МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное учреждение
высшего профессионального образования
"Казанский (Приволжский) федеральный университет"
Высшая школа информационных технологий и информационных систем

УТВЕРЖДАЮ
Проректор
по образовательной деятельности КФУ
Проф. Минзарипов Р.Г.
_____
"___"_____20___ г.

Программа дисциплины

Скриптинг Б3.ДВ.3

Направление подготовки: 230700.62 - Прикладная информатика
Профиль подготовки: не предусмотрено
Квалификация выпускника: бакалавр
Форма обучения: очное
Язык обучения: русский

Автор(ы):
Насрутдинов М.Ф. , Егорчев А.А.
Рецензент(ы):
Чикрин Д.Е.

СОГЛАСОВАНО:
Заведующий(ая) кафедрой: Чикрин Д. Е.
Протокол заседания кафедры No ___ от "____" _____ 201__г
Учебно-методическая комиссия Высшей школы информационных технологий и информационных систем:
Протокол заседания УМК No ____ от "____" _____ 201__г

Регистрационный No

Казань
2014

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) ассистент, б/с Егорчев А.А. кафедра Автономные робототехнические системы Высшая школа информационных технологий и информационных систем , AAEgorchev@kpfu.ru ; заместитель директора Насрутдинов М.Ф. Высшая школа информационных технологий и информационных систем КФУ , Marat.Nasrutdinov@kpfu.ru

1. Цели освоения дисциплины

Цель дисциплины заключается в формировании у студентов компетенции в области автоматизации рутинных процессов, связанных с администрированием операционных систем (на примере *NIX систем)

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

Данная учебная дисциплина включена в раздел " Б3.ДВ.3 Профессиональный" основной образовательной программы 230700.62 Прикладная информатика и относится к дисциплинам по выбору. Осваивается на 3 курсе, 6 семестр.

Данная учебная дисциплина включена в раздел " Б3.ДВ.3 Профессиональный" основной образовательной программы 230700.62 Прикладная информатика и относится к дисциплинам по выбору. Осваивается на 3 курсе, 6 семестр.

Курс "Скриптинг" предполагает знание студентами основ программирования и базовых навыков работы в *NIX-системах.

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)	способен использовать, обобщать и анализировать информацию, ставить цели и находить пути их достижения в условиях формирования и развития информационного общества
ОК-3 (общекультурные компетенции)	способен работать в коллективе, нести ответственность за поддержание партнерских, доверительных отношений
ОК-4 (общекультурные компетенции)	способен находить организационно-управленческие решения и готов нести за них ответственность
ОК-5 (общекультурные компетенции)	способен самостоятельно приобретать и использовать в практической деятельности новые знания и умения, стремится к саморазвитию
ОК-6 (общекультурные компетенции)	способен осознавать социальную значимость своей будущей профессии, обладать высокой мотивацией к выполнению профессиональной деятельности
ПК-10 (профессиональные компетенции)	способен применять к решению прикладных задач базовые алгоритмы обработки информации, выполнять оценку сложности алгоритмов, программировать и тестировать программы
ПК-4 (профессиональные компетенции)	способен ставить и решать прикладные задачи с использованием современных информационно-коммуникационных технологий

Шифр компетенции	Расшифровка приобретаемой компетенции
ПК-5 (профессиональные компетенции)	способен осуществлять и обосновывать выбор проектных решений по видам обеспечения информационных систем
ПК-6 (профессиональные компетенции)	способен документировать процессы создания информационных систем на всех стадиях жизненного цикла

В результате освоения дисциплины студент:

4. должен демонстрировать способность и готовность:

- понимать основные концепции скриптовых языков
- уметь выбирать скриптовый язык в зависимости от задач
- знать основные системные команды оболочки shell
- уметь писать скрипты.

4. Структура и содержание дисциплины/ модуля

Общая трудоемкость дисциплины составляет 4 зачетных(ые) единиц(ы) 144 часа(ов).

Форма промежуточного контроля дисциплины экзамен 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. Тема 1. Введение в программирование на Shell. Начальные понятия и примеры. Запуск сценариев.	6	1	2	0	0	домашнее задание
2.	Тема 2. Тема 2. Специальные символы.	6	2	2	4	0	домашнее задание
3.	Тема 3. Тема 3. Переменные и параметры. Подстановка переменных	6	3	2	6	0	контрольная работа

N	Раздел Дисциплины/ Модуля	Семестр	Неделя семестра	Виды и часы аудиторной работы, их трудоемкость (в часах)			Текущие формы контроля
				Лекции	Практические занятия	Лабораторные работы	
4.	Тема 4. Конструкции проверки условий.	6	4-5	4	8	0	домашнее задание
5.	Тема 5. Внутренние команды	6	6-7	4	8	0	домашнее задание
6.	Тема 6. Внешние команды	6	8-9	4	10	0	контрольная точка
.	Тема . Итоговая форма контроля	6		0	0	0	экзамен
	Итого			18	36	0	

4.2 Содержание дисциплины

Тема 1. Тема 1. Введение в программирование на Shell. Начальные понятия и примеры. Запуск сценариев.

лекционное занятие (2 часа(ов)):

Введение. Актуальность, мотивация. Сферы применения. Примеры приложений.

Тема 2. Тема 2. Специальные символы.

лекционное занятие (2 часа(ов)):

Представление текста в ЭВМ. Понятие набора знаков. Кодовая страница, элемента кода. Кодировки. Методы автоматического определения кодировки. Методы автоматического определения языка текста.

практическое занятие (4 часа(ов)):

Введение в платформу Apache UIMA. Обзор альтернативных программных библиотек, их преимуществ и недостатков. Обзор архитектуры UIMA, основных компонент: CAS, конвейер, аннотаторы, дескрипторы. Создание простейших аннотаторов и их дескрипторов. Запуск простого конвейера.

Тема 3. Тема 3. Переменные и параметры. Подстановка переменных

лекционное занятие (2 часа(ов)):

Понятие токенизации. Основные подходы. Автоматические разбиение текста на предложения и другие структурные блоки. Основные проблемы и методы. Особенности сегментации русскоязычных текстов.

практическое занятие (6 часа(ов)):

Разработка сегментатора текста. Работа с UIMA CAS Editor. Создание небольшого тестового документа Написание программы оценки качества сегментатора. Консультация по задачам, заданными для самостоятельной работы.

Тема 4. Конструкции проверки условий.

лекционное занятие (4 часа(ов)):

Задачи морфологического анализа, морфологический разбор, стемминг, лемматизация. Понятия лексемы, словоформы, леммы, морфемы, псевдо-основы и псевдо-окончания. Грамматические категории. Словоизменительная парадигма. Морфотактика. Структура данных морфологического словаря, лексикона. Грамматические модели русского языка в контексте автоматической обработки.

практическое занятие (8 часа(ов)):

Регулярные выражения. Конечные автоматы. Конечные преобразователи. Их применение для словарного морфологического анализа. Минимальное расстояние редактирования. Алгоритм подсчета расстояния Левенштейна. Практика по подсчету минимального расстояния Левенштейна. Работа с морфологическим словарем OpenCorpora. Понятие внешних ресурсов в UIMA. Практика по реализации и применению внешних ресурсов на примере словаря OpenCorpora. Обзор доступного инструментария для морфологического анализа русского языка. Консультация по задачам, заданными для самостоятельной работы.

Тема 5. Внутренние команды

лекционное занятие (4 часа(ов)):

Основные внутренние команды и сценарии их использования

практическое занятие (8 часа(ов)):

Нюансы программной реализации языковых моделей. Обзор программных инструментариев для реализации и применения языковых моделей. Сбор экспериментального корпуса. Практическая работа с инструментарием SRILM. Написание скриптов с использованием внутренних команд

Тема 6. Внешние команды

лекционное занятие (4 часа(ов)):

Basic Commands Complex Commands Time / Date Commands Text Processing Commands File and Archiving Commands Communications Commands Terminal Control Commands Math Commands Miscellaneous Commands

практическое занятие (10 часа(ов)):

Написание скриптов с использованием внешних команд

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. Тема 1. Введение в программирование на Shell. Начальные понятия и примеры. Запуск сценариев.	6	1	подготовка домашнего задания	2	домашнее задание
2.	Тема 2. Тема 2. Специальные символы.	6	2	подготовка домашнего задания	2	домашнее задание
3.	Тема 3. Тема 3. Переменные и параметры. Подстановка переменных	6	3	подготовка к контрольной работе	6	контрольная работа
4.	Тема 4. Конструкции проверки условий.	6	4-5	подготовка домашнего задания	8	домашнее задание
5.	Тема 5. Внутренние команды	6	6-7	подготовка домашнего задания	18	домашнее задание
6.	Тема 6. Внешние команды	6	8-9	подготовка к контрольной точке	18	контрольная точка
	Итого				54	

5. Образовательные технологии, включая интерактивные формы обучения

На практических занятиях и при выполнении курсовых предполагается разбиение студентов на группы для командной работы. В команде выделяются роли, ответственные за разработку конкретного модуля, участники команды должны предварительно договориться о разграничении работ и сборке целостного результата командной работы. При командной работе используется система контроля версий (Git), имитируется реальный процесс промышленной разработки программного комплекса.

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

Тема 1. Тема 1. Введение в программирование на Shell. Начальные понятия и примеры. Запуск сценариев.

домашнее задание , примерные вопросы:

Введение. Актуальность, мотивация. Сферы применения. Примеры приложений.

Тема 2. Тема 2. Специальные символы.

домашнее задание , примерные вопросы:

Представление текста в ЭВМ. Понятие набора знаков. Кодовая страница, элемента кода. Кодировки. Методы автоматического определения кодировки. Методы автоматического определения языка текста.

Тема 3. Тема 3. Переменные и параметры. Подстановка переменных

контрольная работа , примерные вопросы:

Понятие токенизации. Основные подходы. Автоматические разбиение текста на предложения и другие структурные блоки. Основные проблемы и методы. Особенности сегментации русскоязычных текстов.

Тема 4. Конструкции проверки условий.

домашнее задание , примерные вопросы:

Задачи морфологического анализа, морфологический разбор, стемминг, лемматизация. Понятия лексемы, словоформы, леммы, морфемы, псевдо-основы и псевдо-окончания. Грамматические категории. Словоизменительная парадигма. Морфотактика. Структура данных морфологического словаря, лексикона. Грамматические модели русского языка в контексте автоматической обработки.

Тема 5. Внутренние команды

домашнее задание , примерные вопросы:

Основные внутренние команды и сценарии их использования Нюансы программной реализации языковых моделей. Обзор программных инструментариев для реализации и применения языковых моделей. Сбор экспериментального корпуса. Практическая работа с инструментарием SRILM.

Тема 6. Внешние команды

контрольная точка , примерные вопросы:

Basic Commands Complex Commands Time / Date Commands Text Processing Commands File and Archiving Commands Communications Commands Terminal Control Commands Math Commands Miscellaneous Commands Написание скриптов с использованием внешних команд

Тема . Итоговая форма контроля

Примерные вопросы к экзамену:

Экзаменационные билеты

Билет ◆1

1. Актуальность использования shell, мотивация. Сферы применения.
2. Написание скриптов с использованием внутренних команд.

Билет ◆2

1. Примеры приложений. Представление текста в ЭВМ..
2. Практическая работа с инструментарием SRILM.

Билет ◆3

1. Понятие набора знаков. Кодовая страница, элемента кода.
2. Сбор экспериментального корпуса.

Билет ◆4

1. Кодировки. Методы автоматического определения кодировки.
2 Обзор доступного инструментария для морфологического анализа русского языка.

Билет ◆5

1. Методы автоматического определения языка текста.
2. Практика по реализации и применению внешних ресурсов на примере словаря OpenCorpora.

Билет ◆6

1. Apache UIMA. Обзор альтернативных программных библиотек, их преимуществ и недостатков.
2. Понятие внешних ресурсов в UIMA.

Билет ◆7

1. Обзор архитектуры UIMA, основных компонент: CAS, конвейер.
2. Работа с морфологическим словарем OpenCorpora.

Билет ◆8

1. Аннотаторы, дескрипторы. Создание простейших аннотаторов и их дескрипторов.
2. Построение графиков (GNUPLOT и XGRACE).

Билет ◆9

1. Запуск простого конвейера. Понятие токенизации. Основные подходы.
2. Использование пакетов прикладных программ. .

Билет ◆10

1. Автоматические разбиение текста на предложения и другие структурные блоки.
2. Программирование на языке командного интерпретатора.

Билет ◆11

1. Основные проблемы и методы. Особенности сегментации русскоязычных текстов.
2. Запуск программ на исполнение.

Билет ◆12

1. Разработка сегментатора текста.
2. Система управления процессами.

Билет ◆13

1. Работа с UIMA CAS Editor.
2. Устройство файловой системы и работа с файлами.

Билет ◆14

1. Создание небольшого тестового документа Написание программы оценки качества сегментатора
2. Алгоритм подсчета расстояния Левенштейна. Практика по подсчету минимального расстояния Левенштейна.

Билет ◆15

1. Задачи морфологического анализа, морфологический разбор, стемминг, лемматизация.
2. Минимальное расстояние редактирования.

Билет ◆16

1. Понятия лексемы, словоформы, леммы, морфемы, псевдо-основы и псевдо-окончания.
2. Конечные преобразователи. Их применение для словарного морфологического анализа..

Билет ◆17

1. Грамматические категории.
2. Конечные автоматы.

Билет ◆18

1. Словоизменительная парадигма. Морфотактика.
2. Регулярные выражения.

Билет ◆19

1. Структура данных морфологического словаря, лексикона.
2. Аннотаторы, дескрипторы. Создание простейших аннотаторов и их дескрипторов.

Билет ◆20

1. Запуск простого конвейера. Понятие токенизации. Основные подходы.
2. Программирование на языке командного интерпретатора.

Билет ◆21

1. Apache UIMA. Обзор альтернативных программных библиотек, их преимуществ и недостатков.
2. Обзор архитектуры UIMA, основных компонент: CAS, конвейер.

Билет ◆22

3. Программирование на языке командного интерпретатора.
4. Аннотаторы, дескрипторы. Создание простейших аннотаторов и их дескрипторов.

Билет ◆23

1. Конечные автоматы.
2. Apache UIMA. Обзор альтернативных программных библиотек, их преимуществ и недостатков.

Билет ◆24

1. Работа с UIMA CAS Editor.
2. Программирование на языке командного интерпретатора.

Билет ◆25

1. Обзор архитектуры UIMA, основных компонент: CAS, конвейер.
2. Запуск простого конвейера. Понятие токенизации. Основные подходы.

Вопросы по темам для контрольных работ.

Тема 1.

Язык. Текст. Основы лингвистики и теории речевой коммуникации. Язык или языки. Текст или Тексты. Основы речевой коммуникации. Лингвистика и лингвистики. Принцип моделирования. Цели. Слово ? коллокация - синтаксические конструкции - текст. Единица анализа и контекст. Инвентарные и конструктивные единицы. Понятие "текущего словаря". Избыточность. Контекстная предсказуемость. Единица анализа и контекст. Коллокации и конструкции. Типы коллокаций и конструкций. Принцип шкалирования. Семантическая и информационная структуры при анализе текстов и/или коллекций. Основные элементы этих структур. Анализ текста в парадигме когнитивных исследований. Анализ текста в парадигмах автоматического понимания текста. Коммуникативная и информационная (смысловая) структуры текста. Избыточность. Компрессия текста. Свертки текста.

ЭЛЕКТРОННЫЙ УНИВЕРСИТЕТ
ИНФОРМАЦИОННО-АНАЛИТИЧЕСКАЯ СИСТЕМА КФУ

Тема 2.

Морфологический анализ и синтез. Словарный морфологический анализ и синтез. Автоматизированное пополнение морфологического словаря. Коррекция орфографических ошибок. Постморфологический и предсинтаксический анализ. Автоматизированное снятие омонимии. Постморфологический анализ. Синтаксическая сегментация. Инструментальные системы разработки приложений по автоматической обработке текстов на естественном языке. Программные средства лингвистической обработки. Представление лингвистических данных. Подходы к представлению данных. Лингвистическая разметка. Лингвистические аннотации. Представления, основанные на абстракции. Недоспецифицированные представления.

Тема 3.

Архитектура инструментальных ЕЯ-систем Компонентная организация. Процессы обработки текста. Системы обработки ЕЯ-текстов. Системы на базе разметки. Системы на базе аннотаций. Системы интеграции поверхностной и глубокой. Системы, развивающие отдельные аспекты обработки текста. Прочие системы. Алгоритмы классификации полнотекстовых документов. Алгоритмы классификации с учителем. Представление данных в задачах классификации текстов. Отбор терминов для классификации. Алгоритм "наивной" байесовской классификации. Алгоритм Роккио. Алгоритм k-ближайших соседей. Алгоритм опорных векторов. Алгоритм деревьев принятия решений. Алгоритм наименьших квадратов. Экспериментальная оценка результата классификации с учителем.

Тема 4.

Алгоритмы классификации без учителя. Иерархические алгоритмы. Алгоритм k-средних Плотностный алгоритм DBSCAN. Нечёткий алгоритм c-средних. Инкрементный алгоритм C2ICM. Нейросетевой алгоритм SOM. Экспериментальная оценка результата классификации без учителя. Выбор метода классификации без учителя. Информационные потоки и сложные сети. Основы анализа информационного пространства и информационных потоков. Понятие информационного пространства. Информационный поток как объект исследования.Тематические информационные потоки. Моделирование информационных потоков. Модель диффузии информации. Самоподобие в информационном пространстве.

7.1. Основная литература:

1. Климов, А. П. JavaScript на примерах [Электронный ресурс] / А.П. Климов . ? 2-е изд., перераб. и доп. ? СПб.: БХВ-Петербург, 2009. ? 336 с.: ил. - ISBN 978-5-9775-0361-7

http://znanium.com/bookread.php?book=489238

2. Колисниченко Д. Н. Руководство по командам и shell-программированию в Linux. ? СПб.: БХВ-Петербург, 2011. ? 283 с. ? (БЛЦ). - ISBN 978-5-9775-0619-9.

http://znanium.com/bookread.php?book=355000

3. Операционные системы, среды и оболочки: Учебное пособие / Т.Л. Партыка, И.И. Попов. - 5-е изд., перераб. и доп. - М.: Форум: НИЦ ИНФРА-М, 2013. - 560 с.: ил.; 60x90 1/16. - (Профессиональное образование). (п) ISBN 978-5-91134-743-7, 1000 экз.

http://znanium.com/bookread.php?book=405821

7.2. Дополнительная литература:

1. Коробко, И. В. Администрирование сетей Windows с помощью сценариев [Электронный ресурс] / И. В. Коробко. - СПб.: БХВ-Петербург, 2007. - 368 с.: ил. - (Системный администратор). - ISBN 978-5-9775-0140-8.

http://znanium.com/bookread.php?book=489562

2. Прохоренок, Н. А. HTML, JavaScript, PHP и MySQL. Джентльменский набор Web-мастера / Николай Прохоренок. ? 3-е изд., перераб. и доп. ? СПб.: БХВ-Петербург, 2010. ? 900 с. ? (Профессиональное программирование). - ISBN 978-5-9775-0540-6.

http://znanium.com/bookread.php?book=350905

7.3. Интернет-ресурсы:

Интернет-университет 'Интуит' - http://www.intuit.ru/

Портал КФУ - http://www.kpfu.ru/

СДО КФУ - http://tulpar.kpfu.ru/

ЭБС 'Знаниум' - http://znanium.com/

ЭБС 'Лань' - http://e.lanbook.com/

8. Материально-техническое обеспечение дисциплины(модуля)

Освоение дисциплины "Скриптинг" предполагает использование следующего материально-технического обеспечения:

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.

Компьютерный класс с проектором и подключением к сети Интернет.

Программа составлена в соответствии с требованиями ФГОС ВПО и учебным планом по направлению 230700.62 "Прикладная информатика" и профилю подготовки не предусмотрено .

Автор(ы):
Насрутдинов М.Ф. _____
Егорчев А.А. _____
"__" _____ 201 __ г.

Рецензент(ы):
Чикрин Д.Е. _____
"__" _____ 201 __ г.