

Министерство науки и высшего образования РФ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Национальный исследовательский университет «МЭИ»

Направление подготовки/специальность: 09.03.01 Информатика и вычислительная техника

Наименование образовательной программы: Системы автоматизированного проектирования

Уровень образования: высшее образование - бакалавриат

Форма обучения: Очная

Рабочая программа дисциплины
ТЕХНОЛОГИИ ОБРАБОТКИ БОЛЬШИХ ДАННЫХ

| | |
|--|---|
| Блок: | Блок 1 «Дисциплины (модули)» |
| Часть образовательной программы: | Часть, формируемая участниками образовательных отношений |
| № дисциплины по учебному плану: | Б1.Ч.12 |
| Трудоемкость в зачетных единицах: | 6 семестр - 2; |
| Часов (всего) по учебному плану: | 72 часа |
| Лекции | 6 семестр - 14 часов; |
| Практические занятия | не предусмотрено учебным планом |
| Лабораторные работы | 6 семестр - 12 часов; |
| Консультации | проводится в рамках часов аудиторных занятий |
| Самостоятельная работа | 6 семестр - 45,7 часа; |
| в том числе на КП/КР | не предусмотрено учебным планом |
| Иная контактная работа | проводится в рамках часов аудиторных занятий |
| включая: Лабораторная работа Контрольная работа | |
| Промежуточная аттестация: | |
| Зачет с оценкой | 6 семестр - 0,3 часа; |

Москва 2023

ПРОГРАММУ СОСТАВИЛ:

Преподаватель

| | | |
|--|--|------------------------------|
| | Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ» | |
| | Сведения о владельце ЦЭП МЭИ | |
| | Владелец | Бородин Г.А. |
| | Идентификатор | R607fd388-BorodinGA-3d6314d0 |

Г.А. Бородин

СОГЛАСОВАНО:

Руководитель
образовательной программы

| | | |
|--|--|-------------------------------|
| | Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ» | |
| | Сведения о владельце ЦЭП МЭИ | |
| | Владелец | Андреева И.Н. |
| | Идентификатор | Rb5322c60-AndreevaIN-0472a135 |

И.Н. Андреева

Заведующий выпускающей
кафедрой

| | | |
|--|--|-------------------------------|
| | Подписано электронной подписью ФГБОУ ВО «НИУ «МЭИ» | |
| | Сведения о владельце ЦЭП МЭИ | |
| | Владелец | Топорков В.В. |
| | Идентификатор | Rc76a6458-ToporkovVV-1f71a135 |

В.В. Топорков

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цель освоения дисциплины: изучение базовых принципов организации, архитектуры и функционирования современных систем поддержки принятия решений, языков запросов и алгоритмов анализа больших данных, администрирования и эксплуатации OLAP кубов.

Задачи дисциплины

- знакомство с составом программных компонент и модулей и принципами развёртывания систем поддержки принятия решений;
- приобретение навыков разработки хранилищ и витрин данных, OLAP кубов, анализа больших данных, составления запросов на выборку данных, подготовки отчётов;
- знакомство с методами системного администрирования и обеспечения информационной безопасности OLAP систем.

Формируемые у обучающегося **компетенции** и запланированные **результаты обучения** по дисциплине, соотнесенные с **индикаторами достижения компетенций**:

| Код и наименование компетенции | Код и наименование индикатора достижения компетенции | Запланированные результаты обучения |
|---|--|--|
| ПК-1 Способен применять методы и средства сборки модулей и компонент программного обеспечения, разработки процедур для развёртывания программного обеспечения, миграции и преобразования данных, создания программных интерфейсов | ИД-1 _{ПК-1} Демонстрирует умение использовать современные технологии разработки ПО | знать: - основы построения хранилищ и витрин данных, миграции и преобразования данных из OLTP баз данных в OLAP кубы; - состав программных модулей OLAP систем и принципы развёртывания компонент систем поддержки принятия решений. уметь: - работать с современными системами программирования для разработки хранилищ и витрин данных, OLAP кубов и отчётов на их основе. |
| ПК-2 Способен определять конфигурацию и технические характеристики оборудования, необходимые для установки программного продукта | ИД-2 _{ПК-2} Демонстрирует умение устанавливать программное обеспечение для информационных и автоматизированных систем | знать: - методику настройки сетевого взаимодействия различных программных компонент и модулей хранилищ, витрин данных и OLAP кубов; - методику инсталляции аппаратного и программного обеспечения для современных OLAP систем. уметь: - устанавливать, тестировать, развёртывать и использовать современные OLAP системы для анализа больших данных. |

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ ВО

Дисциплина относится к основной профессиональной образовательной программе Системы автоматизированного проектирования (далее – ОПОП), направления подготовки 09.03.01

Информатика и вычислительная техника, уровень образования: высшее образование - бакалавриат.

Требования к входным знаниям и умениям:

- знать теорию баз данных
- знать основы объектно-ориентированное программирования
- уметь устанавливать и администрировать систему управления базами данных
- уметь применять современные компиляторы, отладчики и оптимизаторы программного кода

Результаты обучения, полученные при освоении дисциплины, необходимы при выполнении выпускной квалификационной работы.

3. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

3.1 Структура дисциплины

Общая трудоемкость дисциплины составляет 2 зачетных единицы, 72 часа.

| № п/п | Разделы/темы дисциплины/формы промежуточной аттестации | Всего часов на раздел | Семестр | Распределение трудоемкости раздела (в часах) по видам учебной работы | | | | | | | | | | Содержание самостоятельной работы/ методические указания | |
|-------|--|-----------------------|-----------|--|-----------|----|--------------|---|-----|----|------------|-------------------|-----------------------------------|---|---|
| | | | | Контактная работа | | | | | | | СР | | | | |
| | | | | Лек | Лаб | Пр | Консультация | | ИКР | | ПА | Работа в семестре | Подготовка к аттестации /контроль | | |
| КПР | ГК | ИККП | ТК | | | | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | |
| 1 | Системы поддержки принятия решений | 15 | 6 | 4 | - | - | - | - | - | - | - | 11 | - | <p><u>Подготовка к лабораторной работе:</u> Разработка и создание OLAP кубов</p> <p><u>Изучение материалов литературных источников:</u> [1], стр. 7-24, 26-36, 45-52</p> <p><u>Подготовка к лабораторной работе:</u> Изучение алгоритмов и моделей Data Mining</p> <p><u>Изучение материалов литературных источников:</u> [2], стр. 9-59</p> <p><u>Подготовка к контрольной работе:</u> Запросы MDX к OLAP кубам</p> <p><u>Подготовка к лабораторной работе:</u> Создание отчетов для баз данных</p> <p><u>Изучение материалов литературных источников:</u> [1], стр. 25-52 [3], стр. 5-18, 38-61 [4], стр. 22-142</p> <p><u>Подготовка к контрольной работе:</u> Запросы DMX к моделям и структурам</p> <p><u>Изучение материалов литературных источников:</u> [2], стр. 22-33, 47-53</p> | |
| 1.1 | Основы систем обработки данных | 15 | | 4 | - | - | - | - | - | - | - | - | 11 | | - |
| 2 | OLAP и MDX | 18 | | 3 | 4 | - | - | - | - | - | - | - | 11 | | - |
| 2.1 | Основы построения OLAP-кубов | 18 | | 3 | 4 | - | - | - | - | - | - | - | 11 | | - |
| 3 | Data Mining и DMX | 20 | | 3 | 4 | - | - | - | - | - | - | - | 13 | | - |
| 3.1 | Введение в анализ данных | 20 | | 3 | 4 | - | - | - | - | - | - | - | 13 | | - |
| 4 | Отчёты для баз данных | 18.7 | | 4 | 4 | - | - | - | - | - | - | - | 10.7 | | - |
| 4.1 | Виды и формы представления результатов | 18.7 | | 4 | 4 | - | - | - | - | - | - | - | 10.7 | | - |
| | Зачет с оценкой | 0.3 | | - | - | - | - | - | - | - | - | 0.3 | - | | - |
| | Всего за семестр | 72.0 | | 14 | 12 | - | - | - | - | - | - | 0.3 | 45.7 | | - |
| | Итого за семестр | 72.0 | 14 | 12 | - | - | - | - | - | - | 0.3 | 45.7 | - | | |

Примечание: Лек – лекции; Лаб – лабораторные работы; Пр – практические занятия; КПр – аудиторные консультации по курсовым проектам/работам; ИККП – индивидуальные консультации по курсовым проектам/работам; ГК- групповые консультации по разделам дисциплины; СР – самостоятельная работа студента; ИКР – иная контактная работа; ТК – текущий контроль; ПА – промежуточная аттестация

3.2 Краткое содержание разделов

1. Системы поддержки принятия решений

1.1. Основы систем обработки данных

Понятие, формальные характеристики и архитектура СППР. Информационные системы руководителя. Склады данных. Цели создания хранилища, основные функциональные узлы, принципы организации склада, правила для хранилищ данных. Уровни обобщения данных. Концепция разработки склада. Модель и анализ данных предприятия. Анализ процессов и событий. Достоинства и недостатки хранилищ. Загрузка данных (ETL). Этапы подготовки данных: сбор, очистка, агрегирование. Витрины данных. Достоинства и недостатки.

2. OLAP и MDX

2.1. Основы построения OLAP-кубов

Основы OLAP систем. Правила для OLAP систем и тест FASMI. Архитектура многомерного анализа данных. Категории OLAP инструментов (MOLAP, ROLAP, HOLAP). Особенности многомерного хранения данных, схемы «звезда» и «снежинка». Основные понятия OLAP кубов: меры, измерения, элементы, иерархии, наборы, кортежи, атрибуты, срезы, вычисляемые элементы, ключевые показатели эффективности. Служба SSAS SQL Server. Технология интерактивной работы с OLAP кубом. Расширения языка SQL: RISQL и MDX. Основы языка MDX: идентификаторы, выражения, операции, спецсимволы, операторы и функции. Оператор SELECT и его основные разделы. PivotTable Service. Взаимодействие с офисными приложениями.

3. Data Mining и DMX

3.1. Введение в анализ данных

Введение в анализ данных. Способы поиска зависимостей между данными. Методы выявления знаний. Стандарты. OLE DB for Data Mining и XMLA. Java Data Mining API (JSR-73). ISO/IEC SQL/MM Part 6 Data Mining. Основы и назначение языка DMX: идентификаторы, выражения, операторы, типы данных и содержимого, классифицированные столбцы, типы использования, флаги моделирования. Функции DMX. Алгоритмы DMX: взаимосвязей, кластеризации, дерева принятия решений, линейной и логистической регрессии, Байеса, нейронной сети, временных рядов, кластеризации последовательностей. Методика создания и подключения пользовательских алгоритмов. Структуры и модели данных. Источники данных. Обработка моделей и структур. Виды запросов к моделям и структурам. Надстройка для Excel. Обзор функций аналитики.

4. Отчёты для баз данных

4.1. Виды и формы представления результатов

Crystal Reports. Основные возможности, создание файл-макета отчёта, варианты вывода отчёта на печать. Crystal Reports for .NET. Microsoft Reporting Services. Архитектура, варианты доступа, сценарии использования. Публикация отчётов. Безопасность и права доступа к сервису. Язык описания отчётов RDL. Макеты отчётов. Алгоритм создания и публикации отчёта. Построители отчётов Report Builder 1 и 3. Конструктор отчётов в Visual Studio. Развёртывание, запуск, экспорт отчёта. Диспетчер отчётов. Отчёты на Java. Crystal Reports, Pentaho Reports, Jasper Reports, i-net Clear Reports.

3.3. Темы практических занятий

не предусмотрено

3.4. Темы лабораторных работ

1. Разработка и создание OLAP кубов;
2. Изучение алгоритмов и моделей Data Mining;
3. Создание отчётов для баз данных.

3.5 Консультации

3.6 Тематика курсовых проектов/курсовых работ

Курсовой проект/ работа не предусмотрены

3.7. Соответствие разделов дисциплины и формируемых в них компетенций

| Запланированные результаты обучения по дисциплине (в соответствии с разделом 1) | Коды индикаторов | Номер раздела дисциплины (в соответствии с п.3.1) | | | | Оценочное средство (тип и наименование) |
|---|------------------|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | |
| Знать: | | | | | | |
| состав программных модулей OLAP систем и принципы развёртывания компонент систем поддержки принятия решений | ИД-1ПК-1 | | + | | | Лабораторная работа/Изучение алгоритмов и моделей Data Mining |
| основы построения хранилищ и витрин данных, миграции и преобразования данных из OLTP баз данных в OLAP кубы | ИД-1ПК-1 | + | | | | Лабораторная работа/Разработка и создание OLAP кубов |
| методику инсталляции аппаратного и программного обеспечения для современных OLAP систем | ИД-2ПК-2 | | | + | | Лабораторная работа/Создание отчётов для баз данных |
| методику настройки сетевого взаимодействия различных программных компонент и модулей хранилищ, витрин данных и OLAP кубов | ИД-2ПК-2 | | | | + | Контрольная работа/Запросы DMX к моделям и структурам |
| Уметь: | | | | | | |
| работать с современными системами программирования для разработки хранилищ и витрин данных, OLAP кубов и отчётов на их основе | ИД-1ПК-1 | | | + | | Контрольная работа/Составление запросов MDX к OLAP кубам |
| инсталлировать, тестировать, развёртывать и использовать современные OLAP системы для анализа больших данных | ИД-2ПК-2 | + | + | + | + | Контрольная работа/Запросы DMX к моделям и структурам Лабораторная работа/Изучение алгоритмов и моделей Data Mining Лабораторная работа/Разработка и создание OLAP кубов Лабораторная работа/Создание отчётов для баз данных |

4. КОМПЕТЕНТНОСТНО-ОРИЕНТИРОВАННЫЕ ОЦЕНОЧНЫЕ СРЕДСТВА ДЛЯ КОНТРОЛЯ ОСВОЕНИЯ РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (ТЕКУЩИЙ КОНТРОЛЬ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНАЯ АТТЕСТАЦИЯ ПО ДИСЦИПЛИНЕ)

4.1. Текущий контроль успеваемости

6 семестр

Форма реализации: Компьютерное задание

1. Запросы DMX к моделям и структурам (Контрольная работа)
2. Изучение алгоритмов и моделей Data Mining (Лабораторная работа)
3. Разработка и создание OLAP кубов (Лабораторная работа)
4. Создание отчетов для баз данных (Лабораторная работа)
5. Составление запросов MDX к OLAP кубам (Контрольная работа)

Балльно-рейтинговая структура дисциплины является приложением А.

4.2 Промежуточная аттестация по дисциплине

Зачет с оценкой (Семестр №6)

Оценка определяется в соответствии с Положением о балльно-рейтинговой системе для студентов НИУ «МЭИ» по совокупности результатов текущего контроля успеваемости.

В диплом выставляется оценка за 6 семестр.

Примечание: Оценочные материалы по дисциплине приведены в фонде оценочных материалов ОПОП.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

5.1 Печатные и электронные издания:

1. Бородин, Г. А. Основы OLAP и MDX : учебное пособие по курсу "Проектирование баз данных" по направлению "Информатика и вычислительная техника" / Г. А. Бородин, И. Н. Андреева, Нац. исслед. ун-т "МЭИ" (НИУ"МЭИ") . – М. : Изд-во МЭИ, 2018 . – 55 с. - ISBN 978-5-7046-1905-5 .
<http://elibr.mpei.ru/elibr/view.php?id=10083>;
2. Бородин, Г. А. Основы Data Mining и DMX : учебное пособие по курсу "Базы данных" по направлению "Информатика и вычислительная техника" / Г. А. Бородин, И. Н. Андреева, Нац. исслед. ун-т "МЭИ" (НИУ"МЭИ") . – М. : Изд-во МЭИ, 2018 . – 61 с. - ISBN 978-5-7046-1992-5 .
<http://elibr.mpei.ru/elibr/view.php?id=10230>;
3. Бородин, Г. А. Отчеты для баз данных : учебное пособие по курсу "Базы данных" по направлению "Информатика и вычислительная техника" / Г. А. Бородин, И. Н. Андреева, Нац. исслед. ун-т "МЭИ" (НИУ"МЭИ") . – М. : Изд-во МЭИ, 2018 . – 64 с. - ISBN 978-5-7046-2059-4 .
<http://elibr.mpei.ru/elibr/view.php?id=10346>;
4. Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н.- "Большие данные. Big Data", Издательство: "Лань", Санкт-Петербург, 2021 - (188 с.)
<https://e.lanbook.com/book/165835>.

5.2 Лицензионное и свободно распространяемое программное обеспечение:

1. СДО "Прометей";
2. Office / Российский пакет офисных программ;

3. Windows / Операционная система семейства Linux;
4. Видеоконференции (Майнд, Сберджаз, ВК и др);
5. Visual Studio.

5.3 Интернет-ресурсы, включая профессиональные базы данных и информационно-справочные системы:

1. ЭБС Лань - <https://e.lanbook.com/>
2. ЭБС "Университетская библиотека онлайн" - http://biblioclub.ru/index.php?page=main_ub_red
3. Научная электронная библиотека - <https://elibrary.ru/>
4. База данных Web of Science - <http://webofscience.com/>
5. База данных Scopus - <http://www.scopus.com>
6. Национальная электронная библиотека - <https://rusneb.ru/>
7. ЭБС "Консультант студента" - <http://www.studentlibrary.ru/>
8. Электронная библиотека МЭИ (ЭБ МЭИ) - <http://elib.mpei.ru/login.php>
9. Портал открытых данных Российской Федерации - <https://data.gov.ru>
10. База открытых данных Министерства труда и социальной защиты РФ - <https://rosmintrud.ru/opendata>
11. База открытых данных профессиональных стандартов Министерства труда и социальной защиты РФ - <http://profstandart.rosmintrud.ru/obshchiy-informatsionnyy-blok/natsionalnyy-reestr-professionalnykh-standartov/>
12. База открытых данных Министерства экономического развития РФ - <http://www.economy.gov.ru>
13. База открытых данных Росфинмониторинга - <http://www.fedsfm.ru/opendata>
14. Электронная открытая база данных "Polpred.com Обзор СМИ" - <https://www.polpred.com>
15. Информационно-справочная система «Кодекс/Техэксперт» - <Http://proinfosoft.ru;http://docs.cntd.ru/>

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

| Тип помещения | Номер аудитории, наименование | Оснащение |
|---|---------------------------------------|--|
| Учебные аудитории для проведения лекционных занятий и текущего контроля | Е-405, Учебная аудитория каф. "ВТ" | парта, стол преподавателя, стул, шкаф для документов, доска меловая, компьютерная сеть с выходом в Интернет, доска маркерная передвижная |
| | Ж-120, Машинный зал ИВЦ | сервер, кондиционер |
| Учебные аудитории для проведения лабораторных занятий | Е-412, Учебная лаборатория кафедры ВТ | стол, стол компьютерный, стул, компьютерная сеть с выходом в Интернет, доска маркерная передвижная, сервер, компьютер персональный |
| | Ж-120, Машинный зал ИВЦ | сервер, кондиционер |
| Учебные аудитории для проведения промежуточной аттестации | Е-405, Учебная аудитория каф. "ВТ" | парта, стол преподавателя, стул, шкаф для документов, доска меловая, компьютерная сеть с выходом в Интернет, доска маркерная передвижная |
| | Ж-120, Машинный зал ИВЦ | сервер, кондиционер |
| Помещения для | НТБ-201, | стол компьютерный, стул, стол |

| | | |
|--|------------------------------|--|
| самостоятельной работы | Компьютерный читальный зал | письменный, вешалка для одежды, компьютерная сеть с выходом в Интернет, компьютер персональный, принтер, кондиционер |
| Помещения для консультирования | Е-411, Лаборатория каф. "ВТ" | стол, стол компьютерный, стул, шкаф для документов, шкаф для одежды, тумба, компьютерная сеть с выходом в Интернет, компьютер персональный |
| Помещения для хранения оборудования и учебного инвентаря | Е-403, Склад | стол для работы с документами, шкаф, шкаф для документов, книги, учебники, пособия, дипломные и курсовые работы студентов |

БАЛЛЬНО-РЕЙТИНГОВАЯ СТРУКТУРА ДИСЦИПЛИНЫ

Технологии обработки больших данных

(название дисциплины)

6 семестр

Перечень контрольных мероприятий текущего контроля успеваемости по дисциплине:

- КМ-1 Разработка и создание OLAP кубов (Лабораторная работа)
- КМ-2 Изучение алгоритмов и моделей Data Mining (Лабораторная работа)
- КМ-3 Запросы DMX к моделям и структурам (Контрольная работа)
- КМ-4 Создание отчётов для баз данных (Лабораторная работа)
- КМ-5 Составление запросов MDX к OLAP кубам (Контрольная работа)

Вид промежуточной аттестации – Зачет с оценкой.

| Номер раздела | Раздел дисциплины | Индекс КМ: | КМ-1 | КМ-2 | КМ-3 | КМ-4 | КМ-5 |
|---------------|--|------------|------|------|------|------|------|
| | | Неделя КМ: | 4 | 8 | 8 | 12 | 12 |
| 1 | Системы поддержки принятия решений | | | | | | |
| 1.1 | Основы систем обработки данных | | + | + | + | + | |
| 2 | OLAP и MDX | | | | | | |
| 2.1 | Основы построения OLAP-кубов | | + | + | + | + | |
| 3 | Data Mining и DMX | | | | | | |
| 3.1 | Введение в анализ данных | | + | + | + | + | + |
| 4 | Отчёты для баз данных | | | | | | |
| 4.1 | Виды и формы представления результатов | | + | + | + | + | |
| Вес КМ, %: | | | 30 | 20 | 15 | 20 | 15 |