

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Уральский федеральный университет имени первого Президента России Б.Н. Ельцина»

УТВЕРЖДАЮ
Директор по образовательной
деятельности


С.Т. Князев
«19» октября 2020 г.



РАБОЧАЯ ПРОГРАММА МОДУЛЯ

Код модуля	Модуль
1149266	Лингвистика и информационные технологии

Екатеринбург, 2020

Перечень сведений о рабочей программе модуля	Учетные данные
Образовательная программа Теория языка и прикладная лингвистика	Код ОП 45.04.03/33.02
Направление подготовки Фундаментальная и прикладная лингвистика	Код направления и уровня подготовки 45.04.03

Программа модуля составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Клепинин Александр Владимирович	кандидат физико- математических наук	доцент	Департамент математики, механики и компьютерных наук
2	Пьянзина Елена Сергеевна	кандидат физико- математических наук	доцент	математической физики

Согласовано:

Учебный отдел

Е.С.Комарова

1. ОБЩАЯ ХАРАКТЕРИСТИКА МОДУЛЯ

1.1. Аннотация содержания модуля

Модуль включает дисциплины «Лингвистические основы информационного поиска», «Системы управления базами данных» и формирует представление о междисциплинарном характере современных лингвистических методов — в первую очередь пересечении лингвистики и информатики. Освоение дисциплин модуля создает предпосылки для работы в сфере современных информационных технологий, формирует компетенции, связанные с формализацией представления языковых данных.

1.2. Структура и объем модуля

Таблица 1

№ п/п	Перечень дисциплин модуля в последовательности их освоения	Объем дисциплин модуля и всего модуля в зачетных единицах
1	Лингвистические основы информационного поиска	3
2	Системы управления базами данных	3
ИТОГО по модулю:		6

1.3. Последовательность освоения модуля в образовательной программе

Пререквизиты модуля	
Постреквизиты и кореквизиты модуля	

1.4. Распределение компетенций по дисциплинам модуля, планируемые результаты обучения (индикаторы) по модулю

Таблица 2

Перечень дисциплин модуля	Код и наименование компетенции	Планируемые результаты обучения (индикаторы)
1	2	3
Лингвистические основы информационного поиска	ПК-7 - Способен осуществлять эффективное управление разработкой программных средств и информационных проектов	РО-1-З: Идентифицировать технологии и программы, используемые для формализации данных, в т.ч. лингвистических, для обработки естественного языка РО-1-У: Анализировать возможности и ограничения современных информационных

		<p>ресурсов и программных средств, предназначенных для анализа, систематизации и извлечения языковых данных.</p> <p>РО-1-ЛК: Способность критически оценивать продукты профессиональной деятельности, быть готовым к смене технологий в профессиональной деятельности, повышать свою цифровую грамотность</p>
Системы управления базами данных	ПК-7 - Способен осуществлять эффективное управление разработкой программных средств и информационных проектов	<p>РО-1-З: Идентифицировать технологии и программы, используемые для формализации данных, в т.ч. лингвистических, для обработки естественного языка</p> <p>РО-1-У: Анализировать возможности и ограничения современных информационных ресурсов и программных средств, предназначенных для анализа, систематизации и извлечения языковых данных.</p> <p>РО-1-В: Создавать техническое задание для информационного проекта.</p> <p>РО-1-ЛК: Способность критически оценивать продукты профессиональной деятельности, быть готовым к смене технологий в профессиональной деятельности, повышать свою цифровую грамотность</p>

1.5. Форма обучения

Обучение по дисциплинам модуля может осуществляться _____.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Лингвистические основы
информационного поиска

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
1	Клепинин Александр Владимирович	кандидат физико- математических наук	доцент	Департамент математики, механики и компьютерных наук

Рекомендовано учебно-методическим советом Уральского гуманитарного института

1. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы: Клепинин Александр Владимирович

1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
 - Продвинутый уровень

**Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания; Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.*

1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины	Содержание
P1	Булев поиск. Структуры булевого поиска и синтаксис запросов в булевом поиске.	Классическая модель поиска. Булевы запросы. Осуществление информационного поиска через прямое чтение файлов (программа grep). Матрица (таблица) "термин-документ". Инвертированный индекс. Биграммный индекс (biword index) и координатный индекс (positional index).
P2	Основные этапы построения индекса. Лемматизация и морфологический анализ.	Токенизация. Стоп-слова. Нормализация. Стемминг. Лемматизация.
P3	Поиск с ранжированием	Ограничения булева поиска. Jaccard coefficient. Term frequency. Inverted document frequency. Значение tf-idf. Векторное пространство документов и запросов.
P4	Классификация текста	Применение классификации. Решение задачи классификации с помощью экспертов и машинного обучения. Представление текста с помощью набора слов. Наивный байесовский классификатор. Применение формулы Байеса к задаче классификации. Упрощения, принимаемые при применении наивного байесовского классификатора. Обучение наивного байесовского классификатора.

P5	Оценка классификаторов	Причины оценки классификаторов. Оценка бинарных классификаторов. Оценка многоклассовых классификаторов.
P6	Оценка поиска	Возможные критерии оценки поиска. Ключевой критерий оценки поиска. Оценка поиска без ранжирования. Оценка поиска с ранжированием. Mean average precision.

1.3. Программа дисциплины реализуется на государственном языке Российской Федерации.

2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Электронные ресурсы (издания)

Гусякова, А. В. Информационные технологии и лингвистика XXI века : учебное пособие / А. В. Гусякова. – Москва : Московский педагогический государственный университет (МПГУ), 2016. – 96 с. : ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=469675>

Майстренко, А. В. Информационные технологии поддержки инженерной и научно-образовательной деятельности : учебное пособие / А. В. Майстренко, Н. В. Майстренко, И. В. Дидрих ; Тамбовский государственный технический университет. – Тамбов : Тамбовский государственный технический университет (ТГТУ), 2014. – 81 с. : табл., схем. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=277948>

Печатные издания

Романенко, В. Н. Сетевой информационный поиск : практ. пособие / В. Н. Романенко, Г. В. Никитина. — СПб. : Профессия, 2005.

Маннинг К., Рагхаван П., Шютце Х. Введение в информационный поиск. – М.: Вильямс, 2011.

Лукашевич, Н. В. Тезаурусы в задачах информационного поиска / Н. В. Лукашевич. — М. : Изд-во Моск. ун-та, 2011.

Паршукова Г. Б. Методика поиска профессиональной информации : учеб.-метод. пособие для студентов вузов / Г. Б. Паршукова. — СПб. : Профессия, 2006.

Профессиональные базы данных, информационно-справочные системы

Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

Базы данных, информационно-справочные и поисковые системы

MS Access, MS Excel

Яндекс, Google и др. поисковые системы

3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Подключение к сети Интернет	Операционная система Microsoft Windows, договор № 43-12/1670-2017 от 01.12.2017. Офисные пакеты Microsoft Office, договор № 43-12/1670-2017 от 01.12.2017. Браузер Google Chrome – свободное ПО; Браузер Mozilla Firefox – свободное ПО; Daemon Tools Lite – свободное ПО; BurnAware Free – свободное ПО; Acrobat Reader DC – свободное ПО.
2	Практические занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная	Операционная система Microsoft Windows, договор № 43-12/1670-2017 от 01.12.2017. Офисные пакеты Microsoft Office, договор № 43-12/1670-2017 от 01.12.2017. Браузер Google Chrome – свободное ПО; Браузер Mozilla Firefox – свободное ПО; Daemon Tools Lite – свободное ПО; BurnAware Free – свободное ПО; Acrobat Reader DC – свободное ПО.
3	Текущий контроль и промежуточная аттестация	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная	Операционная система Microsoft Windows, договор № 43-12/1670-2017 от 01.12.2017. Офисные пакеты Microsoft Office, договор № 43-12/1670-2017 от 01.12.2017. Браузер Google Chrome – свободное ПО; Браузер Mozilla Firefox – свободное ПО; Daemon Tools Lite – свободное ПО;

			BurnAware Free – свободное ПО; Acrobat Reader DC – свободное ПО.
4	Самостоятельная работа студентов	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Подключение к сети Интернет	Операционная система Microsoft Windows, договор № 43-12/1670-2017 от 01.12.2017. Офисные пакеты Microsoft Office, договор № 43-12/1670-2017 от 01.12.2017. Браузер Google Chrome – свободное ПО; Браузер Mozilla Firefox – свободное ПО; Daemon Tools Lite – свободное ПО; BurnAware Free – свободное ПО; Acrobat Reader DC – свободное ПО.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ

Вопросы к зачету/экзамену по дисциплине

1. Классическая модель поиска. Булевы запросы. Осуществление информационного поиска через прямое чтение файлов (программа grep).
2. Матрица (таблица) "термин-документ". Инвертированный индекс. Биграммный индекс (biword index) и координатный индекс (positional index).
3. Токенизация. Стоп-слова. Нормализация.
4. Нормализация. Стеemming. Лемматизация.
5. Ограничения булева поиска. Jaccard coefficient. Term frequency. Inverted document frequency.
6. Term frequency. Inverted document frequency. Значение tf-idf. Векторное пространство документов и запросов.
7. Применение классификации. Решение задачи классификации с помощью экспертов и машинного обучения.
8. Представление текста с помощью набора слов. Наивный байесовский классификатор. Применение формулы Байеса к задаче классификации. Упрощения, принимаемые при применении наивного байесовского классификатора. Обучение наивного байесовского классификатора.
9. Причины оценки классификаторов. Оценка бинарных классификаторов.
10. Оценка многоклассовых классификаторов.
11. Возможные критерии оценки поиска. Ключевой критерий оценки поиска.
12. Оценка поиска без ранжирования. Оценка поиска с ранжированием. Mean average precision.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Системы управления базами данных

Рабочая программа дисциплины составлена авторами:

№ п/п	Фамилия Имя Отчество	Ученая степень, ученое звание	Должность	Подразделение
2	Пьянзина Елена Сергеевна	кандидат физико- математических наук	доцент	математической физики

Рекомендовано учебно-методическим советом Уральского гуманитарного института

2. СОДЕРЖАНИЕ И ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ

Авторы: Пьянзина Елена Сергеевна

1.1. Технологии реализации, используемые при изучении дисциплины модуля

- Традиционная (репродуктивная) технология
- Разноуровневое (дифференцированное) обучение
 - Продвинутый уровень

**Базовый I уровень – сохраняет логику самой науки и позволяет получить упрощенное, но верное и полное представление о предмете дисциплины, требует знание системы понятий, умение решать проблемные ситуации. Освоение данного уровня результатов обучения должно обеспечить формирование запланированных компетенций и позволит обучающемуся на минимальном уровне самостоятельности и ответственности выполнять задания; Продвинутый II уровень – углубляет и обогащает базовый уровень как по содержанию, так и по глубине проработки материала дисциплины. Это происходит за счет включения дополнительной информации. Данный уровень требует умения решать проблемы в рамках курса и смежных курсов посредством самостоятельной постановки цели и выбора программы действий. Освоение данного уровня результатов обучения позволит обучающемуся повысить уровень самостоятельности и ответственности до творческого применения знаний и умений.*

1.2. Содержание дисциплины

Таблица 1.1

Код раздела, темы	Раздел, тема дисциплины	Содержание
Раздел 1	Теоретические основы баз данных	Представление о базах данных и системах управления базами данных. Основные понятия: база данных, таблица, поле, ключ, запрос, форма, отчет. Предметная область, данные. Жизненный цикл базы данных. Проектирование БД. Назначение и принципы построения баз данных. Реляционная модель данных. Описание проблемной области: объекты, атрибуты, связи. Разные типы связи между таблицами базы данных.
Раздел 2	Разработка базы данных	Пользователи, администраторы БД, их функции. Нормализация данных и обеспечение их целостности. Таблица: поля, ключевые поля, логические записи. Связь, типы связей, нормализация отношений. Схема данных, создание модели данных. Требования и ограничения к проблемной области. Целостность и последствия ее нарушения.

Раздел 3	Пользовательский интерфейс базы данных	<p>Форма. Виды форм: автоформа, главная / подчиненная, табличная, диаграмма. Создание формы с помощью мастера, с помощью конструктора.</p> <p>Отчет. Виды отчетов: автоотчет, многоуровневый, почтовые наклейки, диаграмма. Создание отчета с помощью мастера, с помощью конструктора.</p> <p>Разработка структурной схемы меню. Создание главной кнопочной формы. Диспетчер кнопочных форм. Создание пользовательского меню.</p>
Раздел 4	Процессы обработки данных в базе данных	<p>Организация процессов обработки данных в БД. Ограничения целостности данных. Технология оперативной обработки транзакции. Независимость данных. Принципы независимости данных. Целостность баз данных. Безопасность данных. Запросы и их виды: на выборку, обновление, создание таблицы, добавление записей, удаление записей, перекрестный, параметрический, с группировкой. Создание запроса с помощью мастера, с помощью конструктора. Организация расчетов в таблицах.</p>
Раздел 5	Основы программирования средствами SQL	<p>Основные инструкции языка. Структура инструкции: команда, ключевое слово.</p> <p>Язык SQL: инструкция SELECT. Запросы на выборку. Необходимость их применения. Ключевые слова: FROM, WHERE. Создание условий отбора, используемые операторы. Использование вычисляемых полей. Ключевые слова: GROUP BY, HAVING. Требования при написании групповых запросов.</p> <p>Язык SQL: инструкция UPDATE. Последовательность выполнения запросов на обновление.</p> <p>Язык SQL: инструкция DELETE. Последовательность выполнения запросов на удаление.</p> <p>Запросы с использованием нескольких таблиц.</p>

2.3. Программа дисциплины реализуется на государственном языке Российской Федерации.

2. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Электронные ресурсы (издания)

Гущин А.Н. Базы данных : учебно-методическое пособие [Электронный ресурс] / А.Н. Гущин. - 2-е изд., испр. и доп. - М. ; Берлин : Директ-Медиа, 2015. - Режим доступа : <http://biblioclub.ru/index.php?page=book&id=278093>

Чурбанова О.В. Базы данных и знаний. Проектирование баз данных в Microsoft Access [Электронный ресурс] : учебно-методическое пособие / О.В. Чурбанова, А.Л. Чурбанов. - Архангельск : САФУ, 2015. - Режим доступа : <http://biblioclub.ru/index.php?page=book&id=436230>

Щелоков С. А. Разработка и создание баз данных средствами СУБД Access и SQL Server [Электронный ресурс] / С. А. Щелоков. - Оренбург, ОГУ, 2014. - Режим доступа : <http://biblioclub.ru/index.php?page=book&id=260754>

Бедердинова, О. И. Информационные технологии общего назначения : учебное пособие / О. И. Бедердинова, Ю. А. Водовозова ; Северный (Арктический) федеральный университет им. М. В. Ломоносова. – Архангельск : Северный (Арктический) федеральный университет (САФУ), 2015. – 84 с. : ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=436288>

Печатные издания

Малыхина, Мария П. Базы данных: основы, проектирование, использование : [учеб. пособие для вузов] / М. П. Малыхина. — СПб. : БХВ-Петербург, 2004. — 512 с.

Профессиональные базы данных, информационно-справочные системы

Материалы для лиц с ОВЗ

Весь контент ЭБС представлен в виде файлов специального формата для воспроизведения синтезатором речи, а также в тестовом виде, пригодном для прочтения с использованием экранной лупы и настройкой контрастности.

Базы данных, информационно-справочные и поисковые системы

MS Access

3. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

Сведения об оснащённости дисциплины специализированным и лабораторным оборудованием и программным обеспечением

Таблица 3.1

№ п/п	Виды занятий	Оснащённость специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
1	Лекции	Мебель аудиторная с количеством рабочих мест в	Операционная система Microsoft Windows, договор № 43-12/1670-2017 от 01.12.2017. Офисные

		соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная Подключение к сети Интернет	пакеты Microsoft Office, договор № 43-12/1670-2017 от 01.12.2017. Браузер Google Chrome – свободное ПО; Браузер Mozilla Firefox – свободное ПО; Daemon Tools Lite – свободное ПО; BurnAware Free – свободное ПО; Acrobat Reader DC – свободное ПО.
2	Практические занятия	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная	Операционная система Microsoft Windows, договор № 43-12/1670-2017 от 01.12.2017. Офисные пакеты Microsoft Office, договор № 43-12/1670-2017 от 01.12.2017. Браузер Google Chrome – свободное ПО; Браузер Mozilla Firefox – свободное ПО; Daemon Tools Lite – свободное ПО; BurnAware Free – свободное ПО; Acrobat Reader DC – свободное ПО.
3	Текущий контроль и промежуточная аттестация	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Рабочее место преподавателя Доска аудиторная	Операционная система Microsoft Windows, договор № 43-12/1670-2017 от 01.12.2017. Офисные пакеты Microsoft Office, договор № 43-12/1670-2017 от 01.12.2017. Браузер Google Chrome – свободное ПО; Браузер Mozilla Firefox – свободное ПО; Daemon Tools Lite – свободное ПО; BurnAware Free – свободное ПО; Acrobat Reader DC – свободное ПО.
4	Самостоятельная работа студентов	Мебель аудиторная с количеством рабочих мест в соответствии с количеством студентов Подключение к сети Интернет	Операционная система Microsoft Windows, договор № 43-12/1670-2017 от 01.12.2017. Офисные пакеты Microsoft Office, договор № 43-12/1670-2017 от 01.12.2017. Браузер Google Chrome – свободное ПО; Браузер Mozilla Firefox – свободное ПО; Daemon Tools Lite – свободное ПО; BurnAware Free – свободное ПО; Acrobat Reader DC – свободное ПО.

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПО ДИСЦИПЛИНЕ

Вопросы к зачету/экзамену по дисциплине

1. Модели данных, нормализация отношений.
2. Основные объекты реляционной модели данных: отношение, связь, тип, домен, атрибут, кортеж, первичный ключ, внешний ключ.
3. Проектирование структуры базы, связи между таблицами, контроль целостности.
4. Выделение сущностей предметной области и выявление связей между ними.
5. Типы связей: один к одному, один ко многим, многие ко многим.
6. Поддержка целостности в реляционной модели данных: структурный, языковой и семантический аспекты.
7. Создание базы данных в среде MS Access.
8. Формирование структуры таблиц разными способами.
9. Основные типы данных, обязательные и дополнительные свойства полей.
10. Схема данных и создание связей между таблицами.
11. Условия включения автоматического контроля целостности данных в связанных таблицах.
12. Сортировка записей по одному или нескольким полям.
13. Фильтрация записей в таблице: команды Фильтр «по выделенному», Фильтр для, Изменить фильтр.
14. Отбор информации по нескольким критериям: построение расширенного фильтра, использование выражений.
15. Простые и многотабличные запросы.
16. Классификация запросов на выборку и запросов на действие.
17. Вычисляемые поля и построитель выражений.
18. Формирование динамических условий отбора в запросах с параметрами.
19. Особенности построения запросов, основанных на других запросах.
20. Виды форм, создание формы с помощью мастера, конструктора, функции Автоформа.
21. Структура формы и основные виды элементов управления.
22. Флажки, кнопки выбора, поля со списком.
23. Отчеты и группировка данных.
24. Модификация отчетов в режиме конструктора и создание отчета с нуля.
25. Организация пользовательского интерфейса.
26. Создание главной кнопочной формы с помощью диспетчера.
27. Особенности модификации кнопочных форм.
28. Разработка системы меню и панелей инструментов.
29. Создание вводного экрана.
30. Настройка параметров запуска.