

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное учреждение
высшего профессионального образования
"Казанский (Приволжский) федеральный университет"
Высшая школа татаристики и тюркологии им.Габдуллы Тукая



УТВЕРЖДАЮ

Проректор по образовательной деятельности КФУ

Проф. Талюцкий Д.А.

КАЗАНСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
ДЕПАРТАМЕНТ
ОБРАЗОВАНИЯ
(ДО КФУ)

_____ 20__ г.

подписано электронно-цифровой подписью

Программа дисциплины

Актуальные проблемы компьютерной лингвистики Б1.В.ДВ.7

Направление подготовки: 45.03.01 - Филология

Профиль подготовки: Прикладная филология: татарский язык и литература, информационные технологии с углубленным изучением иностранного языка

Квалификация выпускника: бакалавр

Форма обучения: очное

Язык обучения: русский

Автор(ы):

Хакимов Б.Э.

Рецензент(ы):

Данилов А.В.

СОГЛАСОВАНО:

Заведующий(ая) кафедрой: Салехова Л. Л.

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

Учебно-методическая комиссия Института филологии и межкультурной коммуникации (Высшая школа татаристики и тюркологии им.Габдуллы Тукая):

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

Регистрационный No 902332618

Казань

2018

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

Данная дисциплина ставит своей целью знакомство студентов с актуальными проблемами современной компьютерной лингвистики.

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

Данная учебная дисциплина включена в раздел " Б1.В.ДВ.7 Дисциплины (модули)" основной образовательной программы 45.03.01 Филология и относится к дисциплинам по выбору. Осваивается на 3 курсе, 6 семестр.

Дисциплина содержательно и логически связана с другими дисциплинами данного направления, связанными с информационными технологиями.

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)	владение культурой мышления; способность к восприятию, анализу, обобщению информации, постановке цели и выбору путей ее достижения
ОК-9 (общекультурные компетенции)	умение использовать основные положения и методы социальных, гуманитарных и экономических наук в профессиональной деятельности; способность анализировать социально значимые проблемы и процессы
ПК-1 (профессиональные компетенции)	способность демонстрировать знание основных положений и концепций в области теории и истории основного изучаемого языка (языков) и литературы (литератур), теории коммуникации, филологического анализа и интерпретации текста, представление об истории, современном состоянии и перспективах развития филологии
ПК-14 (профессиональные компетенции)	владение навыками перевода различных типов текстов (в основном научных и публицистических, а также документов) с иностранных языков и на иностранные языки; аннотирование и реферирование документов, научных трудов и художественных произведений на иностранных языках
ПК-2 (профессиональные компетенции)	владение базовыми навыками сбора и анализа языковых и литературных фактов с использованием традиционных методов и современных информационных технологий
ПК-5 (профессиональные компетенции)	способность применять полученные знания в области теории и истории основного изучаемого языка (языков) и литературы (литератур), теории коммуникации, филологического анализа и интерпретации текста в собственной научно-исследовательской деятельности

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-7 (профессиональные компетенции)	владение навыками подготовки научных обзоров, аннотаций, составления рефератов и библиографий по тематике проводимых исследований, приемами библиографического описания; знание основных библиографических источников и поисковых систем
ПК-8 (профессиональные компетенции)	владение навыками участия в научных дискуссиях, выступления с сообщениями и докладами, устного, письменного и виртуального (размещение в информационных сетях) представления материалов собственных исследований
ОК-11 (общекультурные компетенции)	владение основными методами, способами и средствами получения, хранения, переработки информации, навыки работы с компьютером как средством управления информацией
ОК-12 (общекультурные компетенции)	способность работать с информацией в глобальных компьютерных сетях

В результате освоения дисциплины студент:

4. должен демонстрировать способность и готовность:

применять полученные знания в научно-исследовательской и прикладной деятельности в области теории и истории основного изучаемого языка и литературы, а также в устной, письменной и электронной коммуникации.

4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 2 зачетных(ые) единиц(ы) 72 часа(ов).

Форма промежуточного контроля дисциплины зачет в 6 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. Понятие семантических технологий. Возможности и трудности						

автоматической обработки семантической информации.

6

1-2

2

4

0

Эссе

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
2.	Тема 2. Онтологии и тезаурусы. Технология Wordnet.	6	3-5	4	6	0	Тестирование
3.	Тема 3. Семантическая разметка в корпусах текстов	6	6-9	4	8	0	Тестирование
4.	Тема 4. Технологии Semantic Web. Информационный поиск.	6	10-12	2	6	0	Эссе
	Тема . Итоговая форма контроля	6		0	0	0	Зачет
	Итого			12	24	0	

4.2 Содержание дисциплины

Тема 1. Введение. Понятие семантических технологий. Возможности и трудности автоматической обработки семантической информации.

лекционное занятие (2 часа(ов)):

Введение. Понятие о семантических технологиях. Автоматическая обработка семантической информации. Потенциальные возможности учета семантической информации в системах автоматической обработки текстов. Сложности в автоматизации анализа семантики как следствие особенностей семантического уровня языка.

практическое занятие (4 часа(ов)):

Выполнение заданий по теме соответствующей лекции

Тема 2. Онтологии и тезаурусы. Технология Wordnet.

лекционное занятие (4 часа(ов)):

Онтологии в лингвистических исследованиях и прикладной деятельности. Тезаурусы и их компьютерные версии. Особенности тезаурусов типа Wordnet. Создание тезауруса Wordnet для русского и татарского языков. Использование тезаурусов в задачах информационного поиска и других задачах.

практическое занятие (6 часа(ов)):

Выполнение заданий по теме соответствующей лекции

Тема 3. Семантическая разметка в корпусах текстов

лекционное занятие (4 часа(ов)):

Семантическая разметка как одна из разновидностей корпусной лингвистической разметки. Таксономическая разметка, тематическая разметка. Разметка корпуса с помощью семантического словаря. Взаимосвязь различных видов лингвистической разметки.

практическое занятие (8 часа(ов)):

Выполнение заданий по теме соответствующих лекций

Тема 4. Технологии Semantic Web. Информационный поиск.

лекционное занятие (2 часа(ов)):

Особенности технологий Semantic Web. Применение технологий Semantic Web. Обработка семантической информации в процессе информационного поиска. Развитие технологий информационного поиска на татарском языке.

практическое занятие (6 часа(ов)):

Выполнение заданий по теме соответствующей лекции

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. Понятие семантических технологий. Возможности и трудности автоматической обработки семантической информации.	6	1-2	подготовка к эссе	8	Эссе
2.	Тема 2. Онтологии и тезаурусы. Технология Wordnet.	6	3-5	подготовка к тестированию	10	Тестирование
3.	Тема 3. Семантическая разметка в корпусах текстов	6	6-9	подготовка к тестированию	10	Тестирование
4.	Тема 4. Технологии Semantic Web. Информационный поиск.	6	10-12	подготовка к эссе	8	Эссе
	Итого				36	

5. Образовательные технологии, включая интерактивные формы обучения

Освоение курса 'Актуальные проблемы компьютерной лингвистики' 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. Понятие семантических технологий. Возможности и трудности автоматической обработки семантической информации.

Эссе , примерные вопросы:

Примерные темы эссе: Понятие о семантических технологиях. Автоматическая обработка семантической информации. Потенциальные возможности учета семантической информации в системах автоматической обработки текстов. Сложности в автоматизации анализа семантики как следствие особенностей семантического уровня языка.

Тема 2. Онтологии и тезаурусы. Технология Wordnet.

Тестирование , примерные вопросы:

Тематика тестирования: Онтологии в лингвистических исследованиях и прикладной деятельности. Тезаурусы и их компьютерные версии. Особенности тезаурусов типа Wordnet. Создание тезауруса Wordnet для русского и татарского языков. Использование тезаурусов в задачах информационного поиска и других задачах.

Тема 3. Семантическая разметка в корпусах текстов

Тестирование , примерные вопросы:

Тематика тестирования: Семантическая разметка как одна из разновидностей корпусной лингвистической разметки. Таксономическая разметка, тематическая разметка. Разметка корпуса с помощью семантического словаря. Взаимосвязь различных видов лингвистической разметки.

Тема 4. Технологии Semantic Web. Информационный поиск.

Эссе , примерные вопросы:

Примерные темы эссе: Особенности технологий Semantic Web. Применение технологий Semantic Web. Обработка семантической информации в процессе информационного поиска. Развитие технологий информационного поиска на татарском языке.

Тема . Итоговая форма контроля

Примерные вопросы к зачету:

Примерные вопросы к зачету:

- 1) Понятие о семантических технологиях
- 2) Автоматическая обработка семантической информации.
- 3) Возможности учета семантической информации в системах автоматической обработки текстов
- 4) Сложности в автоматизации анализа семантики.
- 5) Онтологии
- 6) Применение онтологий в лингвистических исследованиях и прикладной деятельности.
- 7) Тезаурусы
- 8) Создание компьютерных электронных тезаурусов.
- 9) Тезаурус типа Wordnet.
- 10) Wordnet на разных языках.
- 11) Разработка Wordnet для русского и татарского языков.
- 12) Тезаурусы в задачах информационного поиска.
- 13) Семантическая разметка в корпусе текстов.
- 14) Таксономическая разметка
- 15) Тематическая и другие виды разметки.
- 16) Разметка корпуса с помощью семантического словаря.
- 17) Взаимосвязь различных видов лингвистической разметки.
- 18) Технологии Semantic Web
- 19) Обработка семантической информации в процессе информационного поиска.
- 20) Развитие технологий информационного поиска на татарском языке.

7.1. Основная литература:

Федотова Е.Л., Федотов А.А. Информационные технологии в науке и образовании: Учебное пособие / Е.Л. Федотова, А.А. Федотов. - М.: ИД ФОРУМ: НИЦ ИНФРА-М, 2015. - 336 с.
<http://znanium.com/catalog.php?bookinfo=487293>

Потапова Р.К. Новые информационные технологии и лингвистика : учебное пособие для студентов вузов, обучающихся по специальности 021800 'Теоретическая и прикладная лингвистика' направления 620200 'Лингвистика и новые информационные технологии' / Р.К. Потапова ; Моск. гос. лингвист. ун-т .? Изд. 5-е .? Москва : URSS : [ЛИБРОКОМ, 2012] .? 364 с.

Щипицина, Л. Ю. Информационные технологии в лингвистике [Электронный ресурс] : учеб. пособие / Л. Ю. Щипицина. ? М. : ФЛИНТА, 2013. ? 128 с.
<http://znanium.com/catalog.php?bookinfo=462989>

Грудева, Е. В. Корпусная лингвистика [Электронный ресурс] : учеб. пособие / Е. В. Грудева. - 2-е изд., стер. - М. : ФЛИНТА, 2012. - 165 с. <http://znanium.com/catalog.php?bookinfo=455049>

7.2. Дополнительная литература:

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

Хроленко А.Т., Денисов А.В. Современные информационные технологии для гуманитария: практическое руководство. - М.: Флинта: Наука, 2008. - 128 с.
<http://znanium.com/catalog.php?bookinfo=320764>

7.3. Интернет-ресурсы:

WordNet - <https://wordnet.princeton.edu/>

Искусственный интеллект и экспертные системы - http://expro.ksu.ru/materials/ii_i_es/book.html

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

Российская ассоциация искусственного интеллекта - <http://www.raai.org/resurs/resurs.shtml>

Татарский национальный корпус ?Туган тел? - <http://corpus.antat.ru>

8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Актуальные проблемы компьютерной лингвистики" 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.

Освоение дисциплины "Актуальные проблемы компьютерной лингвистики" предполагает использование следующего материально-технического обеспечения:

Мультимедийная аудитория, состоящая 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.

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 45.03.01 "Филология" и профилю подготовки Прикладная филология: татарский язык и литература, информационные технологии с углубленным изучением иностранного языка

Автор(ы):

Хакимов Б.Э. _____

"__" _____ 201__ г.

Рецензент(ы):

Данилов А.В. _____

"__" _____ 201__ г.