

Уч 2015-2016

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»
(ВлГУ)



УТВЕРЖДАЮ
Проректор
по образовательной деятельности

А.А.Панфилов

« 29 » 12 2016 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Технологии обработки данных

Специальность 10.05.04 "Информационно-аналитические системы безопасности"

Специализация "Автоматизация информационно-аналитической деятельности"

Уровень высшего образования специалитет

Форма обучения очная

Семестр	Трудоем- кость зач. ед./час.	Лек- ций, час.	Практич. занятий, час.	Лаборат. работ, час.	СРС, час.	Форма промежуточного контроля (экз./зачет)
4	3/108	18		36	54	Зачет
5	6/216	36		36	99	Экзамен (45ч), КР
Итого	9/324	54		72	153	Зачет, Экзамен (45ч), КР

Владимир 2016

1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целями освоения дисциплины «Технологии обработки данных» являются обеспечение профессиональной подготовки студентов в соответствии с требованиями ФГОС ВО и учебного плана по специальности 10.05.04 «Информационно-аналитические системы безопасности», формирование у студентов обобщенного представления о понятийном аппарате информационных технологий (ИТ) и технологий обработки данных и информации.

Задачей изучения дисциплины «Технологии обработки данных» является изучение: классификации информации и данных; основных процессах обработки данных и информации; методах и проблемах интеграции данных и информации; методах обработки информации в зависимости от ее типа и формы представления.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО СПЕЦИАЛИТЕТА

Данная дисциплина относится к дисциплинам по выбору вариативной части Блока Б1 (код Б1.В.ДВ.2). В учебном плане предусмотрены виды учебной деятельности, обеспечивающие синтез теоретических лекций и лабораторных работ, ориентированных на освоение студентами современных информационных технологий и технологий обработки данных и информации, а также методов и способов их применения в профессиональной деятельности. Курс тесно взаимосвязан с другими дисциплинами данного цикла, например, «Основы информационной безопасности», «Базы данных и экспертные системы», «Безопасность операционных систем».

Дисциплина изучается на втором и третьем курсах, требования к «входным» знаниям, умениям и готовностям (пререквизитам) обучающегося определяются требованиями к уровню подготовки по специальности 10.05.04 «Информационно-аналитические системы безопасности» по курсам «Информатика», «Структуры данных», «Технологии и методы программирования», «Языки программирования», «Базы данных и экспертные системы», «Сети и системы передачи информации», «Системное программное обеспечение».

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате освоения дисциплины студент должен обладать следующими общепрофессиональными компетенциями:

ОПК-4 - способностью применять в профессиональной деятельности языки и системы программирования, инструментальные средства разработки программного обеспечения, современные методы и технологии программирования; профессионально-специализированными компетенциями:

ПСК-1.2 – способностью разрабатывать и применять автоматизированные технологии обработки естественно-языковых текстов и формализованных данных при решении информационно-аналитических задач.

В результате освоения дисциплины обучающийся должен демонстрировать следующие результаты образования:

1) Знать: - понятийно-категориальный аппарат дисциплины; основные виды и процедуры обработки информации, модели и методы решения задач обработки информации (ОПК-4, ПСК-1.2);

2) Уметь: - ставить и решать типовые задачи с помощью современных информационных технологий обработки данных; применять на пользовательском уровне основные инструментальные средства обработки данных; использовать информационно-поисковые средства локальных и глобальных вычислительных и информационных сетей; использовать типовые алгоритмы обработки информации; применять прикладные информационные технологии обработки информации (ОПК-4, ПСК-1.2);

3) Владеть: - навыками применения современных технологий обработки данных к текущим реальным ситуациям; навыками развертывания основных программных комплексов и программ, реализующих ту или иную технологию обработки данных; методологией использования инструментальных средств информационных технологий при обработке информации и данных; технологиями интеллектуального анализа данных (ОПК-4, ПСК-1.2).

У обучаемых в процессе изучения дисциплины должны вырабатываться дополнительные компетенции, с учетом требований работодателей: - Способность выбора оптимального системного программного обеспечения для решения прикладных задач в области обеспечения информационной безопасности.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Общая трудоемкость дисциплины составляет 9 зачетных единиц, 324 часа.

№ п/п	Раздел (тема) дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)						Объем учебной работы, с применением интерактивных методов (в часах / %)	Формы текущего контроля успеваемости (по неделям семестра), форма промежуточной аттестации (по семестрам)	
				Лекции	Практические занятия	Лабораторные работы	Контрольные работы,	СРС	КП / КР			
1.	Введение. Классификация технологий обработки информации.	4	1,2	2		4			6		3 ч/ 50%	
2.	Технологии обработки текстовой информации. Программные средства обработки текстовой информации. Эргономические требования к оформлению текстовых документов. Форматирование текста в текстовом редакторе. Сканирование и распознавание текста.	4	3,4	2		4			6		3 ч/ 50%	
3.	Технологии обработки графической информации. Программные средства обработки графической информации. Технология обработки растровых изображений. Технология обработки векторных изображений. 3-D графика и технологии ее обработки.	4	5,6	2		4			6		3 ч/ 50%	Рейтинг-контроль №1
4.	Технологии обработки числовой информации. Технологии обработки числовой информации средствами электронных таблиц. Режим форматирования электронных таблиц. Режим управления вычислениями. Графический режим.	4	7,8	2		4			6		3 ч/ 50%	
5.	Технологии хранения, поиска и сортировки информации. Методы и алгоритмы архивирования данных. Алгоритмы сортировки данных. Методы фильтрации в процессе поиска информации.	4	9,10	2		4			6		3 ч/ 50%	
6.	Понятие базы данных (БД). Понятие системы управления базами данных (СУБД). Классификация моделей данных, поддерживаемых СУБД. Основные операции над данными, поддерживаемые СУБД.	4	11, 12	2		4			6		3 ч/ 50%	Рейтинг-контроль №2

№ п/п	Раздел (тема) дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)						Объем учебной работы, с применением интерактивных методов (в часах / %)	Формы текущего контроля успеваемости (по неделям семестра), форма промежуточной аттестации (по семестрам)		
				Лекции	Практические занятия	Лабораторные работы	Контрольные работы,	СРС	КП / КР				
7.	Технологии обработки мультимедийной информации. Понятие медиа-серверных систем. Создание медиа данных средствами медиа-серверных систем. Упорядоченное хранение и выборка медиа данных.	4	13, 14	2		4			6		3 ч/ 50%		
8.	Основы статистического анализа данных. Основные характеристики описательной статистики, введение в корреляционный и регрессионный анализ.	4	15, 16	2		4			6		3 ч/ 50%		
9.	Методы и стадии Data Mining. Классификация методов Data Mining. Сравнительная характеристика методов. Стандарты CWM, CRISP, PMML.	4	17, 18	2		4			6		3 ч/ 50%	Рейтинг-контроль №3	
Всего по 4 семестру:						18			36		54	27 /50%	ЗАЧЕТ
10.	Процесс подготовки данных для анализа. Понятия качества данных, грязных данных, этапы очистки данных. Классификация инструментов очистки данных, классификация ошибок в данных в ходе очистки данных.	5	1,2	4		4			11		4 ч/ 50%		
11.	Методы классификации и прогнозирования. Метод деревьев решений, процесс его построения. Методы опорных векторов, "ближайшего соседа" и байесовской классификации.	5	3,4	4		4			11		4 ч/ 50%		
12.	Основы кластерного анализа, математические характеристики кластера. Обзор методов кластеризации. Итеративные методы на примере алгоритма k-means.	5	5,6	4		4			11		4 ч/ 50%	Рейтинг-контроль №1	
13.	Метод поиска ассоциативных правил. Алгоритм Apriori и др. Секвенциальный анализ и его место в методах поиска ассоциативных правил.	5	7,8	4		4			11		4 ч/ 50%		
14.	Методы визуального анализа данных, характеристики средств визуализации данных. Инфографика.	5	9,10	4		4			11		4 ч/ 50%		

№ п/п	Раздел (тема) дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)					Объем учебной работы, с применением интерактивных методов (в часах / %)	Формы текущего контроля успеваемости (по неделям семестра), форма промежуточной аттестации (по семестрам)
				Лекции	Практические занятия	Лабораторные работы	Контрольные работы,	СРС		
15.	Технологии OLAP. Многомерное представление данных и многомерный куб MOLAP. Операции с многомерным кубом ROLAP, схемы «снежинка» и «звезда», HOLAP. Виртуальные хранилища данных. Организация процесса извлечения данных. Уровни очистки данных. Преобразование структур данных. Организация процесса загрузки в хранилище данных. Многопоточная загрузка.	5	11, 12	4		4		11	4 ч/50%	Рейтинг-контроль №2
16.	Введение в web-технологии функционирования систем (приложений). Основные методы представления и обработки данных в web.	5	13, 14	4		4		11	4 ч/50%	
17.	Информационные технология разработки и эксплуатации баз знаний, классификация. Обзор существующих моделей обработки и представления знаний.	5	15, 16	4		4		11	4 ч/50%	
18.	ГИС технологии. Обзор технологий, классификация, архитектурные решения, сферы применения, основные операции над данными в ГИС.	5	17, 18	4		4		11	4 ч/50%	Рейтинг-контроль №3
Всего по 5 семестру:						36		99	36 /50%	ЭКЗАМЕН
ИТОГО:				4,5		54		153	63 ч/50%	Зачет, Экзамен, КР

5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

Изучение дисциплины предполагает не только запоминание и понимание, но и анализ, синтез, рефлексию, формирует универсальные умения и навыки, являющиеся основой становления специалиста по специальности 10.05.04 «Информационно-аналитические системы безопасности».

Для реализации компетентностного подхода предлагается интегрировать в учебный процесс интерактивные образовательные технологии, включая информационные и коммуникационные технологии (ИКТ), при осуществлении различных видов учебной работы:

- учебную дискуссию;
- электронные средства обучения (слайд-лекции, электронные тренажеры, компьютерные тесты);
- дистанционные (сетевые) технологии.

Как традиционные, так и лекции инновационного характера могут сопровождаться компьютерными слайдами или слайд-лекциями. Основное требование к слайд-лекции – применение динамических эффектов (анимированных объектов), функциональным назначением которых является наглядно-образное представление информации, сложной для понимания и осмысления студентами, а также интенсификация и диверсификация учебного процесса.

Удельный вес занятий, проводимых в интерактивных формах, определяется главной целью ОПОП специальности 10.05.04 «Информационно-аналитические системы безопасности», особенностью контингента обучающихся и содержанием конкретных дисциплин, и в целом, в учебном процессе, они составляют не менее 30 процентов аудиторных занятий.

Занятия лекционного типа для соответствующих групп студентов согласно требованиям стандарта высшего образования не могут составлять более 55 процентов аудиторных занятий. Программа дисциплины соответствует данным требованиям.

Таким образом, применение интерактивных образовательных технологий придает инновационный характер практически всем видам учебных занятий, включая лекционные. При этом делается акцент на развитие самостоятельного, продуктивного мышления, основанного на диалогических дидактических приемах, субъектной позиции обучающегося в образовательном процессе. Тем самым создаются условия для реализации компетентностного подхода при изучении данной дисциплины.

6. ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ИТОГАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ И УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Для текущего контроля успеваемости предлагается использование рейтинговой системы оценки, которая носит интегрированный характер и учитывает успешность студента в различных видах учебной деятельности, степень сформированности у студента общекультурных и профессиональных компетенций.

Примерный перечень заданий для текущих контрольных мероприятий:

Вопросы рейтинг-контроля №1 для 4 семестра:

- Технология HTML DOM: расшифровать и дать определение для HTML и HTML DOM, сравнить модели HTML и HTML DOM. Перечислить основные объекты в составе технологии.
- Технология XML: расшифровать и дать определение для XML, сравнить типизированную и нетипизированную модели XML. Перечислить и описать основные объекты в составе XML.

- Технологии управления версиями в процессе разработки информационных систем: назначение, способы организации управления версиями, примеры программного обеспечения для управления версиями.
- Технологии BigData: дать определение для BigData, назначение BigData, история появления и основные принципы BigData. Достоинства и недостатки BigData.
- Технологии управления знаниями, визуализации знаний и интеллектуальные карты. Дать определение понятиям, назначение технологии, привести примеры программного обеспечения для визуализации знаний и построения интеллектуальных карт.
- Технологии ETL: дать определение и расшифровку, область применения. Перечислить основные классы информационных ресурсов, к которым применима технология. Привести примеры программного обеспечения, реализующего ETL.
- ГИС-технологии: дать определение и расшифровку, область применения. Перечислить основные классы информационных ресурсов, к которым применима технология. Привести примеры программного обеспечения, реализующего ГИС.
- Классификация знаний. Способы представления знаний: графы, системы продукций, фреймы, предикаты, семантические сети, онтологии, ситуации.
- Модель данных. Определение. Типы моделей структурированных, неструктурированных, частично структурированных.
- Понятие хранилищ данных: типовая архитектура, особенности, назначение и области применения, примеры систем управления хранилищами данных.
- Хранилище и витрины данных. Противоречивость требований к OLTP-системам и БД для обработки данных

Вопросы рейтинг-контроля №2 для 4 семестра

- Классификация знаний. Способы представления знаний: графы, системы продукций, фреймы, предикаты, семантические сети, онтологии, ситуации.
- Модель данных. Определение. Типы моделей структурированных, неструктурированных, частично структурированных.
- Понятие хранилищ данных: типовая архитектура, особенности, назначение и области применения, примеры систем управления хранилищами данных.
- Хранилище и витрины данных. Противоречивость требований к OLTP-системам и БД для обработки данных
- Многомерное представление и модель данных. Операции с кубом данных.
- Бизнес-аналитика. Примеры инструментария.
- Стандарты Data Mining. Основные типы моделей.
- Инструментальные средства поддержки Data Mining.
- Служба Microsoft Integration Service. Правила создания пакетов.
- Задачи классификации и регрессии.
- Поиск ассоциативных правил. Отличие от задач кластеризации.
- Deskриптивные и предсказательные модели.
- Представление результатов анализа данных.

Вопросы рейтинг-контроля №3 для 4 семестра

- Представление результатов анализа данных.
- Методы и стадии Data Mining.
- Классификация задач Data Mining, их характерные особенности.
- Методы классификации и кластеризации.
- Методы прогнозирования и визуализации данных.
- Этапы развития технологий обработки информации.
- Современные технологии обработки информации.
- Классификация технологий обработки информации.
- Кодирование различных видов информации

- Хранение информационных объектов.
- Поиск информации с использованием компьютера. Программные поисковые сервисы.
- Автоматизированные системы обработки информации.
- Возможности настольных издательских систем: создание, организация и основные способы преобразования (верстки) текста.
- Электронные таблицы.

Перечень вопросов к зачету 4 семестра (промежуточной аттестации по итогам освоения дисциплины):

1. Технология HTML DOM: расшифровать и дать определение для HTML и HTML DOM, сравнить модели HTML и HTML DOM. Перечислить основные объекты в составе технологии.
2. Технология XML: расшифровать и дать определение для XML, сравнить типизированную и нетипизированную модели XML. Перечислить и описать основные объекты в составе XML.
3. Технологии управления версиями в процессе разработки информационных систем: назначение, способы организации управления версиями, примеры программного обеспечения для управления версиями.
4. Технологии BigData: дать определение для BigData, назначение BigData, история появления и основные принципы BigData. Достоинства и недостатки BigData.
5. Технологии управления знаниями, визуализации знаний и интеллектуальные карты. Дать определение понятиям, назначение технологии, привести примеры программного обеспечения для визуализации знаний и построения интеллектуальных карт.
6. Технологии ETL: дать определение и расшифровку, область применения. Перечислить основные классы информационных ресурсов, к которым применима технология. Привести примеры программного обеспечения, реализующего ETL.
7. ГИС-технологии: дать определение и расшифровку, область применения. Перечислить основные классы информационных ресурсов, к которым применима технология. Привести примеры программного обеспечения, реализующего ГИС.
8. Классификация знаний. Способы представления знаний: графы, системы продукций, фреймы, предикаты, семантические сети, онтологии, ситуации.
9. Модель данных. Определение. Типы моделей структурированных, неструктурированных, частично структурированных.
10. Понятие хранилищ данных: типовая архитектура, особенности, назначение и области применения, примеры систем управления хранилищами данных.
11. Хранилище и витрины данных. Противоречивость требований к OLTP-системам и БД для обработки данных
12. Многомерное представление и модель данных. Операции с кубом данных.
13. Бизнес-аналитика. Примеры инструментария.
14. Стандарты Data Mining. Основные типы моделей.
15. Инструментальные средства поддержки Data Mining.
16. Служба Microsoft Integration Service. Правила создания пакетов.
17. Задачи классификации и регрессии.
18. Поиск ассоциативных правил. Отличие от задач кластеризации.
19. Deskриптивные и предсказательные модели.
20. Представление результатов анализа данных.
21. Методы и стадии Data Mining.
22. Классификация задач Data Mining, их характерные особенности.
23. Методы классификации и кластеризации.
24. Методы прогнозирования и визуализации данных.
25. Этапы развития технологий обработки информации.
26. Современные технологии обработки информации.
27. Классификация технологий обработки информации.
28. Кодирование различных видов информации

29. Хранение информационных объектов.
30. Поиск информации с использованием компьютера. Программные поисковые сервисы.
31. Автоматизированные системы обработки информации.
32. Возможности настольных издательских систем: создание, организация и основные способы преобразования (верстки) текста.
33. Электронные таблицы.

Вопросы рейтинг-контроля №1 для 5 семестра

- Возможности динамических (электронных) таблиц. Математическая обработка числовых данных.
- Обработка и форматирование данных в электронных таблицах.
- Программные средства компьютерной графики
- Растровая, векторная и фрактальная графика. 3D графика.
- Структура интернет-страницы.
- Способы создания и сопровождения сайта с использованием различных сервисов.
- Создание, редактирование и форматирование текстовых документов в текстовых редакторах.
- Работа с колонтитулами, таблицами, списками и объектами в текстовых редакторах.
- Работа со списками и логическими функциями в текстовых редакторах.
- Возможности динамических (электронных) таблиц. Математическая обработка числовых данных.
- Обработка и форматирование данных в электронных таблицах.
- Классификация баз данных.
- Организация баз данных и системах управления базами данных.
- Медиа-серверные системы.

Вопросы рейтинг-контроля №2 для 5 семестра

- Классификация баз данных.
- Организация баз данных и системах управления базами данных.
- Медиа-серверные системы.
- Создание медиа данных средствами медиа-серверных систем.
- Упорядоченное хранение и выборка медиа данных.
- Классификация источников информации.
- Данные как вид информации. Формы данных. Типы данных.
- Модели структурированных данных.
- Неструктурированные и слабоструктурированные данные.
- Основные компоненты технологии обработки информации. Информационные процессы обработки информации.
- Системы сбора и обработки информации. Виды систем сбора информации.

Вопросы рейтинг-контроля №3 для 5 семестра

- Системы сбора и обработки информации. Виды систем сбора информации.
- Общие визуализаторы: графики, диаграммы, гистограммы, статистика.
- OLAP- анализ, манипуляции с OLAP- кубами. Матрицы классификации, диаграммы рассеяния.
- Регрессионный анализ, ретропрогноз.
- Детализованные и агрегированные данные, метаданные. Многомерное представление данных и многомерный куб MOLAP.
- Операции с многомерным кубом ROLAP, схемы «снежинка» и «звезда».
- Оценка качества данных по происхождению, профайлинг данных. Визуальная оценка качества данных.

- Виды аномалий данных. Обнаружение аномальных значений специальными методами. Происхождение пропусков данных, восстановление пропущенных значений.
- Централизованная и распределенная обработка данных. Сравнительная характеристика.
- Обработка двумерных сигналов (изображений), обработка графической информации.
- Обработка информации с целью получения знаний. Задача ассоциации, кластеризации, классификация и регрессия, статистические методы, машинное обучение.
- Системы обработки входящей текстовой информации, методы поиска текстовой информации. Качество информационно-поисковых систем.

Перечень вопросов к экзамену 5 семестр (промежуточной аттестации по итогам освоения дисциплины):

1. Возможности динамических (электронных) таблиц. Математическая обработка числовых данных.
2. Обработка и форматирование данных в электронных таблицах.
3. Программные средства компьютерной графики
4. Растровая, векторная и фрактальная графика. 3D графика.
5. Структура интернет-страницы.
6. Способы создания и сопровождения сайта с использованием различных сервисов.
7. Создание, редактирование и форматирование текстовых документов в текстовых редакторах.
8. Работа с колоннитулами, таблицами, списками и объектами в текстовых редакторах.
9. Работа со списками и логическими функциями в текстовых редакторах.
10. Возможности динамических (электронных) таблиц. Математическая обработка числовых данных.
11. Обработка и форматирование данных в электронных таблицах.
12. Классификация баз данных.
13. Организация баз данных и системах управления базами данных.
14. Медиа-серверные системы.
15. Создание медиа данных средствами медиа-серверных систем.
16. Упорядоченное хранение и выборка медиа данных.
17. Классификация источников информации.
18. Данные как вид информации. Формы данных. Типы данных.
19. Модели структурированных данных.
20. Неструктурированные и слабоструктурированные данные.
21. Основные компоненты технологии обработки информации. Информационные процессы обработки информации.
22. Системы сбора и обработки информации. Виды систем сбора информации.
23. Общие визуализаторы: графики, диаграммы, гистограммы, статистика.
24. OLAP- анализ, манипуляции с OLAP- кубами. Матрицы классификации, диаграммы рассеяния.
25. Регрессионный анализ, ретропрогноз.
26. Детализированные и агрегированные данные, метаданные. Многомерное представление данных и многомерный куб MOLAP.
27. Операции с многомерным кубом ROLAP, схемы «снежинка» и «звезда».
28. Оценка качества данных по происхождению, профайлинг данных. Визуальная оценка качества данных.
29. Виды аномалий данных. Обнаружение аномальных значений специальными методами. Происхождение пропусков данных, восстановление пропущенных значений.
30. Централизованная и распределенная обработка данных. Сравнительная характеристика.
31. Обработка двумерных сигналов (изображений), обработка графической информации.
32. Обработка информации с целью получения знаний. Задача ассоциации, кластеризации, классификация и регрессия, статистические методы, машинное обучение.
33. Системы обработки входящей текстовой информации, методы поиска текстовой информации. Качество информационно-поисковых систем.

Перечень лабораторных работ:

Темы лабораторных работ соответствуют темам, указанным в таблице, отражающей структуру и содержание дисциплины.

Перечень лабораторных работ 4 семестр

- Технологии обработки текстовой информации. Программные средства обработки текстовой информации. Сканирование и распознавание текста.
- Технологии обработки графической информации. Программные средства обработки графической информации. Технология обработки растровых изображений. Технология обработки векторных изображений. 3-D графика и технологии ее обработки.
- Технологии обработки числовой информации. Технологии обработки числовой информации средствами электронных таблиц. Режим форматирования электронных таблиц. Режим управления вычислениями. Графический режим.
- Технологии хранения, поиска и сортировки информации. Методы и алгоритмы архивирования данных. Алгоритмы сортировки данных. Методы фильтрации в процессе поиска информации.
- Понятие базы данных (БД). Понятие системы управления базами данных (СУБД). Классификация моделей данных, поддерживаемых СУБД. Основные операции над данными, поддерживаемые СУБД.
- Технологии обработки мультимедийной информации. Понятие медиа-серверных систем. Создание медиа данных средствами медиа-серверных систем. Упорядоченное хранение и выборка медиа данных.
- Основы статистического анализа данных. Основные характеристики описательной статистики, введение в корреляционный и регрессионный анализ.
- Методы и стадии Data Mining. Классификация методов Data Mining. Сравнительная характеристика методов. Стандарты CWM, CRISP, PMML

Перечень лабораторных работ 5 семестр

- Процесс подготовки данных для анализа. Понятия качества данных, грязных данных, этапы очистки данных. Классификация инструментов очистки данных, классификация ошибок в данных в ходе очистки данных.
- Методы классификации и прогнозирования. Метод деревьев решений, процесс его построения. Методы опорных векторов, "ближайшего соседа" и байесовской классификации.
- Основы кластерного анализа, математические характеристики кластера. Обзор методов кластеризации. Итеративные методы на примере алгоритма k-means.
- Метод поиска ассоциативных правил. Алгоритм Apriori и др. Секвенциальный анализ и его место в методах поиска ассоциативных правил.
- Методы визуального анализа данных, характеристики средств визуализации данных. Инфографика.
- Технологии OLAP. Многомерное представление данных и многомерный куб MOLAP. Операции с многомерным кубом ROLAP, схемы «снежинка» и «звезда», HОLAP. Виртуальные хранилища данных. Организация процесса извлечения данных. Уровни очистки данных. Преобразование структур данных. Организация процесса загрузки в хранилище данных. Многопоточная загрузка.
- Введение в web-технологии функционирования систем (приложений). Основные методы представления и обработки данных в web.
- Информационные технология разработки и эксплуатации баз знаний, классификация. Обзор существующих моделей обработки и представления знаний.
- ГИС технологии. Обзор технологий, классификация, архитектурные решения, сферы применения, основные операции над данными в ГИС.

Единое задание на лабораторные работы. Выполнить сравнительный анализ типов инструментальных средств, в которых используется изучаемая технология.

Состав отчета по лабораторной работе:

- титульный лист;

- задание на лабораторную работу;
- классификация типов средств, в которых используется рассматриваемая технология;
- выбор критериев для сравнения инструментальных средств;
- сравнительная таблица инструментальных средств;
- пример практического использования изучаемой технологии на базе одного из выбранных инструментальных средств.

Вопросы и задания для самостоятельной работы студентов 4 семестр:

1. Информационные технологии обработки данных в интернет-магазинах.
2. Информационные технологии обработки данных платежей через Интернет.
3. Информационные технологии обработки данных на основе Google-технологий.
4. Информационные технологии обработки данных на основе Yandex-технологий.
5. Информационные технологии удаленной работы с данными.
6. Информационные технологии интеграции данных.
7. Современные технологии электронных таблиц.
8. Модели представления данных, СУБД и принципы обработки данных.
9. Форматы представления звуковых и видеофайлов.
10. Искусственный интеллект и классификация решаемых им задач, методы представления данных в таких задачах.
11. Знания и данные. Модели представления знаний и методы обработки знаний.
12. Организация данных в технологии параллельных вычислений.
13. Самоорганизующиеся карты Кохонена.
14. Нейронные сети и их применение для задач Data Mining.
15. Виды OLAP – MOLAP, ROLAP, HOLAP.
16. Основные требования к многомерному хранилищу данных.
17. Обзор современных программных средств сканирования и распознавания текстов.
18. Свободно распространяемое программное обеспечение для обработки текстовой информации.

Вопросы и задания для самостоятельной работы студентов 4 семестр:

1. Свободно распространяемое программное обеспечение для обработки графической информации.
2. Свободно распространяемое программное обеспечение для обработки медиа данных.
3. Этапы развития технологий хранения информации.
4. Анализ временных рядов. Трансформация временных рядов: скользящее окно, интервал и горизонт прогноза и др.
5. Цифровая обработка цифровых изображений: двумерное дискретное преобразование Фурье, получение спектра изображения.
6. Цифровая обработка цифровых изображений: гистограммные методы преобразования яркости изображения.
7. Фильтрация двумерных дискретных сигналов: низкочастотная и высокочастотная фильтрация изображений.

Для курсовой работы 5 семестр предлагаются следующие тематики:

По заданию преподавателя на базе полученного набора данных разработать приложение, реализующие основные операции обработки данных.

В целом, должно быть получено законченное приложение, в котором обязательны к реализации следующие функции:

- 1) форма регистрации/аутентификации пользователя.
- 2) меню приложения, которое должно дублироваться;
- 3) форма ввода данных;
- 4) форма вывода данных;
- 5) форма с поисковыми запросами;
- 6) форма статистической обработки данных;

- 7) форма с инфографикой;
- 8) форма со справочными данными с возможностями ввода/редактирования/удаления/сохранения.
- 9) форма подготовки данных на печать.

Примерное содержание пояснительной записки к курсовой работе и вопросы для проработки:

- Задание на курсовой проект/работу.
- Оглавление.
- Список сокращений и обозначений
- Реферат (аннотация).
- Введение.
- 1 Анализ состояния вопроса, и постановка задач исследования (образ будущего продукта/разрабатываемого приложения, описываются общие современные потребности в системах вашего типа, исходя из темы; обзор существующих аналогов выполняется с анализом функционала аналогов или классов аналогов).
- 2 Выбор и обоснование технологий для проектирования и реализации приложений (сравнительный анализ технологий, которые могут быть использованы при разработке данного проекта; обоснование выбора технологий для проекта; обоснование выбора методологии и CASE-средств для разработки проекта на этапе проектирования).
- 3 Проектирование приложения на базе одной из методологий (RUP, ARIS, Agile, MSF с применением UML и др.).
- 4 Реализация системы (реализация выявленных требований).
- 5 Оценка реализованных решений (оценка разработанной системы на соответствие требованиям).
- Заключение.
- Список использованных источников.
- Приложения.

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

а) Основная литература:

1. Структуры и алгоритмы обработки данных: Учебное пособие / В.Д. Колдаев. - М.: ИЦ РИОР: НИЦ ИНФРА-М, 2014. - 296 с.: 60x90 1/16. - (Высшее образование: Бакалавриат). (переплет) ISBN 978-5-369-01264-2, 500 экз. Режим доступа: <http://znanium.com/catalog.php?bookinfo=418290>
2. Базовые и прикладные информационные технологии: Учебник / В.А. Гвоздева. - М.: ИД ФОРУМ: НИЦ ИНФРА-М, 2014. - 384 с.: 60x90 1/16. - (Высшее образование). (переплет) ISBN 978-5-8199-0572-2, 500 экз. Режим доступа: <http://znanium.com/catalog.php?bookinfo=428860>
3. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов и др. - М.: НИЦ ИНФРА-М, 2015. - 890 с.: 60x90 1/16 ISBN 978-5-16-103267-1 (online) Режим доступа: <http://znanium.com/catalog.php?bookinfo=515227>
4. Прикладные методы анализа статистических данных [Электронный ресурс]: учеб. пособие / Горяинова Е.Р., Панков А.Р., Платонов Е.Н. - М. : ИД Высшей школы экономики, 2012. - <http://www.studentlibrary.ru/book/ISBN9785759808664.html> Электронное издание на основе: Прикладные методы анализа статистических данных [Текст] : учеб. пособие / Е.Р. Горяинова, А.Р. Панков, Е.Н. Платонов ; Нац. исслед. ун-т "Высшая школа экономики". - М.: Изд. дом Высшей школы экономики, 2012. - 310, [2] с. - 1000 экз. - ISBN 978-5-7598-0866-4.

б) Дополнительная литература:

1. Богданова, С.В. Информационные технологии: учебное пособие / С.В. Богданова, А.Н. Ермакова. - Ставрополь: Сервисшкола, 2014. - 211 с. - Режим доступа: <http://znanium.com/catalog.php?bookinfo=514867> Режим доступа: <http://znanium.com/>
2. Петрунин Ю.Ю. Информационные технологии анализа данных. Data analysis: учебное пособие - 2-е изд.- КДУ, 2010.- 293 с. ISBN:978-5-98227-701-5.- Режим доступа: <https://vlsu.bibliotech.ru>
3. Батоврин В. К. Системная и программная инженерия. Словарь-справочник: учеб. пособие для вузов.- ДМК Пресс, 2010.- 281 с. ISBN:978-5-94074-592-1.- Режим доступа: <https://vlsu.bibliotech.ru>
4. Левчук, Е.А. Технологии организации, хранения и обработки данных [Электронный ресурс] : учеб. пособие / Е.А. Левчук. - 3-е изд. - Минск: Выш. шк., 2007. - 239 с.: ил. ISBN 978-985-06-1409-4. Режим доступа: <http://znanium.com/>
5. Основы эксплуатации автоматизированных систем обработки информации и управления. Краткий курс [Электронный ресурс] : Учеб. пособие / В. М. Постников. - М. : Издательство МГТУ им. Н. Э. Баумана, 2013. - <http://www.studentlibrary.ru/book/ISBN9785703836552.html> Электронное издание на основе: Постников В.М. Основы эксплуатации автоматизированных систем обработки информации и управления. Краткий курс : учеб. пособие / В. М. Постников. - М.: Изд-во МГТУ им. Н. Э. Баумана, 2013. - 177, [2] с.: ил. ISBN 978-5-7038-3655-2.
6. Тезаурусы в задачах информационного поиска [Электронный ресурс] / Лукашевич Н.В. - М. : Издательство Московского государственного университета, 2011. - <http://www.studentlibrary.ru/book/ISBN9785211059269.html> Электронное издание на основе: Тезаурусы в задачах информационного поиска. - М.: Издательство Московского университета, 2011. - 512 с., ил. - ISBN 978-5-211-05926-9.
7. Статистические методы обработки экспериментальных данных с использованием пакета MathCad: Учебное пособие/Ф.И.Карманов, В.А.Острейковский - М.: КУРС, НИЦ ИНФРА-М, 2015. - 208 с.: 60x90 1/16 (Обложка) ISBN 978-5-905554-96-4, 100 экз. Режим доступа: <http://znanium.com/>
8. Барсегян, А. А. Технологии анализа данных: Data Mining, Visual Mining, Text Mining, OLAP [Электронный ресурс] / А. А. Барсегян, М. С. Куприянов, В. В. Степаненко. - 2-е изд.,

перераб. и доп. - СПб.: БХВ-Петербург, 2007. - 384 с.: ил. + CD-ROM - ISBN 5-94157-991-8.
Режим доступа: <http://znanium.com/>

9. Методы многомерного анализа статистических данных [Электронный ресурс] : учеб. пособие/ В.М. Симчера. - М. : Финансы и статистика, 2008. - <http://www.studentlibrary.ru/book/ISBN9785279031849.html> Электронное издание на основе: Методы многомерного анализа статистических данных: учеб. пособие. - М.: Финансы и статистика, 2008. - 400 с.: ил. - ISBN 978-5-279-03184-9.

в) Периодические издания:

1. Ежемесячный теоретический и прикладной научно-технический журнал «Информационные технологии». Режим доступа <http://novtex.ru/IT/>;

2. Журнал «Информационные технологии и вычислительные системы». Журнал выпускается при научно-методическом руководстве Отделения нанотехнологий и информационных технологий Российской академии наук и поддержке Российской ассоциации искусственного интеллекта. ISSN 2071-8632. Режим доступа

http://www.jitcs.ru/index.php?option=com_content&view=article&id

г) Программное обеспечение и Интернет-ресурсы:

1. Внутривузовские издания ВлГУ.— Режим доступа: <http://e.lib.vlsu.ru/>

2. ИНТУИТ. Национальный открытый университет.— Режим доступа: <http://www.intuit.ru/>

3. Чубукова И. Data Mining.— ИНТУИТ. Национальный открытый университет.— Режим доступа: <http://www.intuit.ru/studies/courses/6/6/info>

4. Академия BIG DATA: Введение в аналитику больших массивов данных.— ИНТУИТ. Национальный открытый университет/ Ю. Аникин, И. Борисова, Н. Загоруйко, А. Зырянов, Е. Павловский.— Режим доступа: <http://www.intuit.ru/studies/courses/12385/1181/info>

5. Microsoft Office, расширение Data Mining для Microsoft Excel.

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

ауд. 408-2, Лекционная аудитория, количество студенческих мест – 50, площадь 60 м², оснащение: мультимедийное оборудование (интерактивная доска Hitachi FX-77WD, проектор BenQ MX 503 DLP 2700ANSI XGA), ноутбук Lenovo Idea Pad B5045

ауд. 427а-2, лаборатория сетевых технологий, количество студенческих мест – 14, площадь 36 м², оснащение: компьютерный класс с 8 рабочими станциями Core 2 Duo E8400 с выходом в Internet, 3 маршрутизатора Cisco 2800 Series, 6 маршрутизаторов Cisco 2621, 6 коммутаторов Cisco Catalyst 2960 Series, 3 коммутатора Cisco Catalyst 2950 Series, коммутатор Cisco Catalyst Express 500 Series, проектор BenQ MP 620 P, экран настенный рулонный. Лицензионное программное обеспечение: операционная система Windows 7 Профессиональная, офисный пакет приложений Microsoft Office Профессиональный плюс 2007, бесплатно распространяемое программное обеспечение: линейка интегрированных сред разработки Visual Studio Express 2012, программный продукт виртуализации Oracle VM VirtualBox 5.0.4, симулятор сети передачи данных Cisco Packet Tracer 7.0, интегрированная среда разработки программного обеспечения IntelliJ IDEA Community Edition 15.0.3.

ауд. 427б-2, УНЦ «Комплексная защита объектов информатизации», количество студенческих мест – 15, площадь 52 м², оснащение: компьютерный класс с 7 рабочими станциями Alliance Optima P4 с выходом в Internet, коммутатор D-Link DGS-1100-16 мультимедийный комплект (проектор Toshiba TLP X200, экран настенный рулонный), прибор ST-031P «Пирания-Р» многофункциональный поисковый, прибор «Улан-2» поисковый, виброакустический генератор шума «Соната АВ 1М», имитатор работы средств нелегального съема информации, работающих по радиоканалу «Шиповник», анализатор спектра «GoodWill GSP-827», индикатор поля «SEL SP-75 Black Hunter», устройство блокирования работы систем мобильной связи «Мозайка-3», устройство защиты телефонных переговоров от прослушивания «Прокруст 2000», диктофон Edic MINI Hunter, локатор «Родник-2К» нелинейный, комплекс проведения акустических и виброакустических измерений «Спрут мини-А», видеорегистратор цифровой Best DVR-405, генератор Шума «Гном-3», учебно-исследовательский комплекс «Сверхширокополосные беспроводные сенсорные сети» (Nano Chaos), сканирующий приемник «Icom IC-R1500», анализатор сетей Wi-Fi Fluke AirCheck с активной антенной. Лицензионное программное обеспечение: Windows 8 Профессиональная, офисный пакет приложений Microsoft Office Профессиональный плюс 2010, бесплатно распространяемое программное обеспечение: линейка интегрированных сред разработки Visual Studio Express 2012, инструмент имитационного моделирования AnyLogic 7.2.0 Personal Learning Edition, интегрированная среда разработки программного обеспечения IntelliJ IDEA Community Edition 14.1.4.

Программа составлена в соответствии с требованиями ФГОС ВО по специальности 10.05.04 "Информационно-аналитические системы безопасности", специализация «автоматизация информационно-аналитической деятельности»

Рабочую программу составил: к.т.н., доцент, доцент каф. ИЗИ Семенова И.И. 

Рецензент

(представитель работодателя) Заместитель руководителя РАИ ООО «ИнфоЦентр»

к.т.н. Вертилевский Н.В. 

(место работы, должность, ФИО, подпись)

Программа рассмотрена и одобрена на заседании кафедры ИЗИ

Протокол № 7 от 28.12.2016 года

Заведующий кафедрой д.т.н., профессор

 /М.Ю. Монахов/

Рабочая программа рассмотрена и одобрена на заседании учебно-методической комиссии специальности 10.05.04 "Информационно-аналитические системы безопасности", специализация «автоматизация информационно-аналитической деятельности»

Протокол № 4 от 23.12.2016 года

Председатель комиссии д.т.н., профессор

 /М.Ю. Монахов/

ЛИСТ ПЕРЕУТВЕРЖДЕНИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ (МОДУЛЯ)

Рабочая программа одобрена на 2017/2018 учебный год

Протокол заседания кафедры № 1 от 28.08.17 года

Заведующий кафедрой д.т.н., профессор

 /М.Ю. Монахов/

(ФИО, подпись)

ЛИСТ ПЕРЕУТВЕРЖДЕНИЯ РАБОЧЕЙ ПРОГРАММЫ ДИСЦИПЛИНЫ (МОДУЛЯ)

Рабочая программа одобрена на _____ учебный год

Протокол заседания кафедры № _____ от _____ года

Заведующий кафедрой д.т.н., профессор

/М.Ю. Монахов/

(ФИО, подпись)

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых»
(ВлГУ)

Институт _____

Кафедра _____

Актуализированная
рабочая программа
рассмотрена и одобрена
на заседании кафедры
протокол № ____ от ____ 20__ г.

Заведующий кафедрой

(подпись, ФИО)

Актуализация рабочей программы дисциплины

(наименование дисциплины)

Специальность

Специализация

Уровень высшего образования

Форма обучения

Владимир 20__

Рабочая программа учебной дисциплины актуализирована в части рекомендуемой литературы.

Актуализация выполнена: _____
(подпись, должность, ФИО)

а) основная литература: _____

б) дополнительная литература: _____

в) периодические издания: _____

в) интернет-ресурсы: _____