

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ «МЭИ»

«УТВЕРЖДАЮ»

Проректор по научной работе

 Драгунов В.К.

« 16 » июня 2015 г.



Программа аспирантуры

Направление 09.06.01 Информатика и вычислительная техника

Направленность (специальность) 05.13.11 Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей

РАБОЧАЯ ПРОГРАММА

дисциплины по выбору

«Организация баз данных и знаний»

Индекс дисциплины по учебному плану: Б1.В.ДВ.1.1

Всего: 108 часов

Семестр 1, в том числе

6 часов – контактная работа,
84 часа – самостоятельная работа,
18 часов – контроль

Программа составлена на основе федерального государственного образовательного стандарта высшего образования (уровень подготовки кадров высшей квалификации) по направлению подготовки 09.06.01 Информатика и вычислительная техника, утвержденного приказом Минобрнауки России от 30 июля 2014 г. № 875, и паспорта специальности, указанной в номенклатуре специальностей научных работников 05.13.11 Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей, утвержденной приказом Минобрнауки России от 25 февраля 2009 г. № 59.

ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью изучения дисциплины является освоение основных подходов к проектированию, реализации, внедрению и эксплуатации баз данных (БД) и знаний (БЗ), соответствующих последним результатам исследований в этой области.

Задачами дисциплины являются:

- изучение подходов и инструментальных средств выполнения всех этапов жизненного цикла БД и БЗ;
- изучение особенностей разработки БД и БЗ различной архитектуры;
- изучение новых подходов и перспективных решений по БД и БЗ.

В процессе освоения дисциплины **формируются следующие компетенции:**

- способность проектировать и осуществлять комплексные исследования, в том числе междисциплинарные, на основе целостного системного научного мировоззрения с использованием знаний в области истории и философии науки (УК-2).
- владение методологией теоретических и экспериментальных исследований в области профессиональной деятельности (ОПК-1).

– способность к разработке новых методов исследования и их применению в самостоятельной научно-исследовательской деятельности в области профессиональной деятельности (ОПК-3).

– способность представлять полученные результаты научно-исследовательской деятельности на высоком уровне и с учетом соблюдения авторских прав (ОПК-6).

– способность разрабатывать эффективные математические методы решения задач естествознания, техники, экономики и управления (ПК-1).

– владение языками и системами программирования, системами управления базами данных и знаний, операционными системами и программными инструментами для организации взаимодействия программ и программных систем (ПК-4).

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБРАЗОВАНИЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ

В результате освоения дисциплины обучающийся должен продемонстрировать **следующие результаты образования:**

знать:

– модели данных, их свойства, преимущества и недостатки, области применения (УК-2);

– методы и средства проектирования и реализации баз данных различной архитектуры (ОПК-1);

– методы и средства аналитической обработки данных (ОПК-3);

– принципы построения сервис-ориентированных систем и средств их реализации (ПК-1);

уметь:

– работать на СУБД Microsoft SQL Server, создать серверную часть баз данных (ПК-4);

– работать в среде Microsoft Visual Studio для разработки клиентской части (ПК-4);

владеть:

- методикой проведения системного анализа с использованием CASE-средств (ОПК-6);
- методикой проектирования компонентов баз данных различной архитектуры (ПК-4);
- методикой внедрения и сопровождения баз данных различной архитектуры (ПК-4).

КРАТКОЕ СОДЕРЖАНИЕ РАЗДЕЛОВ ДИСЦИПЛИНЫ

Введение. Основные понятия

(6 часов самостоятельной работы)

Основные понятия: база данных, система управления базами данных, администратор данных, администратор базы данных. информационные системы. Классификация баз данных. Краткий обзор наиболее распространенных СУБД.

Основы теории реляционных баз данных

(12 часов самостоятельной работы)

Определение реляционной модели, реляционное исчисление и реляционная алгебра. Основные и дополнительные операции на реляционной модели. Функциональные и многозначные зависимости; ключи. Нормальные формы и нормализация, аномалии. Свойства нормализации (соединение без потерь и сохранение зависимостей), проверка их наличия. Декомпозиция отношений, алгоритмы декомпозиции. Преимущества и недостатки нормализации.

Проектирование баз данных

(12 часов самостоятельной работы)

Жизненный цикл базы данных, основные этапы и модели. Анализ предметной области для проектирования баз данных. Проектирование на основе данных и на основе решаемых задач. Этапы проектирования: даталогическое, логическое и физическое проектирование. Использование

объектно-ориентированного анализа и проектирования, языка UML для проектирования баз данных. Диаграммы языка UML и их применение при проектировании баз данных. Переход от объектно-ориентированной модели к реляционной. Специальные CASE-средства проектирования баз данных.

Язык SQL: средства описания, поиска и обновления данных. Алгоритмические средства язык T-SQL.

Базы данных архитектуры «Клиент-сервер»

(12 часов самостоятельной работы)

Проектирование баз данных архитектуры «клиент-сервер». Создание серверной части. Реализация структуры баз данных на СУБД Microsoft SQL SERVER (таблиц и отношений между ними). Создание представлений, хранимых процедур, функций и триггеров. Решение вопросов, связанных с одновременным доступом пользователей.

Разработка клиентской части в среде Microsoft Visual Studio. «толстый» и «тонкий» клиент. Разработка многооконных интерфейсов.

Язык LINQ и его использование при работе с базами данных.

Сервис-ориентированные системы

(10 часов самостоятельной работы)

Понятие распределенной базы данных. 12 принципов распределенных баз данных. Фрагментация и репликация в распределенных базах данных.

Язык XML и его использование при построении систем обработки данных. Представление реляционной модели средствами XML. Обработка представленных на XML данных алгоритмическим языком C#.

Язык LINQ и его использование для обработки данных, представленных на XML.

Понятие сервис – ориентированной архитектуры (SOA). Принципы разработки и реализации информационных систем согласно SOA. Основные проблемы при их создании и пути их решения. Средства реализации систем SOA.

Хранилище данных и аналитическая обработка данных

(10 часов самостоятельной работы)

Отличия хранилищ данных и баз данных: OLTP и OLAP системы, их особенности. Назначение и характеристики OLAP-систем. Многомерный куб как основа построения систем аналитической обработки. Построение многомерных кубов. Классические задачи анализа данных: классификация, регрессия, поиск ассоциативных связей и кластеризация.

Аналитические средства СУБД Microsoft SQL Server.

Нереляционные базы данных

(10 часов самостоятельной работы)

Ограниченность реляционной модели. Объектные базы данных, их преимущества и недостатки. СУБД CACHE. Базы данных NoSQL (Not Only SQL). Тенденции развития баз данных.

От баз данных к базам знаний. Принципы обработки знаний

(12 часов самостоятельной работы)

Отличия данных и знаний, от обработки данных к обработке знаний. Методы и модели представления и оперирования знаниями, представление поверхностных (простых) и глубинных (структурированных) знаний. Принципы организации баз знаний. Использование баз знаний и методов обработки знаний для аналитической обработки. Data Mining: извлечение знаний из данных, задачи и методы решения. Использование методов искусственного интеллекта для анализа данных.

ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ РЕЗУЛЬТАТОВ ОБРАЗОВАНИЯ ПО ДИСЦИПЛИНЕ

Промежуточная аттестация по итогам освоения дисциплины:

1 семестр – дифференцированный зачет.

Вопросы для самоконтроля

1. Что такое «база данных»?
2. Что такое «система управления базами данных – СУБД»?
3. Что может быть инвариантным относительно предметной области база данных и/или СУБД?
4. Какие функции выполняет администратор данных?
5. Какие функции выполняет администратор базы данных?
6. Из каких компонентов состоит информационная система?
7. В чем суть реляционной модели данных? Что такое отношение, кортеж?
8. Реляционная алгебра, основные понятия.
9. Реляционное исчисление, основные понятия.
10. Базовые и дополнительные операции реляционной алгебры и реляционного исчисления.
11. Какую роль играют в реляционной модели функциональные зависимости?
12. Что такое ключ (простой и составной ключ)?
13. Как связаны ключ и функциональные зависимости?
14. В чем отличие функциональной и многозначной зависимостей?
15. Какие виды аномалии существуют? В чем их вред?
16. Дайте определения нормальных форм.
17. С какой целью выполняется нормализация?
18. В чем суть свойства соединения без потерь?
19. В чем суть свойства сохранения зависимостей?
20. В чем заключаются преимущества и недостатки нормализации?
21. Какие существуют подходы к построению баз данных?

22. Какие задачи выполняют на этапе даталогического проектирования?
23. Какие задачи выполняют на этапе логического проектирования?
24. Какие задачи выполняют на этапе физического проектирования?
25. Какие из названных этапов могут быть выполнены независимо от среды реализации?
26. Какие диаграммы UML могут быть применены при проектировании баз данных? Для решения каких задач проектирования?
27. Как описать структуру базы данных на языке SQL?
28. Как функционирует база данных архитектуры «клиент-сервер»?
29. Что такое «толстый клиент» и «тонкий клиент»?
30. Для решения каких задач применим язык LINQ?
31. В чем особенности распределенных баз данных?
32. В чем суть операции фрагментации?
33. В чем суть операции репликации?
34. В каких случаях требуется применение только фрагментации? Только репликации? Обеих операций?
35. В чем заключаются принципы функционирования сервис-ориентированной архитектуры?
36. В чем преимущества и недостатки сервис-ориентированной архитектуры по сравнению с архитектурой «клиент-сервер»?
37. Какова структура языка XML?
38. Как могут быть представлены на XML данные из реляционной модели?
39. В чем отличие баз данных и хранилищ данных?
40. Какие наиболее существенные отличия систем OLTP и OLAP?
41. Какие Вы знаете классические задачи OLAP, в чем их суть?
42. Какие средства аналитической обработки имеются в Microsoft SQL Server?
43. Какие Вы знаете постреляционные модели данных?

44. Какие основные отличия между реляционными и объектными базами данных?
45. Что понимается под термином NoSQL?
46. В чем основные отличия между данными и знаниями?
47. Чем отличается обработка данных от обработки знаний?
48. Что такое Data Mining?
49. Какие требования предъявляют к результатам Data Mining?
50. В чем заключается интеллектуальный анализ данных?

Вопросы для проведения зачета

1. Основные понятия, связанные с базами данных.
2. Реляционная модель данных.
3. Основные и дополнительные операции на реляционной модели данных.
4. Функциональные зависимости и их роль в реляционной модели данных?
5. Нормальные формы и нормализация.
6. Декомпозиция как средство нормализации.
7. Свойства декомпозиции.
8. Алгоритмы декомпозиции.
9. Преимущества и недостатки декомпозиции.
10. Жизненный цикл базы данных.
11. Этапы проектирования баз данных, их содержание.
12. Средства проектирования баз данных.
13. Средства описания структуры данных языка SQL.
14. Средства поиска данных на языке SQL.
15. Средства изменения данных на языке SQL.
16. Представления: их назначение и создание.
17. Алгоритмические средства языка T- SQL.
18. Архитектура «клиент-сервер»
19. Хранимые процедуры и функции, их назначение и создание.
20. Триггеры: их назначение и создание.

- 21.Создание клиентской части, вызов хранимых процедур и функций.
- 22.Принципы построения сервис-ориентированных систем.
- 23.Использование языка XML для обработки данных.
- 24.Язык LINQ и его использование при обработке данных.
- 25.Отличие баз данных и хранилищ данных. Сравнение OLAP и OLTP.
- 26.Многомерный куб как основа аналитической обработки данных.
- 27.Аналитические средства в современных СУБД.
- 28.Отличие данных и знаний.
- 29.Обработка знаний, ее отличие от обработки данных.
- 30.Применение методов искусственного интеллекта для анализа данных.

Критерии оценки за освоение дисциплины определены в Инструктивном письме И-23 от 14 мая 2012 г.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

Основная литература:

1. Советов Б.Я., Цехановский В.В., Чертовской В.Д. Базы данных: теория и практика: учебник. – 2-е изд. – М.: Юрайт, 2012. – 464 с.
2. Бондарь А.Г. Microsoft SQL Server 2012. – СПб.: БХВ-Петербург, 2013. – 608 с.
3. Кузин А.В., Левонисова С.В. Базы данных: учебное пособие для вузов. – 5-е изд., испр. – М.: Академия, 2012. – 320 с.
4. Осипов Г.С. Лекции по искусственному интеллекту / Г.С. Осипов, Российская академия наук. – М.: Эдиториал УРСС, 2014. – 272 с.

Дополнительная литература:

5. Баженова И.Ю. Основы проектирования приложений баз данных: учебное пособие. – М.: Интернет-Ун-т информ. Технологий: БИНОМ. Лаборатория знаний, 2009. – 325 с.
6. Кузнецов С.Д. Основы баз данных: учебное пособие. – 2-е изд., испр. – М.: Интернет-Ун-т информ. технологий: БИНОМ. Лаборатория знаний, 2010. – 484 с.

7. Илюшечкин В.М. Основы использования и проектирования баз данных: учебное пособие для вузов. – М.: Юрайт, 2010. – 213 с.

8. Башмаков А.И., Башмаков И.А. Интеллектуальные информационные технологии: учебное пособие для вузов. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2005. – 304 с.

Лицензионное программное обеспечение: Среда разработки Microsoft Visual Studio, СУБД Microsoft SQL Server.