ТИПОВАЯ ПРОГРАММА
по дисциплине «ВВЕДЕНИЕ В КОМПЬЮТЕРНУЮ ЛИНГВИСТИКУ»

для направления подготовки 531200 «Компьютерная лингвистика»
Академическая степень: бакалавр

Составитель: **Мадраимова Ж. И.** – и. о. доц. кафедры Информационных технологий факультета Лингвистики

СТРУКТУРА ПРОГРАММЫ

I. Организационно-методический раздел
I. 1. Цели и задачи дисциплины
I. 2. Место дисциплины в структуре ООП высшего профессионального образования
I. 3. Требования к результатом освоения содержания дисциплины
I. 4. Объем дисциплины и виды учебной работы
I. 5. Формы контроля
I. 6. Технологическая карта дисциплины
II. Содержание программы учебной дисциплины
II. 1 Содержание разделов дисциплины для 1-курса
II. 2 Вопросы и задания для самостоятельной работы студентов
III. Оценочные средства по итогам освоения дисциплины
III. 1 Примерные темы проектов
III. 2 Примерные вопросы к экзамену
IV Образовательные технологии
IV. 1 Учебно-методические: основная и дополнительная литература
IV. 2 Информационное и материально-техническое обеспечение дисциплины: база-данных, информационно-справочные и поисковые системы
IV. 3. Материально-техническое обеспечение дисциплины

I. ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЙ РАЗДЕЛ

I.1. Цель дисциплины – сформировать у студентов системное представление о лингвистических основах информатики и аспектах моделирования языка и мышления в компьютерной среде с помощью компьютерных программ; сформировать систему знаний, умений и навыков в области использования компьютерных технологий как в обучении языку, так и в научно-исследовательской деятельности будущих бакалавров.

Задачи дисциплины:

- сформировать компетентности в области использования возможностей современных компьютерных средств в образовательной деятельности;
- обучить студентов использованию и применению средств ИКТ в профессиональной деятельности специалиста, работающего в системе образования;
- формировать умения работать с информацией, развитие коммуникативных способностей;

Пререквизиты – это дисциплины, содержащие знания, умения и навыки, необходимые для освоения изучаемой дисциплины, соответственно до изучения данной дисциплины определяют ее преемственность. Пререквизитами данного курса являются знания студентов по информатике и информационным технологиям.

Постреквизиты – это дисциплины, для изучения которых требуются знания, умения и навыки, приобретаемые по завершении 1-го курса, т. е. это обязательные для освоения дисциплины, соответственно после изучения данного предмета. Постреквизитами данного курса являются практическое применение автоматических словарей, осуществлять запросы и извлекать информацию.

I. 2. Место дисциплины в структуре ООП бакалавра

Данная учебная дисциплина входит в базовую часть цикла общенаучных дисциплин ГОСТ по направлению подготовки 531200 «Компьютерная Лингвистика».

Дисциплина «Введение в компьютерную лингвистику» по своему содержанию логически тесно связана с дисциплиной «Информатика», «Новые информационные технологии в учебном процессе», «Лексикография и фразеография», изучаемые в ООП бакалавра компьютерной лингвистики.

По дисциплине «Введение в компьютерную лингвистику» бакалавр направления 531200 «Компьютерная Лингвистика» готовится к своей профессии в области, во-первых, научно-методической деятель-

ности, для чего необходимы формируемые профессиональные компетенции, во- вторых, научно-исследовательской деятельности, при современном уровне развития науки и техники проведение исследовательской работы невозможно без элементарных знаний, умений и навыков, получаемых в процессе усвоения дисциплины. Особое внимание уделяется разработке методов анализа, использованию уже достигнутого в науке, ознакомлению с требованиями, предъявляемыми к научному изложению и доказательности.

I.3. Требования к результатам освоения содержания дисциплины.

В соответствии с требованиями ГОСТ в результате освоения дисциплин обучающийся должен овладеть комплексом компетенций. Выполнение этого требования проверяется при аттестации образовательной программы, в том числе путём контроля остаточных знаний обучающихся.

Дисциплина «Введение в компьютерную лингвистику» способствует формированию студентов следующих компетенций, предусматриваемых ГОСТ:

1) общенаучными (ОК):

- владение культурой мышления, способность к обобщению, анализу, восприятия информации, постановке цели и выбору путей её достижения, умение логически аргументировано и ясно строить устную и письменную речь (ОК-1);
- понимание социальной значимости своей будущей профессии, обладание высоко мотивацией к выполнению профессиональной деятельности (ОК- 3);
- умение применять методы и средства познания, обучения и самоконтроля для интеллектуального развития, повышения культурного уровня, профессионально компетенции, сохранения своего здоровья, нравственного и физическо-самосовершенствования (ОК-5);
- владение широкой общей подготовкой (базовыми знаниями) для решения практических задач в области информационных систем и технологий (ОК-6);

инструментальными (ИК):

- владение иностранными языками в устной и письменной форме для осуществления коммуникации в учебной, научной, профессиональной и социально-культурной сферах общения; владение терминологией специальности на иностранном языке; умение готовить публикации, проводить презентации, вести дискуссии и защищать представленную работу на иностранном языке (ИК-4);
- владение навыками использования программных средств и работы в

компьютерных сетях, использования ресурсов Интернет; владение основными методами, способами и средствами получения, хранения, переработки информации (ИК-5);

- способность использовать современную вычислительную технику и специализированное программное обеспечение в научно-исследовательской работе (ИК-6);

социально-личностными и общекультурными (СЛК);

- способность к письменной, устной и электронной коммуникации на государственном и/или русском языках и необходимое знание иностранного языка (профессиональный английский или др. язык) (СЛК-1);
- умение критически оценивать свои достоинства и недостатки, наметить пути и выбрать средства развития достоинств и устранения недостатков (СЛК-3);
- готов к самооценке, ценностному социо-культурному самоопределению и саморазвитию (СЛК-8);

профессиональными (ПК);

Выпускник должен обладать следующими профессиональными компетенциями (ПК):

- способность разрабатывать алгоритмы частичного синтаксического анализа текстов, написанных на естественном языке (ПК-1);
- способность применять для решения прикладных задач принципов установления лексических отношений, а также свойств и моделей связного текста (ПК-3);
- способность создавать системы автоматической рубрикации текстов, а также информационно-поисковые системы, основанные на разных моделях поиска (ПК-8);
- способность применять современные методы автоматической обработки текстов, статистики, машинного обучения для извлечения знаний из текстов (ПК-10);
- способность работать с информацией в глобальных компьютерных сетях (ПК-14);
- готовностью участвовать в рамках профессиональной компетенции в разработке электронных словарей и других электронных ресурсов для решения лингвистических задач (ПК-23);
- способность разрабатывать лингвистические компоненты систем автоматической обработки естественного языка (синтеза и распознавания устной речи, генерации текста, контент-анализа, автоматического перевода, автоматического реферирования и аннотирования), а также интеллектуальных систем (вопросно- ответных, экспертных) (ПК-28);

- способность собирать и анализировать научно-техническую информацию, учитывать современные тенденции развития и использовать достижения отечественной и зарубежной науки, техники и технологии в профессиональной деятельности (ПК-44);

В результате освоения компетенций студент должен:

знать:

- методы и решения в системах организации знаний; эмпирические законы описания естественных языков; логико-статистические методы анализа языкового описания предметных областей; технологию автоматизированной обработки текстовой информации.

уметь:

- использовать технологию автоматизированной обработки текстовой информации для анализа языковых описаний предметных областей.
- иметь навыки: работы со специальными программными средствами автоматизированной обработки текстов.
- иметь представление: о системе языковедческих дисциплин; предмете, методах и моделях прикладной лингвистики; предмете компьютерной лингвистики.
- знать приемы и методы использования компьютерных технологий в различных видах и формах учебной и научной деятельности;

владеть:

- методикой использования компьютерных технологий в предметной области;
- обладать навыками разработки педагогических технологий, создания инновационных программных и учебных продуктов по лингвистике, основанных на применении компьютера;
- уметь использовать телекоммуникационные технологии в образовательных целях при преподавании филологических дисциплин.

I. 4. Объем дисциплины и виды учебной работы для 1-курса

Виды учебной работы	**Всего**		**Семестры**	
	Часов	**кредиты**	**1**	**2**
Аудиторные занятия (всего)	**120**	**4**	**2**	**2**
В том числе				
лекции	32		16	16
Практические занятия	28		14	14
Самостоятельная работа	60		30	30
Итоговый контроль			**Экзамен**	**Экзамен**
Общая трудоемкость часов	**120**	4		

I.5. Формы контроля: зачет

Текущий контроль осуществляется лектором и преподавателем, ведущим практические занятия.

Контроль за освоением дисциплины осуществляется в каждом дисциплинарном разделе отдельно.

I. 6. Технологическая карта дисциплины

Наименование дисциплины	**Уровень/ступень образования (бакалавриат/ магистратура)**	**Количество зачетных единиц/ кредитов**	**Форма отчетности**	**Курс, семестр**
Введение в компьютерную лингвистику	бакалавриат	4	Экзамен	1 курс 1, 2 семестр

БАЗОВЫЙ МОДУЛЬ (проверка знаний и умений по дисциплине)				
Тема или задание текущей аттестационной работы	**Виды текущей аттестации**	**Аудиторная или внеаудиторная**	**Минимальное количество баллов**	**Максимальное количество баллов**
Выполнение заданий в течение учебного процесса	Своевременное выполнения домашних и проектных работ	Внеаудиторная или аудиторная	30	40
Выполнение заданий самостоятельную работу(Написание реферата, практических работ по задан. теме)	Выполнение определенного задания по СРС	внеаудиторная	5	15
Знание базовых понятий и примеров по основным темам и умение решать практические задачи	Устный опрос, тестирование	Аудиторная и внеаудиторная	25	30
Итого:			**60**	**85**

ДОПОЛНИТЕЛЬНЫЙ МОДУЛЬ				
Тема или задание текущей аттестационной работы	**Виды текущей аттестации**	**Аудиторная и внеаудиторная**	**Минимальное количество баллов**	**Максимальное количество баллов**
Презентация, практикум	Устный ответ	Аудиторная	4	10
Посещение занятий	Присутствие	Аудиторная	1	5
Итого:			5	15
Итого всего:			65	100

Все оценки складываются из результатов накопленных баллов по каждому модулю и для финального контроля выводится средний балл для окончательной оценки за курс:

Шкала перевода баллов в оценки следующее:

1 модуль -70 баллов	СРС-15 баллов
0-51 «неуд»	0-3 «неуд»
52-62 «удовл»	4-8 «удовл»
63-74 «хор»	9-11 «хор»
75-85 «отл»	12-15 «отл»

Проверка знаний: 2 контрольные работы (или тестирование) по основным темам и 2 проектные работы.

Дополнительные требования для студентов, отсутствующих на занятиях по уважительной причине: устное сквозное собеседование с преподавателем по проблеме пропущенных практических занятий, обязательное выполнение аудиторных проверочных и письменных работ.

Итоговая форма контроля – экзамен.

Формы контроля: Экзамен

Текущий контроль осуществляется лектором и преподавателем, ведущим практические занятия.

Контроль за освоением дисциплины осуществляется в каждом дисциплинарном разделе отдельно.

II. СОДЕРЖАНИЕ ПРОГРАММЫ УЧЕБНОЙ ДИСЦИПЛИНЫ

II.1. Содержания разделов дисциплины

Тема 1. Основные направления и предмет компьютерной лингвистики. Когнитивный инструментарий компьютерной лингвистики.

Тема 2. Компьютерная лингвистика как междисциплинарное научное направление. История становления и основные тенденции развития. Когнитивный инструментарий компьютерной лингвистики. Предмет компьютерной лингвистики.

Тема 3. Направления КЛ: автоматизированный анализ, распознание и автоматическая обработка текста; искусственный интеллект и машинный перевод;

Тема 4. Компьютерная лексикография (электронные словари); корпусная лингвистика и создание электронной базы данных; Исходные понятия корпусной лингвистики.

Тема 5. Компьютерная лингводидактика. Введение. История развития компьютерной лингводидактики. Условия эффективной интеграции информационно-коммуникационных технологий (ИКТ) в учебный процесс. Программное обеспечение обучения. Программное обеспечение обучения языку. Типология электронных материалов для обучения языку. Оценка качества электронных учебных материалов.

Тема 6. Естественные и искусственные языки. Компьютерное обеспечение представления знаний. Естественные и искусственные языки. Автоматизированный анализ: распознавание и синтез устной и письменной речи. Проблемы и перспективы, морфологический анализ, проблемы семантического анализа, синтаксический анализ, реферирование и аннотирование текста.

Тема 7. Лингвистические базы данных: модели и типы данных. Создания общих искусственных языков для представления информации; усовершенствования языков доступа к информации

Тема 8. Современный машинный перевод. Предпосылки возникновения и развитие машинного перевода (МП). Последовательность формальных операций, обеспечивающих анализ и синтез в системе машинного перевода. Стратегии машинного перевода. Автоматизированные и полуавтоматизированные системы перевода. Области использования машинного перевода. Важнейшие системы машинного перевода.

Тема 9. Компьютерная лексикография и корпусная лингвистика. Типология традиционных и электронных словарей. Словарные процессоры. Типология словарей. Тезаурусы и терминологические словари. Компьютерные технологии составления и эксплуатации словарей. Специальные программы – базы данных, компьютерные картотеки, программы обработки текста. Корпусная лингвистика, ориентированная на создание национальных корпусов естественных языков. Типы лингвистической разметки.

Тема 10. Информационные технологии в научной деятельности (этапы конструирования логики научного исследования). Поисковые

компьютерные программы и их использование в научных проектах. Использование компьютера при частотных статистических анализах текстов на материалах разных языков. Накопление знаний и фактов, поиск литературы, опытно-экспериментальная работа при фиксации лингвистических данных.

Тема 11. Компьютерная лингводидактика. Дидактическая концепция обучения на основе компьютерных технологий. Принципы создания электронных учебных средств. Классификация электронных средств учебного назначения. Электронные учебники, словари, учебно-методические материалы. Мультимедиа в помощь лингвисту.

Примерный план практических занятий

1. Учебные компьютерные словари
2. Автоматизация подготовки и редактирования текстов
3. Базы данных, компьютерные картотеки
4. Информационный поиск: реферирование, аннотирование
5. Информационный поиск. Поиск в сети Интернет
6. Машинный перевод
7. Электронные учебники, словари
8. Работа с презентациями

II.2. Вопросы и задания для самостоятельной работы студентов

1. Автоматизированный анализ: распознавание и синтез устной и письменной речи.
2. Морфологический анализ, проблемы семантического анализа, синтаксический анализ.
3. Реферирование и аннотирование текста.
4. Лингвистические базы данных: модели и типы данных. Создания общих искусственных языков для представления информации.
5. Современный машинный перевод. Предпосылки возникновения и развитие машинного перевода (МП).
6. Последовательность формальных операций, обеспечивающих анализ и синтез в системе машинного перевода.
7. Области использования машинного перевода. Важнейшие системы машинного перевода.
8. Компьютерная лексикография как одно из направлений прикладной лингвистики. Словарные процессоры.
9. Основные понятия структуры словаря: словник, словарная статья, грамматические, стилистические пометы; иллюстративный материал
10. Типология электронных словарей.
11. Тезаурусы и терминологические словари.

III. ОЦЕНОЧНЫЕ СРЕДСТВА ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

III.1. Примерные вопросы к экзамену

1. Цели и задачи курса «Введение в компьютерную лингвистику». Компьютерная лингвистика как междисциплинарное научное направление. Предмет компьютерной лингвистики.
2. История становления и основные тенденции развития.
3. Основные направления компьютерной лингвистики.
4. Когнитивный инструментарий компьютерной лингвистики. «Фреймы», «сценарии» и «планы».
5. Компьютерное обеспечение представления знаний.
6. Естественные и искусственные языки. Виды искусственных языков.
7. Автоматизированный анализ: распознавание и синтез устной и письменной речи.
8. Морфологический анализ, проблемы семантического анализа, синтаксический анализ.
9. Реферирование и аннотирование текста.
10. Лингвистические базы данных: модели и типы данных. Создания общих искусственных языков для представления информации.
11. Современный машинный перевод. Предпосылки возникновения и развитие машинного перевода (МП).
12. Последовательность формальных операций, обеспечивающих анализ и синтез в системе машинного перевода.
13. Стратегии машинного перевода. Автоматизированные и полуавтоматизированные системы перевода.
14. Области использования машинного перевода. Важнейшие системы машинного перевода.
15. Компьютерная лексикография как одно из направлений прикладной лингвистики. Словарные процессоры.
16. Основные понятия структуры словаря: словник, словарная статья, грамматические, стилистические пометы; иллюстративный материал
17. Типология электронных словарей.
18. Тезаурусы и терминологические словари.
19. Компьютерные технологии составления и эксплуатации словарей. Специальные программы– базы данных, компьютерные картотеки, программы обработки текста.
20. Корпусная лингвистика, ориентированная на создание национальных корпусов естественных языков.

IV. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

IV.1. Учебно-методические: основная и дополнительная литература

1. Баранов А. Н. Введение в прикладную лингвистику. М.,2011. http: //www. twirpx. com/file - электронное пособие
2. Беляев М. И., Григорьев С. Г., Гриншкун В. В., Демкин В. П., Краснова. А.,Макаров С. И., Роберт И. В., Щенников С. А. и др. Теория и практика создания образовательных электронных изданий. // М.: Изд-во РУДН, - 2003, 241 с. Часть 1. 72 с.
3. Гриншкун В. В. Григорьев С. Г. Образовательные электронные издания и ресурсы. // Учебно-методическое пособие для студентов педагогических вузов и слушателей системы повышения квалификации работников образования. Курск: КГУ, Москва: МГПУ - 2010, 98 с.
4. Зубов А. Б., Зубова И. И.. Информационные технологии в лингвистике. /М.: «Академия», - 2004.
5. Козырев В., Черняк В. Русская лексикография: Пособие для вузов – М., 2004.
6. Кузнецов А. А., Григорьев С. Г., Гриншкун В. В. Образовательные электронные издания и ресурсы: методическое пособие. М.: Дрофа,2009. 156 с.
7. Марчук Ю. Н. Основы компьютерной лингвистики. - М., 2007: 18-37с.

Дополнительная литература

1. Башмаков А. И., Старых В. А. Систематизация информационных ресурсов
2. Григорьев С. Г., Гриншкун В. В., Краснова Г. А. Основные принципы и методики использования системы порталов в учебном процессе. // В сб. научн. ст. «Интернет-порталы: содержание и технологии», Вып. 2. ГНИИ ИТТ «Информика», М.: Просвещение - 2004. С. 56-84.
3. Григорьев С. Г., Гриншкун В. В., Макаров С. И. Методико-технологические основы создания электронных средств обучения. // Научное издание. / Самара: Издательство Самарской государственной экономической академии. 2002. 110с.
5. Еляков А. Современное информационное общество. // Высшее образование в России. М., - 2001. №4.
6. Золотая рыбка в «сети». Интернет-технологии в средней школе // Под ред. Ольховской Л. И., Рудаковой Д. Т., Силаевой А. Г. - Москва, Прожект Хармони,Инк., 2001. - 168 с.
7. Интернет в профессиональной деятельности. // Под ред. Рудаковой Д. Т. - М.: ИОСО РАО, 2003. - 220 с.
8. Кобозева И. М. Лингвистическая семантика: учебник. – М.: Эдиториал РСС,2009. – 352 с.
9. Назарова Т. С., Полат Е. С. Средства обучения: технология создания и использования. // М.: Изд-во УРАО, - 1998. 204 с.

10. Панюкова С. В. Информационные и коммуникационные технологии в личностно - ориентированном обучении. // М.: ИОСО РАО - 1998, 225с

IV.2. Информационное и материально-техническое обеспечение дисциплины.

1. Онлайн- энциклопедия «Кругосвет». Гуманитарные науки: Лингвистика http: //www. krugosvet. ru/enc/gumanitarnye_nauki/lingvistika
2. Филологический портал Philology. ru http: //www. philology. ru/
3. Библиотека Гумер–языкознание http: //www. gumer. info/bibliotek_Buks/ Linguist/Index_Ling. php
4. Интернет-ресурс авторов УМК: http: //www. philippovich. ru/ – Учебно-научный кластер CLAIMM «Компьютерная Лингвистика, Искусственный Интеллект, Мультимедиа технологии».
5. http: //www. compling-info. narod. ru/ – Литература по Компьютерной Лингвистике книги, статьи, ссылки в Рунете
6. Dialog-21. ru – Международная конференция Диалог
7. Philology. ru – Русский филологический портал

Предусматриваются следующие образовательные технологии:

1. традиционные и интерактивные лекции с дискурсивной практикой обучения;
2. использование ситуационно-тематических и лингвистических задач, разбор конкретных лингвистических ситуаций, методологические тренинги;
3. семинары и коллоквиумы, на которых обсуждаются основные проблемы, освещенные в лекциях и сформулированные в домашних заданиях;
4. практические и/или устные домашние задания, подготовка доклада, творческого проекта;
5. участие в научно-методологических семинарах и конференциях;
6. консультации преподавателя;
7. самостоятельная работа студентов, в которую входит освоение теоретического материала, подготовка к семинарским занятиям с использованием интернета и электронных библиотек, выполнение письменных работ.

IV.3. Материально-техническое обеспечение дисциплины

Необходимо наличие компьютерного класса со следующим программным обеспечением:

- операционная система WindowsXP или Windows 7;
- пакет Microsoft Office 2003 иливыше;
- браузер Internet Explorer 6. 0 иливыше;браузер Mozilla FireFox 5. 0 иливыше