

МИНОБРНАУКИ РОССИИ

**Федеральное государственное бюджетное образовательное
учреждение высшего образования "Пермский
государственный национальный исследовательский
университет"**

Кафедра радиоэлектроники и защиты информации

Авторы-составители: **Луногов Игорь Владимирович
Скляренко Максим Сергеевич**

Рабочая программа дисциплины

ХРАНИЛИЩА ДАННЫХ

Код УМК 95657

Утверждено
Протокол №4
от «24» июня 2020 г.

Пермь, 2020

1. Наименование дисциплины

Хранилища данных

2. Место дисциплины в структуре образовательной программы

Дисциплина входит в обязательную часть Блока « Б.1 » образовательной программы по направлениям подготовки (специальностям):

Направление: **01.03.02** Прикладная математика и информатика
направленность Инженерия программного обеспечения

3. Планируемые результаты обучения по дисциплине

В результате освоения дисциплины **Хранилища данных** у обучающегося должны быть сформированы следующие компетенции:

01.03.02 Прикладная математика и информатика (направленность : Инженерия программного обеспечения)

ПК.5 Способен разрабатывать требования и проектировать программное обеспечение, в том числе интеллектуальные информационные системы

Индикаторы

ПК.5.2 Проектирует используемые структуры данных и программные интерфейсы, разрабатывает алгоритмы и оценивает эффективность их использования

4. Объем и содержание дисциплины

| | |
|---|--|
| Направления подготовки | 01.03.02 Прикладная математика и информатика (направленность: Инженерия программного обеспечения) |
| форма обучения | очная |
| №№ триместров, выделенных для изучения дисциплины | 7 |
| Объем дисциплины (з.е.) | 4 |
| Объем дисциплины (ак.час.) | 144 |
| Контактная работа с преподавателем (ак.час.), в том числе: | 56 |
| Проведение лекционных занятий | 28 |
| Проведение практических занятий, семинаров | 28 |
| Самостоятельная работа (ак.час.) | 88 |
| Формы текущего контроля | Входное тестирование (1) Защищаемое контрольное мероприятие (4) Итоговое контрольное мероприятие (1) |
| Формы промежуточной аттестации | Экзамен (7 триместр) |

5. Аннотированное описание содержания разделов и тем дисциплины

Хранилища данных

Проектирование хранилищ данных

основы хранилищ данных

Что такое хранилище данных. Компоненты хранилища данных. Понятия кубов, измерений, витрин. Методология построения хранилищ данных: ROLAP, MOLAP, HOLAP. Подходы к разработке хранилищ данных: снизу-вверх, сверху-вниз

проектирование хранилищ данных

1. студенты под руководством преподавателя проектируют пример простого хранилища данных: таблицы фактов, таблицы измерений

2. студенты под руководством преподавателя выбирает предметную область учебного проекта и самостоятельно проектирует реляционное хранилище данных, хранилище наполняет тестовыми данными.

архитектура современных реляционных СУБД

архитектура современных реляционных СУБД

Технологические аспекты построения реляционных хранилищ данных

Физический дизайн реляционных хранилищ.

Способы повышения пропускной способности подсистемы ввода/вывода.

Применение секционированных (партиционированных) таблиц, индексов, материализованных представлений.

Рассматриваются также основы оптимизации хранилища и некая специфика СУБД Oracle

Основы ETL

ETL

Понятие о ETL (extract, transform, load), оперативный склад данных (область подготовки данных).

Конвейерный и секционный параллелизм, основные типы преобразований. Обзор рынка ETL-средств.

Разработка ETL модулей

изучение основного инструментария Pentaho Data Integration.

Разработка ETL-модулей на Pentaho DataIntegration

Разработка процедуру агрегации по иерархии с использованием pl/PGSQL

Анализ данных реляционных хранилищ

Аналитические возможности языка SQL

Аналитические функции, конструкции CUBE, ROLLUP.

Системы бизнес аналитики

Основные компоненты систем бизнес-аналитики. Обзор рынка BI-систем. Построение регламентной и аналитической отчетности на примере средства Pentaho BI.

Итоговый тест

Итоговое тестирование по программе дисциплины.

6. Методические указания для обучающихся по освоению дисциплины

Освоение дисциплины требует систематического изучения всех тем в той последовательности, в какой они указаны в рабочей программе.

Основными видами учебной работы являются аудиторские занятия. Их цель - расширить базовые знания обучающихся по осваиваемой дисциплине и систему теоретических ориентиров для последующего более глубокого освоения программного материала в ходе самостоятельной работы. Обучающемуся важно помнить, что контактная работа с преподавателем эффективно помогает ему овладеть программным материалом благодаря расстановке необходимых акцентов и удержанию внимания интонационными модуляциями голоса, а также подключением аудио-визуального механизма восприятия информации.

Самостоятельная работа преследует следующие цели:

- закрепление и совершенствование теоретических знаний, полученных на лекционных занятиях;
- формирование навыков подготовки текстовой составляющей информации учебного и научного назначения для размещения в различных информационных системах;
- совершенствование навыков поиска научных публикаций и образовательных ресурсов, размещенных в сети Интернет;
- самоконтроль освоения программного материала.

Обучающемуся необходимо помнить, что результаты самостоятельной работы контролируются преподавателем во время проведения мероприятий текущего контроля и учитываются при промежуточной аттестации.

Обучающимся с ОВЗ и инвалидов предоставляется возможность выбора форм проведения мероприятий текущего контроля, альтернативных формам, предусмотренным рабочей программой дисциплины. Предусматривается возможность увеличения в пределах 1 академического часа времени, отводимого на выполнение контрольных мероприятий.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации.

При проведении текущего контроля применяются оценочные средства, обеспечивающие передачу информации, от обучающегося к преподавателю, с учетом психофизиологических особенностей здоровья обучающихся.

7. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

При самостоятельной работе обучающимся следует использовать:

- конспекты лекций;
- литературу из перечня основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля);
- текст лекций на электронных носителях;
- ресурсы информационно-телекоммуникационной сети "Интернет", необходимые для освоения дисциплины;
- лицензионное и свободно распространяемое программное обеспечение из перечня информационных технологий, используемых при осуществлении образовательного процесса по дисциплине;
- методические указания для обучающихся по освоению дисциплины.

8. Перечень основной и дополнительной учебной литературы

Основная:

1. Петрунин Ю. Ю. Информационные технологии анализа данных. Data analysis : учебное пособие / Ю. Ю. Петрунин. — 2-е изд. — М.: КДУ, 2010. — 292 с. : ил., табл. — ISBN 978-5-98227-701-5. — Текст : электронный // Электронно-библиотечная система БиблиоТех : [сайт]. <https://psu.bibliotech.ru/Reader/Book/7107>
2. Туманов В. Е. Проектирование хранилищ данных для систем бизнес-аналитики: учебное пособие / В. Е. Туманов. — М.: Интернет-Университет Информационных Технологий: БИНОМ. Лаборатория знаний, 2010. — 616 с. — ISBN 978-5-9963-0353-3. — Текст : электронный // Электронно-библиотечная система БиблиоТех : [сайт]. <https://psu.bibliotech.ru/Reader/Book/8850>
3. Дьяков, И. А. Базы данных. Язык SQL : учебное пособие / И. А. Дьяков. — Тамбов : Тамбовский государственный технический университет, ЭБС АСВ, 2012. — 81 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/64070.html>

Дополнительная:

1. Чуприна С. И. Базы данных и СУБД. Практические вопросы проектирования и реализации: учебное пособие для студентов, обучающихся по направлению подготовки бакалавров "Прикладная математика и информатика"/С. И. Чуприна.-Пермь:ПГНИУ,2014, ISBN 978-5-7944-2442-3.-100.-Библиогр.: с. 85-87
2. Смирнов С. Н.,Задворьев И. С. Работаем с Oracle:учеб. пособие/С. Н. Смирнов, И. С. Задворьев.- М.:Гелиос АРВ,2002, ISBN 5-85438-048-X.-496.-Библиогр.: с. 489
3. Туманов, В. Е. Основы проектирования реляционных баз данных : учебное пособие / В. Е. Туманов. — 3-е изд. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 502 с. — ISBN 978-5-4497-0683-6. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. <http://www.iprbookshop.ru/97570.html>

9. Перечень ресурсов сети Интернет, необходимых для освоения дисциплины

https://help.pentaho.com/Documentation/8.3/Products/Pentaho_Data_Integration Официальная документация на Pentaho Data Integration

<https://www.postgresqltutorial.com/> краткие курс по основам SQL в PostgreSQL

<https://intuit.ru/studies/courses/568/424/lecture/9641?page=7> Лекция НОУ ИНТУИТ "Введение в основы OLAP"

<https://intuit.ru/studies/courses/1145/214/lecture/5508?page=2> Доп. материалы по OLAP от НОУ ИНТУИ

<https://www.enterprisedb.com/downloads/postgres-postgresql-downloads> Дистрибутивы PostgreSQL

<https://docs.oracle.com/en/database/oracle/oracle-database/12.2/dwhsg/database-data-warehousing-guide.pdf> Oracle Datawarehouse Guide

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине

Образовательный процесс по дисциплине **Хранилища данных** предполагает использование следующего программного обеспечения и информационных справочных систем:

- 1) презентационные материалы (слайды по темам лекционных и практических занятий);
- 2) доступ в режиме on-line в Электронную библиотечную систему (ЭБС);
- 3) доступ в электронную информационно-образовательную среду университета;
- 4) интернет-сервисы и электронные ресурсы (поисковые системы, электронная почта);

Перечень необходимого лицензионного и (или) свободно распространяемого программного обеспечения

Проигрыватели виртуальных машин VirtualBox и VMWare Player (VMware Workstation). Пакеты офисных программ (тестовые процессоры, табличные редакторы, программы для создания презентаций и др.).

C++ Builder или C#, MS Visual Studio с фреймворком .net минимум версии 4.0

Операционная система ALT Linux;

Офисный пакет приложений «LibreOffice».

PostgreSQL

DBeaver

Pentaho Data Integration

Pentaho BI

При освоении материала и выполнения заданий по дисциплине рекомендуется использование материалов, размещенных в Личных кабинетах обучающихся ЕТИС ПГНИУ (**student.psu.ru**).

При организации дистанционной работы и проведении занятий в режиме онлайн могут использоваться:

система видеоконференцсвязи на основе платформы BigBlueButton (<https://bigbluebutton.org/>).

система LMS Moodle (<http://e-learn.psu.ru/>), которая поддерживает возможность использования текстовых материалов и презентаций, аудио- и видеоконтент, а так же тесты, проверяемые задания, задания для совместной работы.

система тестирования Indigo (<https://indigotech.ru/>).

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Для проведения лекционных занятий:

Аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Для проведения практических занятий – Компьютерный класс, оснащенный персональными ЭВМ и соответствующим программным обеспечением. Состав оборудования определен в Паспорте Компьютерного класса.

Аудитории для проведения текущего контроля;
Компьютерный класс, оснащенный персональными ЭВМ и соответствующим программным обеспечением. Состав оборудования определен в Паспорте компьютерного класса.

Аудитории для групповых (индивидуальных) консультаций;
Аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук) с соответствующим программным обеспечением, меловой (и) или маркерной доской.

Аудитория для самостоятельной работы:
Аудитория оснащенная компьютерной техникой с возможностью подключения к сети «Интернет», обеспеченная доступом в электронную информационно-образовательную среду университета.
Помещения Научной библиотеки ПГНИУ

Помещения научной библиотеки ПГНИУ для обеспечения самостоятельной работы обучающихся:

1. Научно-библиографический отдел, корп.1, ауд. 142. Оборудован 3 персональными компьютера с доступом к локальной и глобальной компьютерным сетям.

2. Читальный зал гуманитарной литературы, корп. 2, ауд. 418. Оборудован 7 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

3. Читальный зал естественной литературы, корп.6, ауд. 107а. Оборудован 5 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

4. Отдел иностранной литературы, корп.2 ауд. 207. Оборудован 1 персональным компьютером с доступом к локальной и глобальной компьютерным сетям.

5. Библиотека юридического факультета, корп.9, ауд. 4. Оборудована 11 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

6. Читальный зал географического факультета, корп.8, ауд. 419. Оборудован 6 персональными компьютерами с доступом к локальной и глобальной компьютерным сетям.

Все компьютеры, установленные в помещениях научной библиотеки, оснащены следующим программным обеспечением:

Операционная система ALT Linux;

Офисный пакет Libreoffice.

Справочно-правовая система «КонсультантПлюс»

**Фонды оценочных средств для аттестации по дисциплине
Хранилища данных**

**Планируемые результаты обучения по дисциплине для формирования компетенции.
Индикаторы и критерии их оценивания**

ПК.5

Способен разрабатывать требования и проектировать программное обеспечение, в том числе интеллектуальные информационные системы

| Компетенция (индикатор) | Планируемые результаты обучения | Критерии оценивания результатов обучения |
|---|--|---|
| <p>ПК.5.2 Проектирует используемые структуры данных и программные интерфейсы, разрабатывает алгоритмы и оценивает эффективность их использования</p> | <p>ЗНАТЬ: архитектуру и методологию построения реляционных хранилищ данных.основы ETL, основные подходы к улучшению производительности БД,основные понятия систем бизнес-аналитики</p> <p>УМЕТЬ: проектировать реляционные хранилища данных</p> <p>ВЛАДЕТЬ: навыками работы в IDE DBEvaer, построения ER - диаграмм, навыками правильного выбора и создания партиций, индексов, материализованных представлений, написания аналитических запросов, навыками реализации аналитической и регламентной отчетности в Pentaho BI</p> | <p align="center">Неудовлетворител</p> <p>не знает: архитектуру и методологию построения реляционных хранилищ данных, основы ETL,основные подходы к улучшению производительности БД, основные понятия систем бизнес-аналитики</p> <p>не умеет: проектировать реляционные хранилища данных, настраивать в Pentaho DataIntegration ETL модули по наполнению хранилищ данных</p> <p>не владеет: навыками работы в IDE DBEvaer, построения ER -диаграмм, навыками написания аналитических запросов, правильного выбора и создания партиций, индексов, материализованных представлений, навыками реализации аналитической и регламентной отчетности в Pentaho BI</p> <p align="center">Удовлетворительн</p> <p>частично знает: архитектуру и методологию построения реляционных хранилищ данных, основы ETL,основные подходы к улучшению производительности БД,основные понятия систем бизнес-аналитики</p> <p>частично умеет: проектировать реляционные хранилища данных,настраивать в Pentaho DataIntegration ETL модули по наполнению хранилищ данных</p> <p>частично владеет: навыками работы в IDE DBEvaer, построения ER -диаграмм, навыками написания аналитических запросов, правильного выбора и создания</p> |

| Компетенция (индикатор) | Планируемые результаты обучения | Критерии оценивания результатов обучения |
|----------------------------|------------------------------------|---|
| | | <p style="text-align: center;">Удовлетворительн</p> <p>партиций, индексов, материализованных представлений, навыками реализации аналитической и регламентной отчетности в Pentaho BI</p> <p style="text-align: center;">Хорошо</p> <p>знает: архитектуру и методологию построения реляционных хранилищ данных, основы ETL, основные подходы к улучшению производительности БД, основные понятия систем бизнес-аналитики</p> <p>умеет: проектировать реляционные хранилища данных, настраивать в Pentaho DataIntegration ETL модули по наполнению хранилищ данных</p> <p>владеет: навыками работы в IDE DBEvaer, построения ER -диаграмм, навыками написания аналитических запросов, правильного выбора и создания партиций, индексов, материализованных представлений, навыками реализации аналитической и регламентной отчетности в Pentaho BI</p> <p style="text-align: center;">Отлично</p> <p>в полной мере знает: архитектуру и методологию построения реляционных хранилищ данных, основы ETL, основные подходы к улучшению производительности БД, основные понятия систем бизнес-аналитики</p> <p>умеет: проектировать реляционные хранилища данных, настраивать в Pentaho DataIntegration ETL модули по наполнению хранилищ данных</p> <p>в полном объеме владеет: навыками работы в IDE DBEvaer, построения ER -диаграмм, навыками написания аналитических запросов, правильного выбора и создания партиций, индексов, материализованных представлений, навыками реализации аналитической и регламентной отчетности в</p> |

| Компетенция (индикатор) | Планируемые результаты обучения | Критерии оценивания результатов обучения |
|------------------------------------|--|---|
| | | Отлично Pentaho BI |

Оценочные средства текущего контроля и промежуточной аттестации

Схема доставки : Базовая

Вид мероприятия промежуточной аттестации : Экзамен

Способ проведения мероприятия промежуточной аттестации : Оценка по дисциплине в рамках промежуточной аттестации определяется на основе баллов, набранных обучающимся на контрольных мероприятиях, проводимых в течение учебного периода.

Максимальное количество баллов : 100

Конвертация баллов в отметки

«отлично» - от 81 до 100

«хорошо» - от 61 до 80

«удовлетворительно» - от 49 до 60

«неудовлетворительно» / «незачтено» менее 49 балла

| Компетенция (индикатор) | Мероприятие текущего контроля | Контролируемые элементы результатов обучения |
|---|---|--|
| Входной контроль | основы хранилищ данных Входное тестирование | контролируются знания основ SQL и БД |
| ПК.5.2 Проектирует используемые структуры данных и программные интерфейсы, разрабатывает алгоритмы и оценивает эффективность их использования | проектирование хранилищ данных Защищаемое контрольное мероприятие | контролируется правильность выбора фактов и измерений, владение DBEaver, навыки создания ER диаграмм. Правильность выбора версионных и иерархических измерений. |
| ПК.5.2 Проектирует используемые структуры данных и программные интерфейсы, разрабатывает алгоритмы и оценивает эффективность их использования | Технологические аспекты построения реляционных хранилищ данных Защищаемое контрольное мероприятие | правильность выбора механизма партиционирования, понимание индексации данных. умения реализации: партиций, индексов, материализованных представлений |
| ПК.5.2 Проектирует используемые структуры данных и программные интерфейсы, разрабатывает алгоритмы и оценивает эффективность их использования | Разработка ETL модулей Защищаемое контрольное мероприятие | реализованные ETL-модули |

| Компетенция (индикатор) | Мероприятие текущего контроля | Контролируемые элементы результатов обучения |
|---|---|--|
| ПК.5.2 Проектирует используемые структуры данных и программные интерфейсы, разрабатывает алгоритмы и оценивает эффективность их использования | Системы бизнес аналитики Защищаемое контрольное мероприятие | навыки написания аналитических запросов, владение инструментами Pentaho BI |
| ПК.5.2 Проектирует используемые структуры данных и программные интерфейсы, разрабатывает алгоритмы и оценивает эффективность их использования | Итоговый тест Итоговое контрольное мероприятие | контролируются в форме теста знания курса |

Спецификация мероприятий текущего контроля

основы хранилищ данных

Продолжительность проведения мероприятия промежуточной аттестации: **1 часа**

Условия проведения мероприятия: **в часы самостоятельной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **0**

Проходной балл: **0**

| Показатели оценивания | Баллы |
|---|--------------|
| студент знает основы SQL | 4 |
| студент понимает основы СУБД | 3 |
| студент знает основы реляционной модели | 3 |

проектирование хранилищ данных

Продолжительность проведения мероприятия промежуточной аттестации: **1 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **15**

Проходной балл: **7**

| Показатели оценивания | Баллы |
|--|--------------|
| спроектировано хранилище данных, представлена ER-диаграмма | 10 |
| корректно выбраны измерения/факты, корректно осуществлен выбор версионных и иерархических измерений. студент ориентируется в понятиях OLAP | 5 |

Технологические аспекты построения реляционных хранилищ данных

Продолжительность проведения мероприятия промежуточной аттестации: **1 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **15**

Проходной балл: **7**

| Показатели оценивания | Баллы |
|-----------------------|-------|
| задание 2 | 8 |
| задание 1 | 7 |

Разработка ETL модулей

Продолжительность проведения мероприятия промежуточной аттестации: **1 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **20**

Проходной балл: **10**

| Показатели оценивания | Баллы |
|-----------------------|-------|
| задание 2 | 10 |
| Задание 1 | 10 |

Системы бизнес аналитики

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **20**

Проходной балл: **10**

| Показатели оценивания | Баллы |
|---|-------|
| Задание 2. Реализован отчет, без диаграмм | 10 |
| Задание 2: реализованы диаграммы, студент владеет инструментом в полной мере: может по запросу преподавателя быстро менять структуру отчета | 5 |
| задание 1 | 5 |

Итоговый тест

Продолжительность проведения мероприятия промежуточной аттестации: **2 часа**

Условия проведения мероприятия: **в часы аудиторной работы**

Максимальный балл, выставляемый за мероприятие промежуточной аттестации: **30**

Проходной балл: **15**

| Показатели оценивания | Баллы |
|-----------------------|-------|
| Вопрос 01 | 1 |
| Вопрос 30 | 1 |
| Вопрос 03 | 1 |
| Вопрос 04 | 1 |
| Вопрос 05 | 1 |
| Вопрос 06 | 1 |
| Вопрос 07 | 1 |
| Вопрос 08 | 1 |
| Вопрос 09 | 1 |
| Вопрос 10 | 1 |

| | |
|-----------|---|
| | |
| Вопрос 11 | 1 |
| Вопрос 12 | 1 |
| Вопрос 13 | 1 |
| Вопрос 14 | 1 |
| Вопрос 15 | 1 |
| Вопрос 16 | 1 |
| Вопрос 17 | 1 |
| Вопрос 18 | 1 |
| Вопрос 19 | 1 |
| Вопрос 20 | 1 |
| Вопрос 21 | 1 |
| Вопрос 22 | 1 |
| Вопрос 23 | 1 |
| Вопрос 24 | 1 |
| Вопрос 25 | 1 |
| Вопрос 26 | 1 |
| Вопрос 27 | 1 |
| Вопрос 28 | 1 |
| Вопрос 29 | 1 |
| Вопрос 02 | 1 |