Министерство образования и науки РФ

Государственное образовательное учреждение высшего профессионального образования

**«Санкт-Петербургский государственный электротехнический университет “ЛЭТИ” им. В.И. Ульянова (Ленина)» (СПбГЭТУ)**

РАБОЧАЯ ПРОГРАММА

дисциплины по выбору студентов №5

*«Хранилища и витрины данных»*

Для подготовки специалистов по специальности 230102.65

*«Автоматизированные системы обработки информации и управления»*

Санкт-Петербург

2011

# Санкт-Петербургский государственный электротехнический

университет “ЛЭТИ”

“УТВЕРЖДАЮ”

Проректор по учебной работе

Лысенко Н.В.

“\_\_\_\_\_”\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_2011 г.

РАБОЧАЯ ПРОГРАММА

дисциплины по выбору студентов №5

*«Хранилища и витрины данных»*

Для подготовки специалистов по специальности 230102.65

*«Автоматизированные системы обработки информации и управления»*

Учебный план № 333

Факультет компьютерных технологий и информатики

Кафедра автоматизированных систем обработки информации и управления

Курс – 4

Семестр – 8

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Лекции | 68 ч. |  | Экзамен | 8 семестр |
|  |  |  |  |  |
| Лабораторные занятия | 17 ч. |  | Зачет | 8 семестр |
|  |  |  |  |  |
| Курсовое проектирование | 17 ч. |  |  |  |

|  |  |  |
| --- | --- | --- |
| Аудиторные занятия | 102 ч. |  |
| Самостоятельные занятия | 51 ч. |  |
| Всего часов | 153 ч. |  |

##### Санкт-Петербург

##### 2011

Рабочая программа обсуждена на заседании кафедры *автоматизированных систем обработки информации и управления* “\_\_\_\_”\_\_\_\_\_\_\_\_\_\_2011 г., протокол №\_\_\_\_\_\_.

Рабочая программа составлена в соответствии с государственным образовательным стандартом для подготовки специалистов по специальности

230102.65 – «Автоматизированные системы обработки информации и управления»

Рабочая программа согласована с рабочими программами изученных ранее дисциплин:

1) Базы данных

2) Организация ЭВМ и систем

и является фундаментом для изучения следующих дисциплин:

1) Моделирование систем

Рабочая программа одобрена методической комиссией факультета компьютерных технологий и информатики “\_\_\_\_”\_\_\_\_\_\_\_\_\_\_\_\_\_2011 г.

**Аннотация дисциплины**

Информационная технология и концепция хранилищ данных. Планирование и построение хранилища данных. Моделирование данных. Репозиторий метаданных. Повышение качества информации в хранилищах данных. Концептуальная и логическая архитектура хранилища данных. Физическая архитектура хранилища данных. Преобразование данных. Доступ к данным. Системы оперативной аналитической обработки (OLAP).

**Цели и задачи дисциплины**

1. Ознакомление с методами создания современных систем поддержки принятия решений, в основе которых лежат хранилища данных.
2. Изучение понятий многомерного концептуального представления данных и методов оперативной аналитической обработки (OLAP).

**Требования к уровню освоения дисциплины**

В результате изучения дисциплины студенты должны

* знать способы построения репозитория метаданных;
* уметь работать с системами оперативной аналитической обработки;
* иметь представление о моделировании корпоративных данных, включая оперативные ER-модели и аналитические схемы типа «звезда»; концептуальной, логической и физической архитектуре хранилища данных; многомерной модели данных.

**Содержание**

**Тема 1. Информационная технология и концепция хранилищ данных.**

Краткая история информационных технологий. Корпоративная модель данных. Архитектура хранилища данных. Доступ потребителей к информации.

**Тема 2.** **Планирование и построение хранилища данных.**

Концептуальная архитектура хранилища. Логическая архитектура хранилища. Физическая архитектура хранилища. Архитектура данных. Реализация.

**Тема 3.** **Моделирование данных.**

Корпоративная модель данных. Анализ схемы типа «звезда». Разработка размерностной модели. Схема типа «снежинка». Физическое моделирование. ER-модель и схема типа «звезда».

**Тема 4. Репозиторий метаданных.**

Модель использования метаданных. Размерностная модель метаданных. Справочник пользователя информации.

**Тема 5. Повышение качества информации в хранилищах данных.**

Методы оценки данных. Программные средства оценки качества данных. Аудит данных.

**Тема 6. Концептуальная и логическая архитектура хранилища данных.**

Архитектурные принципы. Концептуальные модели. Логические модели.

**Тема 7. Физическая архитектура хранилища данных.**

Физическая память. Базы данных. Аппаратное обеспечение сервера баз данных. Операционные системы. Сервер запросов и сервер приложений. Сети и средства взаимодействия. Программное обеспечение промежуточного уровня. Рабочие станции специалистов в области инженерных знаний. Развертывание архитектуры.

**Тема 8. Преобразование данных.**

Планирование. Методы извлечения и перемещения данных. Преобразование данных. Загрузка данных.

**Тема 9. Доступ к данным.**

Выбор средств доступа к данным. Распределение информации. Доступ с помощью Web. Электронные таблицы. Средства визуализации. Средства создания запросов. Программные средства СППР.

**Тема 10. Системы оперативной аналитической обработки (OLAP).**

Многомерная модель данных. Основные понятия многомерной модели. Определение OLAP-систем. Архитектура OLAP-систем. Классификация OLAP по способу представления данных. Интеграция OLAP и интеллектуального анализа данных.

**Перечень лабораторных работ**

|  |  |  |
| --- | --- | --- |
| № | Наименование работы | Номер темы |
| 1 | Освоение моделирования данных в нотации IDEF1X | 2 |
| 2 | Разработка логической модели | 3, 4 |
| 3 | Разработка физической модели | 3, 4 |
| 4 | Многомерное моделирование | 8, 10 |
| 5 | Моделирование временных данных и Data Vault | 10 |

**Цели и содержание курсового проекта**

и его ориентировочная трудоемкость

В предлагаемом курсовом проекте рассматриваются проблемы и подходы к проектированию многомерных баз данных. Работа посвящена технологии многомерного анализа данных OLAP, рассматриваются концепции хранилищ данных, требования к многомерным базам данных и OLAP-средствам, логическая организация OLAP-данных, применение OLAP-технологий.

Задача проекта: исследование технологии OLAP, её возможностей и применения многомерных хранилищ данных. Нобходимо рассмотреть основные подходы к проектированию и созданию хранилищ данных с помощью Oracle Express. В качестве практической части предлагается создание OLAP-куба средствами Microsoft Analysis Services.

Ориентировочная трудоемокость — 36 часов.

**Распределение учебных часов по темам и видам занятий**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| №  темы | Название разделов и тем | Объем учебных часов | | | | | | Семестр | Литература по темам |
| Лекции | Лабор.  занятия | Практ.  занятия | Аудит.  занятия | Сам.  работа | Всего |
| 1 | Информационная технология и концепция хранилищ данных. | 5 |  |  | 5 |  | 5 | 8 | Л1, Л2, Д4, Д5 |
| 2 | Планирование и построение хранилища данных. | 5 | 3 |  | 8 | 4 | 12 | 8 | Л1, Л2 |
| 3 | Моделирование данных. | 8 | 3 |  | 11 | 4 | 15 | 8 | Л1, Л2 |
| 4 | Репозиторий метаданных. | 6 | 4 |  | 10 | 4 | 14 | 8 | Л1 |
| 5 | Повышение качества информации в хранилищах данных. | 6 |  |  | 6 | 4 | 10 | 8 | Л1 |
| 6 | Концептуальная и логическая архитектура хранилища данных. | 8 |  |  | 8 | 4 | 12 | 8 | Л1 |
| 7 | Физическая архитектура хранилища данных. | 8 |  |  | 8 | 4 | 12 | 8 | Л1 |
| 8 | Преобразование данных. | 8 | 3 |  | 11 | 4 | 15 | 8 | Л1, Л3 |
| 9 | Доступ к данным. | 4 |  |  | 4 |  | 4 | 8 | Л3, Д2 |
| 10 | Системы оперативной аналитической обработки (OLAP). | 10 | 4 |  | 14 | 4 | 18 | 8 | Л3, Д1, Д3 |
| Курсовое проектирование | |  |  | 17 | 17 | 19 | 36 | 8 |  |
| **ИТОГО:** | | 68 | 17 | 17 | 102 | 51 | 153 |  |  |

**Учебно-методическое обеспечение дисциплины**

**Основная литература**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| № | Название, библиографическое описание | Л | Лр | Пз (С) | Кп (р) | Инд.  зад. | К-во экз. в библ. (на каф.) |
| **1** | Барсегян А. А., Куприянов М. С., Степаненко В. В., Холод И. И. Методы и модели анализа данных: OLAP и Data Min-ing. СПб.: БХВ-Петербург, 2004. 336 с. | 8 | 8 |  | 8 |  | 54 |
| **2** | Архипенков С., Голубев Д., Максименко О. Хранилища данных. От концепции до внедрения. М.: Диалог-МИФИ, 2002. 528 с. | 8 | 8 |  | 8 |  | 37 |
| **3** | Маклаков С.В., Туманов В.Е. Проектирование реляционных хранилищ данных. М.: Диалог-МИФИ, 2007. 333 с. | 8 | 8 |  | 8 |  | 43 |

##### **Дополнительная литература**

|  |  |  |
| --- | --- | --- |
| № | Название, библиографическое описание | К-во экз. в библ. (на каф.) |
| **1** | Спирли Э. Корпоративные хранилища данных. Планирование, разработка, реализация. М.: Издательский дом «Вильямс», 2001. 400 с | **5** |
| **2** | Архипенков С. Oracle Express OLAP. М.: Диалог-МИФИ, 1999. 320 с. | **10** |
| **3** | Корнеев В. В., Гареев А. Ф., Васютин С. В., Райх В. В. Базы данных. Интеллектуальная обработка информации. М.: Нолиндж, 2000. 352 с. | **9** |
| **4** | Елманова Н., Федоров А. Введение в OLAP-технологии Microsoft. М.: Диалог-МИФИ, 2002. 272 с. | **нет** |
| **5** | Гаврилова Т. А., Хорошевский В. Ф. Базы знаний интеллектуальных систем. СПб.: Питер, 2000. 384 с. | **49** |
| **6** | Романов В. П. Интеллектуальные информационные системы в экономике. М.: Экзамен, 2003. 496 с. | **нет** |

|  |  |
| --- | --- |
| Зав. отделом учебной литературы *(для технических дисциплин)* | Киселева Т.В. |

|  |  |
| --- | --- |
| Авторы:  к.т.н., доцент | Савосин С.В. |
|  |  |
| Рецензент:  д.т.н., профессор | Водяхо А.И. |
|  |  |
| Зав. кафедрой автоматизированных систем обработки информации  и управления  д.т.н., профессор | Советов Б.Я. |
|  |  |
| Декан факультета компьютерных технологий и информатики  д.т.н., профессор | Куприянов М.С. |
|  |  |
| Программа согласована: |  |
|  |  |
| Председатель методической комиссии факультета компьютерных  технологий и информатики |  |
| к.т.н., доц. | Михалков В.А. |
|  |  |
| Руководитель методического отдела |  |
| к.т.н., доцент | Марасина Л.А. |