Программа учебной дисциплины «Управление данными»

Утверждена
Академическим руководителем

_____ Н.В. Асеева

_____ 20_____

Автор	Маслова Е.А.
Число кредитов	8
Контактная работа (час.)	36
Самостоятельная работа (час.)	268
Курс	2
Формат изучения дисциплины	без использования онлайн курса

I. ЦЕЛЬ, РЕЗУЛЬТАТЫ ОСВОЕНИЯ ДИСЦИПЛИНЫ И ПРЕРЕКВИЗИТЫ

Целями освоения дисциплины "Управление данными" являются формирование у студентов четкого представления места и роли современных систем управления базами данных, освоение теоретических основ моделирования и обработки информации, понимание тенденций развития отрасли и направления перспективных исследований, изучение студентами принципов построения современных систем поддержки принятия решений на основе технологий Data Warehousing и OLAP, разработки информационно-поисковых систем.

В результате освоения дисциплины студент должен:

знать:

- сложившуюся в отечественной и зарубежной практике терминологию баз данных, виды информационных моделей и соответствующее языковое обеспечение, основные типы систем управления базами данных, их архитектуру, функции и принципы использования СУБД в организации, математические методы, влияющие на принципы разработки СУБД;

уметь:

- применять полученные знания к решению вопросов проектирования логической структуры баз данных, формированию запросов к реляционным базам данных на языке SQL;

владеть:

- навыками формулирования запросов на языке SQL, работы в многопользовательской СУБД с архитектурой «клиент-сервер», информационного моделирования предметной области и интеграции баз данных с существующими информационными системами на предприятии.

Настоящая дисциплина относится к базовой части дисциплин профессионального цикла, обеспечивающих подготовку бакалавра. Изучается на 3-м курсе в 1-4 модулях.

Изучение данной дисциплины базируется на следующих дисциплинах:

- Теоретические основы информатики;
- Программирование;

Для освоения учебной дисциплины, студенты должны владеть следующими знаниями и компетенциями:

- основные принципы управления информационными системами;
- современные методы проектирования и реализации информационных систем;
- основные алгоритмы и структуры данных для быстрого поиска информации;
- программирование на языках Java и C++

Основные положения данного курса используются при изучении дисциплин "Архитектура предприятия", "Информационная безопасность", "Корпоративные информационные системы" и написании КР и ВКР.

II. СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

Раздел 1 Назначение и основные принципы архитектуры систем управления базами данных. Понятие модели данных.

Содержание тем: Предмет и задачи информационного моделирования. Понятие информационной системы. Задачи систем обработки данных. Различные уровни представления данных. Эволюция методов унификации хранения и доступа к информации на примере файлов. Причины перехода к СУБД. Терминология СУБД: банк и база данных, ЯОД, ЯМД. Эталонная архитектура СУБД. Категории СУБД, различия и возможности. Категории пользователей СУБД (администраторы, непостоянные пользователи, конечные пользователи, прикладные программисты, информационные аналитики), их требования к СУБД и необходимые навыки. Жизненный цикл базы данных. Модели данных: логические структуры и механизмы/методы работы. Требования к моделям данных. Концептуальные модели данных. Модель "сущность-связь". Диаграммы классов, сценариев использования UML.

Раздел 2 Теоретические основы реляционных систем управления базами данных. Основные принципы проектирования структуры реляционных баз данных.

Содержание тем: Реляционная модель данных. Реляционная алгебра. Терминология реляционных отношений. Реляционные операции манипулирования данными. Инфологическое и даталогическое проектирование. Понятие нормальных форм в реляционной модели. Потребность в объектно-реляционном отображении для объектно-ориентированных информационных систем. Алгоритмы объектно-ориентированного отображения. Коллекция идентичности.

Раздел 3 Языковое обеспечение современных систем управления базами данных

Содержание тем: Требования к языковым средствам СУБД. История. Классификация. Дополнительные сведения о SQL. Типа данных PostgreSQL 9 и Oracle XE 11g. Основные инструкции SQL в PostgreSQL 9 и Oracle XE 11g. Интерфейсы СУБД с прикладными программами. JDBC, Embedded SQL. Клиент-серверная организация приложений на основе технологий баз данных. Активный сервер баз данных. Понятие хранимой процедуры (persistent stored module). Языки PL/pgSQL СУБД PostgreSQL и PL/SQL СУБД Oracle. Понятие целостности данных. Триггеры, их реализация в PostgreSQL 9 и Oracle XE 11g.

Раздел 4 Современные технологии распределенных систем управления базами данных

Содержание тем: Основные элементы архитектуры «клиент-сервер». Архитектура промышленных СУБД, различные варианты КС. Особенности многопользовательской работы. Работа в гетерогенных системах (ODBC, IDAPI). Понятие распределенной СУБД. Основные принципы размещения схемы распределенной БД. Преимущества и недостатки распределенных СУБД. Понятие транзакции, основные свойства транзакций. Алгоритмы поддержания целостности и восстановления данных на основе механизма транзакций. Мониторы транзакций. Основные стандарты на транзакционную обработку данных в современных информационных системах. Алгоритм двухфазной блокировки.

Раздел 5 Хранилища данных

Содержание тем: Понятие интеллекта бизнеса (Business Intelligence) и его воплощение в технологии DataWarehousing. Основные сведения о базах данных и экспертных системах. Концепция распространения информации. Основные свойства хранимых и обрабатываемых данных. Разные категории программных систем (DataWarehouse и Data Mart). Аналитическая система Basegroup Deductor. Практическая реализация технологий Data Waregousing в системе Deductor.

Раздел 6 Перспективные направления развития технологии баз данных. NoSQL

Содержание тем: Недостатки современных СУБД. Постреляционная модель данных. Объектно-ориентированные СУБД. Ограничения SQL-ориентированных СУБД. Основные виды NoSQL СУБД. Распределенное хранение данных, sharding. Агрегатные модели данных. Хранилища «ключ-значение», документно-ориентированные и столбцовые СУБД. Графовые модели данных.

III. ОЦЕНИВАНИЕ

Промежуточная оценка $О_{промежуточная}$ рассчитывается следующим образом

$$О_{промежуточная} = 0,6 \cdot О_{накопленная} + 0,4 \cdot О_{промежуточный\ экзамен}$$

Накопленная оценка учитывает результаты студента по текущему контролю следующим образом:

$$О_{накопленная} = 0,7 \cdot О_{текущий} + 0,15 \cdot О_{ауд} + 0,15 \cdot О_{сам.работа}$$

где $О_{текущий}$ рассчитывается как взвешенная сумма всех форм текущего контроля

$$О_{текущий} = 0,25 \cdot О_{ауд.} + 0,75 \cdot О_{д/з};$$

Результирующая оценка $О_{результ}$ за дисциплину рассчитывается следующим образом:

$$О_{результ} = 0,8 \cdot (О_{промежуточная} + О_{накопленная}):2 + 0,2 \cdot О_{экзамен}$$

Полученные после округления этих величин до целого значения ***выставляются в диплом*** как ***результирующие оценки по 10-балльной шкале.***

IV. ПРИМЕРЫ ОЦЕНОЧНЫХ СРЕДСТВ

Оценочные средства для текущего контроля студента
Примерные задания *для домашнего задания:*

1. Типичные информационные запросы в информационной системе по книжной торговле.
2. Типичные информационные запросы в информационной системе бюро по трудоустройству.
3. Типичные информационные запросы в информационной системе пункта проката.
4. Типичные информационные е запросы в информационной системе медицинского учреждения.
5. Типичные информационные запросы в информационной системе отдела кадров.
6. Типичные информационные запросы в информационной системе отдела заказов.
7. Типичные информационные запросы в информационной системе морского порта
8. Типичные информационные запросы в информационной системе автопредприятия.
9. Типичные информационные запросы в информационной системе авторемонтного предприятия.
10. Типичные информационные запросы в информационной системе отдела мобильных продаж.
11. Типичные информационные запросы в информационной системе дорожно-ремонтной компании.
12. Типичные информационные запросы в информационной системе отделения потреб. кооперации.
13. Типичные информационные запросы в информационной системе железной дороги.
14. Типичные информационные запросы в информационной системе сетевого маркетинга.
15. Типичные информационные запросы в информационной системе авиа-предприятия.
16. Типичные информационные запросы в информационной системе тур. оператора.
17. Типичные информационные запросы в информационной системе распределенного вычислительного центра.
18. Типичные информационные запросы в информационной системе страховой фирмы.
19. Типичные информационные запросы в информационной системе по обслуживанию банкоматов.
20. Типичные информационные запросы в информационной системе мелкой производственной фирмы.
21. Типичные информационные запросы в информационной системе снабжения.
22. Типичные информационные запросы в информационной системе лесного хозяйства.
23. Типичные информационные запросы в информационной системе поликлиники.
24. Типичные информационные запросы в информационной системе театра.
25. Типичные информационные запросы в информационной системе садоводческого товарищества.
26. Типичные запросы в информационной системе управления электропотреблением.

Тематика *контрольной работы*: применение аппарата реляционной алгебры и реляционного исчисления для составления запросов к базе данных клиентов торговой системы

Темы *коллоквиума* для каждого студента утверждается преподавателем в индивидуальном порядке.

Оценочные средства для промежуточной аттестации

Примерный перечень вопросов к экзамену (итоговый контроль):

1. Понятие информации и информационных процессов. Роль информации в современном мире.
2. Понятие информационно-поисковой системы, классификация ИПС. Роль ИПС в моделировании информационных процессов.
3. Различные уровни представления данных в фактографических информационных системах. Понятие внешних и внутренних моделей данных, представлений.
4. Эволюция методов унификации хранения и доступа к информации в информационных системах. Причины перехода от позадачного подхода в обработке данных к системам управления базами данных.
5. Основные требования к архитектуре СУБД. Трехуровневая архитектура ANSI/SPARC. Главные компоненты и их назначение.
6. Терминология СУБД: банк и база данных. Категории СУБД, различия и возможности.
7. Категории пользователей СУБД, их требования к СУБД и необходимые навыки.
8. Понятие модели данных. Различные примеры моделей данных. Требования к моделям данных различных групп пользователей.
9. Терминология моделей данных: информационный объект, атрибут, отношение. Виды отношений. Основные группы моделей данных.
10. Иерархическая модель данных. Методы описания объектов и связей между ними. Реализация иерархической модели данных в системе IMS.
11. Сетевая модель данных. Методы описания объектов и связей между ними. Предложения группы CODASYL по практической реализации сетевой модели.
12. Практическая ценность ранних моделей данных (сетевой и иерархической). Основные недостатки.
13. Реляционная модель данных. Основные термины и правила над реляционными отношениями.
14. Ключ и суперключ отношения. Неизвестные значения в отношениях. Правила работы с неизвестными отношениями.
15. Принципы моделирования связей объектов в реляционной модели. Понятия кандидатного, первичного и внешнего ключа.
16. Понятие реляционной алгебры. Язык и принципы работы с реляционными отношениями. Методы использования реляционной алгебры для оптимизации запросов в СУБД.
17. Теоретико-множественные операции реляционной алгебры. Определения, примеры.
18. Реляционные операции реляционной алгебры. Определения, примеры.
19. Понятие полной и избыточной системы операций. Примеры полных неизбыточных систем операций реляционной алгебры.
20. Реляционное исчисление с переменными-кортежами. Формальное определение. Редукция реляционной алгебры к реляционному исчислению с переменными-кортежами.

21. Использование теории предикатов первого порядка для преобразования выражений реляционного исчисления с переменными-кортежами (редукция кванторов всеобщности и оператора логического следствия). Практические примеры на языке SQL.

22. Стандартизация языка SQL- причины, прошлое, настоящее, будущее. Основные организации, занимающиеся стандартизацией SQL. Основные производители SQL-ориентированных СУБД.

23. Основная операция определения схемы и способы задания ограничений целостности в языке SQL. Стандартные типы данных.

24. Структура запросов в языке SQL. Основные элементы запроса (список выборки, блок FROM, WHERE, GROUP BY, HAVING, ORDER BY).

25. Операции модификации данных в языке SQL (INSERT, DELETE, UPDATE). Их работа в присутствии введенных ограничений целостности.

26. Операция определения пользовательского представления CREATE VIEW. Назначение, особенности использования.

27. Назначение, и способы использования агрегатных функций в языке SQL. Способы введения новых агрегатных функций в СУБД Postgresql и Oracle.

28. Понятие курсора в SQL. Причины появления механизма курсора. Его отличие от оператора выборки. Специфичные операции с курсором.

29. Назначение хранимых процедур в языке SQL. Способы реализации и использования на примере СУБД Postgresql и Oracle (операция CREATE FUNCTION).

30. Назначение внедренного SQL (Embedded SQL). Принципы использования и сопутствующий инструментарий на примере СУБД Postgresql.

31. Назначение правил в языке SQL. Способы реализации и использования на примере СУБД Postgresql и Oracle (операция CREATE RULE).

32. Способы ограничения прав доступа к реляционным данным на примере операции SQL GRANT/REVOKE.

33. Принципы организации СУБД при использовании технологии «клиент-сервер». Задачи, выполняемые клиентским и серверным приложением. Функции пользователей. Архитектура СУБД Postgresql.

34. Понятие транзакции в СУБД. «Замечательные» свойства транзакций. Возможности, предоставляемые механизмом транзакций для прикладных программистов. Понятие сохранения (commit) и отката (rollback) транзакции.Реализация транзакций в SQL.

35. Восстановление информации в СУБД на основе механизма транзакций. Основные понятия («грязные» страницы, журнал транзакций, точка синхронизации, откат, повторный запуск). Алгоритмы реализации.

36. Организация параллельного доступа к данным в СУБД на основе механизма транзакций. Аномалии доступа, блокировки, взаимный захват. Управление уровнем защиты транзакции в SQL.

37. Недостатки реляционной модели данных. Другие причины появления новых моделей данных. Основные отличия объектно-реляционной модели данных и конкретная реализация на примере СУБД Postgresql и Oracle.

38. Понятие объектно-реляционного отображения. Алгоритмы реализации

39. Назначение документальных информационных систем. Их отличия от фактографических ИС. Основные понятия: информационно-поисковый язык (виды ИПЯ), поисковый образ документа, поисковый образ запроса, дескриптор.

40. Причины широкого распространения, назначение и особенности Хранилищ Данных (data warehouse) и распределенных служб каталогов (напр. LDAP).

41. Отличия аналитических (OLAP) и операционных (OLTP) систем баз данных

42. SnowFlake-схема и Star-схема проектирования хранилищ данных.

V. РЕСУРСЫ

5.1 Основная литература

1. Голицына, О.Л. Базы данных [электронный ресурс]: учебное пособие / О.Л.Голицына, Н.В.Максимов, И.И.Попов; ЭБС Znanium. - 2-е изд., испр. и доп. - М.: Форум: ИНФРА-М, 2009. - 400 с.: ил. - (Профессиональное образование). - ISBN 978-5-91134-098-8. - Режим доступа: http://znanium.com/bookread.php?book=182482. - Загл. с экрана. Гриф МО РФ

5.2 Дополнительная литература

1. Общие принципы построения интеллектуальных систем поддержки принятия решений: учебник / Э.А.Бабкин [и др.]; НГТУ им. Р.Е.Алексеева, Гос. ун-т - Высшая школа экономики, Нижегор. фил. - Н. Новгород: Изд-во НГТУ, 2008. - 269 с. Гриф МО РФ
2. Паклин, Н.Б. Бизнес-аналитика: от данных к знаниям / Н.Б.Паклин, В.И.Орешков. - СПб.: Питер, 2009. - 624 с. + 1опт. диск (CD-ROM): аналитическая платформа Deductor Academic.
3. Фейерштейн, С. Oracle PL/SQL для профессионалов / С.Фейерштейн, Б.Прибыл; пер. с англ. О.Здир. - 3-е изд. - СПб.: Питер, 2005. - 941 с.
4. Голицина, О.Л. Базы данных: учебное пособие / О.Л.Голицина, Н.В.Максимов, И.И.Попов. - М.: ФОРУМ, ИНФРА-М, 2006

5.3 Программное обеспечение

№ п/п	Наименование	Условия доступа
1.	Microsoft Office 2013 Prof + Microsoft Visual Studio 2013 Prof	*Государственный контракт*
2.	MongoDB, HBase, Cassandra, Node4j, Redis, Riak (на выбор)	*Свободно-распространяемые СУБД*
3.	PostgreSQL, Oracle XE, MySQL, MS SQL Server Compact (на выбор)	*Свободно-распространяемые реляционные СУБД*

5.4 Профессиональные базы данных, информационные справочные системы, интернет-ресурсы (электронные образовательные ресурсы)

№ п/п	Наименование	Условия доступа
	Профессиональные базы данных, информационно-справочные системы	
1.	Интернет-портал Basegroup, свободно распространяемая версия аналитической платформы Deductor Academic	URL: http://www.basegroup.ru
2.	Oracle Database 11g Express Edition (XE)	URL: http://www.oracle.com/technetwork/database/express-edition/overview/index.html
3.	Сайт СУБД PostgreSQL	URL: http://www.postgresql.org/
4.	Сайт СУБД MySQL	URL: http://www.mysql.com/
5.	Java SE Downloads	URL: http://www.oracle.com/technetwork/java/javase/downloads/index.html

5.5 Материально-техническое обеспечение дисциплины

Лекционные занятия проходят в аудиториях, оборудованных следующим мультимедийным оборудованием: преподавательским компьютером (или ноутбуком), экраном, проектором.

Практические занятия проходят в компьютерных классах, оснащенных преподавательским компьютером, персональными компьютерами, объединенными в локальную сеть с возможностью выхода в интернет.