анализаторы

Автоматическое разрешение кореференции

#### **5. Критерии оценок**

| Блок | Макс балл | Высокий уровень | Средний уровень | Слабый уровень | Неподготовлен |
|------|-----------|-----------------|-----------------|----------------|---------------|
| 1    | 10        | 8-10            | 5-7             | 2-5            | 0-2           |
| 2    | 20        | 15-20           | 10-15           | 5-10           | 0-5           |
| 3    | 40        | 30-40           | 20-30           | 10-20          | 0-10          |
| 4    | 30        | 22-30           | 15-22           | 7-15           | 0-7           |

## Описание уровней подготовленности абитуриента

**Высокий** уровень подготовленности абитуриента (абитуриент исчерпывающе ответил на все или почти все вопросы блока)

- **Средний** уровень подготовленности абитуриента ( абитуриент исчерпывающе ответил на 50-70% вопросов блока)
- **Слабый** уровень подготовленности абитуриента ( абитуриент исчерпывающе ответил на 30-50% вопросов блока)
- **Неподготовленность** абитуриента (абитуриент исчерпывающе ответил на менее, чем 30% вопросов блока)