

## **Приложение** *(обязательное)*

**Фонд оценочных средств по дисциплине " Математическая лингвистика и обработка текстов на естественном языке "**

ФОС оформлен в виде отдельного файла и является неотъемлемой и обязательной составляющей рабочей программы дисциплины.

### **1. Контроль и оценка результатов освоения учебной дисциплины**

Контроль и оценка результатов освоения учебной дисциплины осуществляются преподавателем в процессе проведения устных опросов, докладов и дискуссий, а также при проверке рефератов, выполненных аспирантами.

Для итоговой оценки усвоения дисциплины учебным планом предусмотрен зачет. Необходимым, но не достаточным, условием допуска к зачету является выполнение рефератов.

### **2. Требования к устным докладам**

К устным докладам аспирантов предъявляются следующие требования:

- содержание доклада должно соответствовать теме доклада;
- доклад должен иметь достаточную полноту и глубину охвата темы;
- выступление должно быть четким, аргументированным и выразительным.

### **3. Требования к оформлению и содержанию рефератов**

Реферат выполняется по выбранной аспирантом теме из «Перечня тем для рефератов». К реферату предъявляются следующие требования:

- содержание реферата должно соответствовать теме;
- объем реферата должен быть в пределах от 3 до 8 листов;
- реферат должен быть выполнен в печатном виде, рукописном виде или в виде вложения в формате "DOC" по e-mail;

- реферат в печатном оформлении должен иметь шрифт Times New Roman 12, междустрочный интервал 1,25;
- титульный лист для рефератов выполняется стандартным способом, т.е. должен содержать наименование, образовательного учреждения, темы реферата, Ф.И.О. исполнителя, Ф.И.О. преподавателя, год.

При оценке реферата учитываются степень соответствия содержания теме реферата, полнота охвата и глубина знания, четкость ответа, уровень изложения материала аспирантами.

#### **Перечень тем рефератов и докладов**

1. Проблемы автоматизации синтаксического анализа предложений.
2. Проблемы обнаружения кореференций и анафоров в текстах на ЕЯ.
3. Применение алгоритмов и методов обработки текстовой информации в технике.
4. Применение алгоритмов и методов обработки текстовой информации в медицине.
5. Применение алгоритмов и методов обработки текстовой информации в системах безопасности.
6. Возможности программных приложений для анализа социальных сетей.
7. Проблемы автоматической идентификации авторов текстов.
8. Методы автоматического построения онтологий.
9. Сравнение алгоритмов обнаружения и исправления ошибок и опечаток.
10. Сравнение алгоритмов морфологического анализа.

#### **4. Требования к зачету**

Для получения допуска к зачету аспиранту необходимо своевременно выполнить реферат. Если реферат не был выполнен вовремя, аспирант не допускается до зачета.

К зачетной работе предъявляются следующие требования:

- работа должна выполняться на базе пройденных тем письменно;
- работа должна быть выполнена в аудитории в течение 45 мин.;
- задания аспирантам должны быть произвольными в пределах списка вопросов по пройденным темам.

При оценке качества зачетной работы учитываются степень соответствия теме вопроса, полнота охвата и глубина знания, четкость ответа, уровень изложения материала аспирантами.

### **Перечень примерных вопросов для зачета**

1. Перечислить направления компьютерной лингвистики.
2. Сформулировать общие принципы построения автоматизированных систем обработки текстов.
3. Разъяснить принципы работы графематического и морфологического анализаторов.
4. Перечислить методы задания синтаксической структуры предложений.
5. Суть теоретико-модельного подхода к исследованию семантики текстов.
6. Изложить основные идеи теории речевых действий.
7. Привести классификацию речевых действий.
8. Привести примеры искусственных языков и нотаций.
9. Сформулировать принципы построения тезаурусов и онтологий. Сходства и отличия.
10. Дать определения семантических сетей, фреймов.
11. Что такое корпусная лингвистика?
12. Применение частотных методов в компьютерной лингвистике. Перечислить, описать, привести примеры.
13. В чем отличие между классификацией и кластеризацией текстов?
14. Перечислить методы классификации и кластеризации текстовой информации. Сформулировать основные принципы.
15. Разъяснить принципы работы автоматических систем извлечения информации.
16. Сформулировать принципы обработки неструктурированной и плохо структурированной информации. Индексация текстов.
17. Что понимается под сентимент-анализом текста?
18. Перечислить формальные методы атрибуции текстов.
19. Перечислить основные направления исследований социальных сетей.
20. Описать методы анализа социальных сетей.
21. Перечислить основные методы обнаружения спам-сообщений. Привести примеры.
22. Пояснить принцип работы наивного Байесовского классификатора.