

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное учреждение
высшего профессионального образования
"Казанский (Приволжский) федеральный университет"
Высшая школа русского языка и межкультурной коммуникации



УТВЕРЖДАЮ

Проректор
по образовательной деятельности КФУ
Проф. Таюрский Д.А.

"__" _____ 20__ г.

Программа дисциплины

Методы автоматической обработки естественного языка Б1.В.ОД.11

Направление подготовки: 44.04.01 - Педагогическое образование

Профиль подготовки: Информационные технологии в филологии и образовании

Квалификация выпускника: магистр

Форма обучения: заочное

Язык обучения: русский

Автор(ы):

Хакимов Б.Э.

Рецензент(ы):

Салехова Л.Л.

СОГЛАСОВАНО:

Заведующий(ая) кафедрой: Салехова Л. Л.

Протокол заседания кафедры No ____ от "____" _____ 201__ г

Учебно-методическая комиссия Института филологии и межкультурной коммуникации (Высшая школа русского языка и межкультурной коммуникации):

Протокол заседания УМК No ____ от "____" _____ 201__ г

Регистрационный No

Казань
2016

Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре основной образовательной программы
3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля
4. Структура и содержание дисциплины/ модуля
5. Образовательные технологии, включая интерактивные формы обучения
6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов
7. Литература
8. Интернет-ресурсы
9. Материально-техническое обеспечение дисциплины/модуля согласно утвержденному учебному плану

Программу дисциплины разработал(а)(и) доцент, к.н. Хакимов Б.Э. Кафедра образовательных технологий и информационных систем в филологии Высшая школа русского языка и межкультурной коммуникации, Boulat.Hakimov@kpfu.ru

1. Цели освоения дисциплины

Цель дисциплины "Методы автоматической обработки естественного языка" - сформировать у обучающихся систему компетенций, знаний, умений и навыков в сфере использования современных компьютерных технологий при работе с языковым материалом. Данная дисциплина призвана объяснить основные причины и условия применения компьютерных технологий в научных филологических исследованиях, прикладной, в том числе образовательной деятельности, познакомить студентов с существующими программными продуктами для осуществления профессиональной деятельности.

2. Место дисциплины в структуре основной образовательной программы высшего профессионального образования

Данная учебная дисциплина включена в раздел " Б1.В.ОД.11 Дисциплины (модули)" основной образовательной программы 44.04.01 Педагогическое образование и относится к обязательные дисциплины. Осваивается на 2 курсе, 4 семестр.

Данная учебная дисциплина логически связана в первую очередь с дисциплинами "Корпусно-ориентированный подход в лингвистике" и "Компьютерный перевод", а также с другими дисциплинами образовательной программы.

3. Компетенции обучающегося, формируемые в результате освоения дисциплины /модуля

В результате освоения дисциплины формируются следующие компетенции:

Шифр компетенции	Расшифровка приобретаемой компетенции
ОК-1 (общекультурные компетенции)	способность к абстрактному мышлению, анализу, синтезу, способностью совершенствовать и развивать свой интеллектуальный и общекультурный уровень
ПК-1 (профессиональные компетенции)	способность применять современные методики и технологии организации образовательной деятельности, диагностики и оценивания качества образовательного процесса по различным образовательным программам
ПК-21 (профессиональные компетенции)	способность формировать художественно-культурную среду
ПК-3 (профессиональные компетенции)	способность руководить исследовательской работой обучающихся
ПК-5 (профессиональные компетенции)	способность анализировать результаты научных исследований, применять их при решении конкретных научно-исследовательских задач в сфере науки и образования, самостоятельно осуществлять научное исследование
ПК-6 (профессиональные компетенции)	готовность использовать индивидуальные креативные способности для самостоятельного решения исследовательских задач

В результате освоения дисциплины студент:

применять полученные знания в научно-исследовательской и педагогической деятельности.

4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 3 зачетных(ые) единиц(ы) 108 часа(ов).

Форма промежуточного контроля дисциплины экзамен в 4 семестре.

Суммарно по дисциплине можно получить 100 баллов, из них текущая работа оценивается в 50 баллов, итоговая форма контроля - в 50 баллов. Минимальное количество для допуска к зачету 28 баллов.

86 баллов и более - "отлично" (отл.);

71-85 баллов - "хорошо" (хор.);

55-70 баллов - "удовлетворительно" (удов.);

54 балла и менее - "неудовлетворительно" (неуд.).

4.1 Структура и содержание аудиторной работы по дисциплине/ модулю

Тематический план дисциплины/модуля

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
1.	Тема 1. Введение. Теоретические, методологические и практические основы использования компьютерных технологий в сфере обработки естественного языка.	4	1-4	2	0	8	
2.	Тема 2. Особенности научно-исследовательской и прикладной деятельности в области автоматической обработки естественного языка.	4	5-9	4	0	10	
	Тема . Итоговая форма контроля	4		0	0	0	экзамен
	Итого			6	0	18	

4.2 Содержание дисциплины

Тема 1. Введение. Теоретические, методологические и практические основы использования компьютерных технологий в сфере обработки естественного языка.

лекционное занятие (2 часа(ов)):

Лингвистические компьютерные технологии. История компьютерной лингвистики.

Компьютерные методы лингвистических исследований. Автоматический анализ текста.

Лингвистические модели. Формализация языковой структуры. Лингвистические аспекты искусственного интеллекта.

лабораторная работа (8 часа(ов)):

Лабораторные работы по тематике раздела

Тема 2. Особенности научно-исследовательской и прикладной деятельности в области автоматической обработки естественного языка.

лекционное занятие (4 часа(ов)):

Компьютерная лексикография, электронные словари. Квантитативная лингвистика. Частотные словари. Корпусная лингвистика. Использование корпусов текстов в научных исследованиях. Лингвистические ресурсы и поиск в Интернет. Лингвистические технологии информационного поиска.

лабораторная работа (10 часа(ов)):

Лабораторные работы по тематике раздела

4.3 Структура и содержание самостоятельной работы дисциплины (модуля)

N	Раздел Дисциплины	Семестр	Неделя семестра	Виды самостоятельной работы студентов	Трудоемкость (в часах)	Формы контроля самостоятельной работы
1.	Тема 1. Введение. Теоретические, методологические и практические основы использования компьютерных технологий в сфере обработки естественного языка.	4	1-4	подготовка домашнего задания	30	проверка домашнего задания
2.	Тема 2. Особенности научно-исследовательской и прикладной деятельности в области автоматической обработки естественного языка.	4	5-9	подготовка домашнего задания	45	проверка домашнего задания
	Итого				75	

5. Образовательные технологии, включая интерактивные формы обучения

Освоение курса "Методы автоматической обработки естественного языка" предполагает использование как традиционных, так и инновационных образовательных технологий. Традиционные образовательные технологии подразумевают использование в учебном процессе таких методов работ, как лекция, семинар, практическое занятие и др. Инновационные образовательные технологии обуславливают внедрение в учебный процесс таких методов и приемов, как различные формы тренингов, деловые игры, дискуссия, моделирование ситуаций и др.

6. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов

Тема 1. Введение. Теоретические, методологические и практические основы использования компьютерных технологий в сфере обработки естественного языка.

проверка домашнего задания , примерные вопросы:

Домашняя работа по темам: Лингвистические компьютерные технологии. История компьютерной лингвистики. Компьютерные методы лингвистических исследований. Автоматический анализ текста. Лингвистические модели. Формализация языковой структуры. Лингвистические аспекты искусственного интеллекта.

Тема 2. Особенности научно-исследовательской и прикладной деятельности в области автоматической обработки естественного языка.

проверка домашнего задания , примерные вопросы:

Домашняя работа по темам: Компьютерная лексикография, электронные словари. Квантитативная лингвистика. Частотные словари. Корпусная лингвистика. Использование корпусов текстов в научных исследованиях. Лингвистические ресурсы и поиск в Интернет. Лингвистические технологии информационного поиска.

Тема . Итоговая форма контроля

Примерные вопросы к экзамену:

Вопросы к экзамену.

- 1) Лингвистические компьютерные технологии.
- 2) История компьютерной лингвистики.
- 3) Компьютерные методы лингвистических исследований.
- 4) Автоматический анализ текста.
- 5) Лингвистические модели.
- 6) Формализация языковой структуры.
- 7) Компьютерная лексикография, электронные словари.
- 8) Квантитативная лингвистика. Частотные словари.
- 9) Корпусная лингвистика.
- 10) Использование корпусов текстов в научных исследованиях.
- 11) Системы машинного перевода.
- 12) Локализация электронных ресурсов и программного обеспечения.
- 13) Лингвистические ресурсы и поиск в Интернет.
- 14) Лингвистические технологии информационного поиска.
- 15) Лингвистические аспекты искусственного интеллекта.

7.1. Основная литература:

Марчук Ю.Н. Компьютерная лингвистика. Учебное пособие. - М.: Восток-Запад, 2007. - 317 с. (1 экз.).

Захаров В.П., Богданова С.Ю. Корпусная лингвистика. - Иркутск: Издательство ИГЛУ, 2011. - 161 с. (1 экз.).

Зубов А.В. Информационные технологии в лингвистике: учебное пособие для студентов вузов. - М.: Academia, 2004. - 205 с. (26 экз.).

Хроленко А.Т., Денисов А.В. Современные информационные технологии для гуманитария: практическое руководство. - М.: Флинта: Наука, 2008. - 128 с. (1 экз.).

7.2. Дополнительная литература:

Гладкий А.В., Мельчук И.А. Элементы математической лингвистики. - М.: Наука, 1969. - 192 с. (3 экз.).

Кибрик А.Е. Очерки по общим и прикладным вопросам языкознания: (универсальное, типовое и специфическое в языке). - М.: Изд-во МГУ, 1992. - 335 с. (2 экз.).

7.3. Интернет-ресурсы:

Британский национальный корпус - <http://www.natcorp.ox.ac.uk>

Компания 'Аби' - <http://abbyy.ru>

Национальный корпус русского языка - <http://www.ruscorpora.ru>

НИИ 'Прикладная семиотика' АН РТ - <http://ips.antat.ru>

Филологический факультет МГУ - <http://www.philol.msu.ru/~lex/main.htm>

8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Методы автоматической обработки естественного языка" предполагает использование следующего материально-технического обеспечения:

Мультимедийная аудитория, вместимостью более 60 человек. Мультимедийная аудитория состоит из интегрированных инженерных систем с единой системой управления, оснащенная современными средствами воспроизведения и визуализации любой видео и аудио информации, получения и передачи электронных документов. Типовая комплектация мультимедийной аудитории состоит из: мультимедийного проектора, автоматизированного проекционного экрана, акустической системы, а также интерактивной трибуны преподавателя, включающей тач-скрин монитор с диагональю не менее 22 дюймов, персональный компьютер (с техническими характеристиками не ниже Intel Core i3-2100, DDR3 4096Mb, 500Gb), конференц-микрофон, беспроводной микрофон, блок управления оборудованием, интерфейсы подключения: USB, audio, HDMI. Интерактивная трибуна преподавателя является ключевым элементом управления, объединяющим все устройства в единую систему, и служит полноценным рабочим местом преподавателя. Преподаватель имеет возможность легко управлять всей системой, не отходя от трибуны, что позволяет проводить лекции, практические занятия, презентации, вебинары, конференции и другие виды аудиторной нагрузки обучающихся в удобной и доступной для них форме с применением современных интерактивных средств обучения, в том числе с использованием в процессе обучения всех корпоративных ресурсов. Мультимедийная аудитория также оснащена широкополосным доступом в сеть интернет. Компьютерное оборудование имеет соответствующее лицензионное программное обеспечение.

Компьютерный класс, представляющий собой рабочее место преподавателя и не менее 15 рабочих мест студентов, включающих компьютерный стол, стул, персональный компьютер, лицензионное программное обеспечение. Каждый компьютер имеет широкополосный доступ в сеть Интернет. Все компьютеры подключены к корпоративной компьютерной сети КФУ и находятся в едином домене.

Учебно-методическая литература для данной дисциплины имеется в наличии в электронно-библиотечной системе "ZNANIUM.COM", доступ к которой предоставлен студентам. ЭБС "ZNANIUM.COM" содержит произведения крупнейших российских учёных, руководителей государственных органов, преподавателей ведущих вузов страны, высококвалифицированных специалистов в различных сферах бизнеса. Фонд библиотеки сформирован с учетом всех изменений образовательных стандартов и включает учебники, учебные пособия, УМК, монографии, авторефераты, диссертации, энциклопедии, словари и справочники, законодательно-нормативные документы, специальные периодические издания и издания, выпускаемые издательствами вузов. В настоящее время ЭБС ZNANIUM.COM соответствует всем требованиям федеральных государственных образовательных стандартов высшего профессионального образования (ФГОС ВПО) нового поколения.

Для изучения данной дисциплины необходима компьютерный класс, оборудованный мультимедийными компьютерами с доступом в Интернет, проектор, экран, интерактивная доска, принтер, сканер, копир.

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 44.04.01 "Педагогическое образование" и магистерской программе Информационные технологии в филологии и образовании .

Автор(ы):

Хакимов Б.Э. _____

"__" _____ 201__ г.

Рецензент(ы):

Салехова Л.Л. _____

"__" _____ 201__ г.