


Утверждено приказом ГОУ ВПО ДонГУУ от 23.08.2016г. №675

ДОНЕЦКАЯ НАРОДНАЯ РЕСПУБЛИКА
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«ДОНЕЦКАЯ АКАДЕМИЯ УПРАВЛЕНИЯ И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ГЛАВЕ ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ»

ФАКУЛЬТЕТ ГОСУДАРСТВЕННОЙ СЛУЖБЫ И УПРАВЛЕНИЯ
КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

УТВЕРЖДАЮ
Проректор по учебной работе

Л.Н.Костина
20.06.2017г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

«Информационные хранилища»

Направление подготовки 09.04.03 «Прикладная информатика»

Донецк
2017

Рабочая программа учебной дисциплины «Информационные хранилища» для студентов 1 курса образовательного уровня «магистр» направления подготовки 09.04.03 «Прикладная информатика» очной формы обучения.

Автор(ы),

разработчик(и): Ст.преподаватель Е.Г.Литвак

должность, ученая степень, ученое звание, инициалы и фамилия

Программа рассмотрена на
заседании ПМК кафедры

«Прикладная информатика»

Протокол заседания ПМК от

08.06.2017 № 10

Председатель ПМК



А.Н.Верзилов

Программа рассмотрена на
заседании кафедры

Информационных технологий

Протокол заседания кафедры от

09.06.2017 № 13

Заведующий кафедрой



дата

Н.В.Брадуд

1. Цель освоения дисциплины и планируемые результаты обучения по дисциплине (соотнесенные с планируемыми результатами освоения образовательной программы)

дать знания по принципам построения информационных хранилищ, внедрению специализированных платформ (Data Warehouse) ориентированных на поддержку решений бизнес аналитики. Студенты получают навыки создания хранилища данных, внедрение ETL (Извлечение, преобразование и загрузка), а также изучат технологии комплексного многомерного анализа (OLAP).

Предполагаемые результаты:

Код соответствующей компетенции по ГОС	Наименование компетенций	Результат освоения (знать, уметь, владеть)
1	2	3
ПК-13	Способность проектировать информационные процессы и системы с использованием инновационных инструментальных средств, адаптировать современные ИКТ к задачам прикладных ИС.	<p>Знать: основные определения, относящиеся к концепции управления хранилищами данных.</p> <p>Уметь: проектировать многомерные кубы данных</p> <p>Владеть: Владеть методами реализации хранилищ данных</p>
ПК-17	Способность управлять информационными ресурсами и ИС	<p>Знать:</p> <ul style="list-style-type: none"> – основные требования (и средства их обеспечения) к хранилищам данных; – технологии, обеспечивающие манипулирование хранилищами данных. <p>Уметь:</p> <p>реализовать средства, обеспечивающие:</p> <ul style="list-style-type: none"> – предоставление пользователю результатов анализа за приемлемое время; – возможность осуществления любого логического и статистического анализа, характерного для данного приложения, и его сохранения в доступном для конечного пользователя виде; – многопользовательский доступ к данным с поддержкой соответствующих механизмов блокировок и средств авторизованного доступа <p>Владеть: Владеть методами реализации хранилищ данных.</p>

ПК-30	Способность анализировать данные средствами многомерного анализа	<p>Знать: архитектуры данных и средства многомерного анализа правила корректной подготовки данных.</p> <p>Уметь:</p> <ul style="list-style-type: none"> – выполнять преобразования данных для подготовки к анализу; – выявлять дубликаты и противоречия. <p>Владеть: программными средствами многомерного анализа.</p>
-------	--	---

2. Место дисциплины в структуре основной образовательной программы

Дисциплина «Информационные хранилища» относится к обязательным дисциплинам вариативной части профессионального цикла учебного плана по направлению подготовки 09.04.03 «Прикладная информатика».

2.1. Требования к предварительной подготовке обучающегося

Перед изучением данной дисциплины студентами должны быть изучены дисциплины «Базы данных», «Корпоративные информационные системы», «Проектирование информационных систем».

2.2. Дисциплины и/или практики, для которых освоение данной дисциплины необходимо как предшествующее:

Дисциплина Информационные хранилища должна быть освоена перед написанием магистерской диссертации.

3. Объем дисциплины в кредитах (зачетных единицах) с указанием количества академических часов, выделенных на аудиторную (по видам учебных занятий) и самостоятельную работу студента

Вариант таблицы при изучении дисциплины в одном семестре

	Зачетные единицы (кредиты ECTS)	Всего часов		Форма обучения (вносятся данные по реализуемым формам)	
		О	З	Очная	Заочная
				Семестр №2	Семестр №
Общая трудоемкость	3	108	X	Количество часов на вид работы:	
Виды учебной работы, из них:					
Аудиторные занятия (всего)				42	X
В том числе:					
Лекции				14	X
Семинарские занятия / Практические занятия				28	X
Самостоятельная работа (всего)				66	X
Промежуточная аттестация					
В том числе:					
зачет /экзамен				экзамен	зачет /экзамен

4. Содержание дисциплины, структурированное по разделам (темам) с указанием отведенного на них количества академических часов и видов учебных занятий

4.1. Разделы (темы) дисциплины с указанием отведенного на них количества академических часов и видов учебных занятий

Наименование раздела, темы дисциплины	Виды учебной работы (бюджет времени) (вносятся данные по реализуемым формам)									
	Очная форма обучения					Заочная форма обучения				
	Лекции	Практические занятия	Семинарские занятия	Самостоятель- ная работа	Всего	Лекции	Практические занятия	Семинарские занятия	Самостоятель- ная работа	Всего
1	2	3	4	5	6	7	8	9	10	11
Раздел 1.										
Тема 1.1.	2	4		9	15					
Тема 1.2.	2	4		9	15					
Тема 1.3.	2	4		9	15					
Итого по разделу:	6	12		27	45					
Раздел 2.										
Тема 2.1.	2	4		9	15					
Тема 2.2.	2	4		10	16					
Тема 2.3.	2	4		10	16					
Тема 2.4.	2	4		10	16					
Итого по разделу:	8	16								
Всего за семестр:	14	28		36	63					

4.2. Содержание разделов дисциплины:

Наименование раздела, темы дисциплины	Содержание разделов дисциплины	Содержание семинарских/практических занятий		
			Кол-во часов	
			0	3
1	2	3	4	5
Раздел 1.				
Тема 1.1.	Архитектуры данных: Базы данных и модели данных.	Практическое занятие №1-2	4	
		1. Лабораторная работа: Формирование необходимых навыков работы с Deductor – мастер-класс	2	
		2. Лабораторная работа: Формирование необходимых навыков работы с Deductor – мастер-класс	2	
Тема 1.2.	Многомерные данные. Концепция хранилищ данных (ХД). OLAP как ключевой компонент ХД.	Практическое занятие №3-4	4	
		1. Лабораторная работа: Многомерные наборы данных –	2	

Наименование раздела, темы дисциплины	Содержание разделов дисциплины	Содержание семинарских/практических занятий		
			Кол-во часов	
			0	3
1	2	3	4	5
		использование платформы Deductor		
		2. Лабораторная работа: Агрегированные показатели – использование платформы Deductor	2	
Тема 1.3.	Архитектуры хранилищ данных. Реляционные хранилища данных (ROLAP – Relational OLAP). Многомерные хранилища данных (MOLAP – Multidimensional OLAP). Гибридные хранилища данных (HOLAP – Hybrid OLAP), Loading).	Практическое занятие №5-6	4	
		1.Лабораторная работа. Проектирование архитектуры хранилища данных	2	
		Проектирование архитектуры хранилища данных	2	
РАЗДЕЛ 2 Проект ХД для выбранной предметной области.				
Тема 2.1.	Основные методы трансформации данных, загружаемых в хранилище. Введение в трансформацию данных. Особенности трансформации временных рядов. Группировка и разгруппировка данных. Слияние данных. Квантование. Нормализация и кодирование данных. Извлечение данных в ETL. Преобразование данных в ETL. Загрузка данных в хранилище данных	Практическое занятие №7-8	4	
		1.Лабораторная работа. Трансформация данных в Deductor Studio.	2	
		2. Лабораторная работа. Трансформация данных в Deductor Studio.	2	
Тема 2.2.	Виртуальные хранилища данных. Введение в визуализацию данных. Визуализаторы общего назначения. OLAP-анализ. Визуализаторы, применяемые для оценки качества моделей. Визуализаторы, применяемые для интерпретации результатов анализа	Практическое занятие №9-10	4	
		1.Лабораторная работа. Визуализация данных в Deductor Studio	2	
		2. Лабораторная работа. Визуализация данных в SQL Server	2	
Тема 2.3.	Оценка качества, очистка и предобработка данных, загружаемых в хранилище. Введение в оценку качества	Практическое занятие №11-12	4	
		1. Лабораторная работа. Использование	2	

Наименование раздела, темы дисциплины	Содержание разделов дисциплины	Содержание семинарских/практических занятий		
			Кол-во часов	
			0	3
1	2	3	4	5
	данных (ADQ – Assessment Data Quality). Технологии и методы оценки качества данных. Очистка и предобработка данных. Фильтрация данных.	Deductor Studio для оценки качества, очистки и предобработки данных. 2. Лабораторная работа. Использование Deductor Studio для оценки качества, очистки и предобработки данных.	2	
Тема 2.4.	Обработка дубликатов и противоречий. Выявление аномальных значений. Восстановление пропущенных значений. Введение в сокращение размерности. Сокращение числа признаков. Сокращение числа значений признаков и записей. Сэмплинг	Практическое занятие №13-14	4	
		1.Лабораторная работа. Использование Deductor Studio для оценки качества, очистки и предобработки данных.	2	
		2. Лабораторная работа. Использование Deductor Studio для оценки качества, очистки и предобработки данных.	2	

5. Перечень учебной литературы, необходимой для освоения дисциплины

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

Раздел «Информационные хранилища» в СДО «Moodle» ГОУ ВПО «ДонАУиГС»:

Режим доступа: elearn.dsum.org

Перечень контрольных вопросов для самоподготовки:

1. Модели и их свойства. Аналитический и информационный подходы к моделированию.

2. Формы представления, типы и виды анализируемых данных.
3. Обучение моделей «с учителем» и «без учителя». Обучающее и тестовое множество. Ошибки обучения. Эффект переобучения.
4. Общая схема анализа данных. Требования к алгоритмам анализа данных.
5. Основные принципы сбора (формализации) данных. Требования к объемам анализируемых данных.
6. Характеристика этапов технологии KDD.
7. Data Mining. Характеристика классов задач, решаемых методами Data Mining.
8. Программный инструментарий для выполнения анализа данных.

9. Цели, задачи и основное содержание консолидации данных. Обобщенная схема процесса консолидации.
10. Характеристика OLTP-систем.
11. Предпосылки появления систем поддержки принятия решений DSS. Понятие ESS, EIS и GDSS.
12. Основные положения концепции хранилищ данных (DW).
13. Реляционные хранилища данных (ROLAP).
14. Технология OLAP. Сущность многомерного представления данных.
15. Структура многомерного куба. Работа с измерениями.
16. Многомерные хранилища данных (MOLAP).
17. Гибридные хранилища данных (HOLAP).
18. Виртуальные хранилища данных.
19. Цели, задачи и основное содержание процесса ETL.
20. Основные виды проблем в данных, из-за которых они нуждаются в очистке.
21. Организация процесса загрузки данных в хранилище. Постзагрузочные операции.
22. Причины отказа от использования хранилищ данных. Особенности загрузки данных из локальных источников.
23. Обогащение данных.
24. Цели, задачи и основное содержание трансформации данных. Трансформация данных на разных этапах аналитического процесса. Типичные средства трансформации.
25. Особенности трансформации временных рядов. Скользящее окно. Преобразование даты и времени.
26. Группировка и разгруппировка данных.
27. Способы слияния данных.
28. Квантование данных.
29. Нормализация и кодирование данных.
30. Цели, задачи и основное содержание визуализации данных. Группы методов визуализации.
31. Визуализаторы общего назначения. OLAP-анализ.
32. Манипуляции с измерениями OLAP-куба.
33. Визуализаторы, применяемые для оценки качества моделей.
34. Визуализаторы, применяемые для интерпретации результатов анализа.
35. Технологии и методы оценки качества данных. Профайлинг.
36. Очистка и предобработка данных.
37. Типичный набор инструментов предобработки данных в аналитическом приложении.
38. Фильтрация данных. Обработка дубликатов и противоречий.
39. Выявление аномальных и восстановление пропущенных значений.
40. Алгоритмы и методы сокращения числа признаков.

5.2. Перечень основной учебной литературы

1. Федин Ф.О. Анализ данных. Часть 1: Подготовка данных к анализу. Учебное пособие. – М: МГПУ, 2012. – 204 с.
2. Харинатх, Сивакумар, Кэррол и др. Microsoft SQL Server 2008 Analysis Services и MDX для профессионалов. Пер. с англ. – М.: ООО «И.Д. Вильямс», 2010. – 1072 с.
3. Маклаков С.В. Создание информационных систем с AllFusion Modeling Suite. – М.: Издательство Диалог-МИФИ, 2007 – 400 с.
4. Туманов В.Е. Основы проектирования реляционных баз данных [Электронный ресурс]/ Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. – Режим доступа: <http://www.iprbookshop.ru/52221.html>

5. Алексеева Е.В., Амириди Ю.В., Дик В.В. Информационные аналитические системы [Электронный ресурс]: учебник / Т. В. Алексеева, Ю. В. Амириди, В. В. Дик и др.; под ред. В. В. Дика. - М.: МФПУ Синергия, 2013. - 384 с. - (Университетская серия). Режим доступа: <http://znanium.com/catalog.php?bookinfo=451186>

5.3. Перечень дополнительной литературы

1. Туманов В.Е., Маклаков С.В. Проектирование реляционных хранилищ данных / М: Диалог-МИФИ, 2007 – 333 с.
2. Архипенков С., Голубев Д., Максименко О. Хранилища данных. От концепции до внедрения / М.: Диалог-МИФИ, 2002. – 528 с.
3. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям – СПб.: Питер, 2009. – 624 с.
4. Инмон Б. Типы хранилищ данных. Перевод Intersoftlab, 2001, <http://www.iso.ru/journal/articles/181.html>
5. Кузнецов С., Артемьев В. Обзор возможностей применения ведущих СУБД для построения хранилищ данных (DataWarehouse). <http://www.citforum.ru/database/kbd98/glava15.shtml>
6. В. П. Божко, А. В. Хорошилов, В. А. Благодатских и др., Предметно-ориентированные экономические информационные системы: учебник для студ. вузов. / М. : Финансы и статистика, 2007. - 224 с. : ил.
7. Коннолли, Томас, Бегг, Карелии. Базы данных. Проектирование, реализация и сопровождение. Теория и практика. 3-е издание.: Пер. с англ. - М.: Издательский дом "Вильямс", 2003. - 1440 с. : ил.

6. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

а) полнотекстовые базы данных

Электронно-библиотечная система IPRbooks. ЭБС содержит более 20 000 учебных и научных изданий по различным дисциплинам, свыше 200 наименований российских и зарубежных журналов, большая часть которых входит в перечень ВАК. Режим доступа: <http://www.iprbookshop.ru/>

б) интернет-ресурсы

<http://www.citforum.ru/>

<http://www.basegroup.ru>

7. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем

7.1. Перечень информационных технологий (при необходимости)

Локальная сеть, Интернет.

7.2. Перечень программного обеспечения (при необходимости)

- IBM Cognos Insight – система хранения данных – хранилище данных;
- Аналитическая платформа Deductor;
- MS SQL Server / MS Access;

MS Excel

7.3. Перечень информационных справочных систем (при необходимости)

В том случае, когда программное обеспечение или информационные справочные системы не применяются при изучении дисциплины, следует указать «Программное

обеспечение не применяется» и «Информационные справочные системы не используются»

8. Фонд оценочных средств для контроля уровня сформированности компетенций

8.1. Виды промежуточной аттестации

Текущий контроль успеваемости позволяет оценить уровень сформированности элементов компетенций (знаний и умений), компетенций с последующим объединением оценок и проводится в форме устного опроса (фронтальный, индивидуальный, комплексный), письменной проверки (ответы на вопросы, тестовые задания), включая задания для самостоятельной работы.

Промежуточная аттестация по итогам освоения дисциплины в форме зачета, позволяет оценить уровень сформированности компетенций и может осуществляться по результатам текущего контроля и итоговой контрольной работы, тестовых заданий и т.п.

Промежуточная аттестация в форме экзамена позволяет оценить уровень сформированности компетенций в целом по дисциплине и может осуществляться как в письменной так и в устной форме.

8.2. Показатели и критерии оценки результатов освоения дисциплины

Средним баллом за дисциплину является средний балл за текущую учебную деятельность.

Механизм конвертации результатов изучения студентом дисциплины в оценки по государственной шкале и шкале ECTS представлен в таблице.

Средний балл по дисциплине (текущая успеваемость)	Отношение полученного студентом среднего балла по дисциплине к максимально возможной величине этого показателя	Оценка по государственной шкале	Оценка по шкале ECTS	Определение
4,5 – 5,0	90% – 100%	5	A	отлично – отличное выполнение с незначительным количеством неточностей (до 10%)
4,0 – 4,49	80% – 89%	4	B	хорошо – в целом правильно выполненная работа с незначительным количеством ошибок (до 20%)
3,75 – 3,99	75% – 79%	4	C	хорошо – в целом правильно выполненная работа с незначительным количеством ошибок (до 25%)
3,25 – 3,74	65% – 74%	3	D	удовлетворительно – неплохо, но со

				значительным количеством недостатков (до 35%)
3,0 – 3,24	60% – 64%	3	Е	достаточно – выполнение удовлетворяет минимальные критерии, но со значительным количеством недостатков (до 40%)
до 3,0	35% – 59%	2	FX	неудовлетворительно с возможностью повторной сдачи (ошибок свыше 40%)
	0 – 34%	2	F	неудовлетворительно – надо поработать над тем, как получить положительную оценку (ошибок свыше 65%)

8.3. Критерии оценки работы студента

При усвоении каждой темы за текущую учебную деятельность студента выставляются оценки по 5-балльной (государственной) шкале. Оценка за каждое задание в процессе текущей учебной деятельности определяется на основе процентного отношения операций, правильно выполненных студентом во время выполнения задания:

- 90-100% – «5»,
- 75-89% – «4»,
- 60-74% – «3»,
- менее 60% – «2».

Если на занятии студент выполняет несколько заданий, оценка за каждое задание выставляется отдельно.

8.3.1. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или опыта деятельности, характеризующие этапы формирования компетенций в процессе освоения образовательной программы)

Приводятся конкретные примеры типовых заданий из оценочных средств, определенных в рамках данной дисциплины для проведения текущей, промежуточной аттестации по пунктам:

- а) типовые вопросы и/или задания;
- б) критерии оценивания компетенций (результатов) по уровням освоения учебного материала: 1 – репродуктивный (освоение знаний, выполнение деятельности по образцу, инструкции или под руководством); 2 – продуктивный (планирование и самостоятельное выполнение деятельности, решение проблемных задач; применение умений в новых условиях); 3 – творческий (самостоятельное проектирование экспериментальной деятельности; оценка и самооценка инновационной деятельности);
- в) критерии оценивания контрольных вопросов, в том числе, для самоподготовки;
- г) описание шкалы оценивания.

8.3.2. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и/или опыта деятельности

Методические материалы, определяющие процедуры оценивания, могут включать в себя следующие основные элементы:

- когда проводится оценивание;
- кто проводит оценивание;
- как предъявляются задания;
- кто собирает и обрабатывает материалы;
- кто и когда предъявляет результаты оценивания;
- и т.п.

9. Методические рекомендации (указания) для обучающихся по освоению дисциплины

Методические рекомендации (указания) для обучающихся по освоению дисциплины (рекомендуемый режим и характер учебной работы, в том числе в части выполнения самостоятельной работы) – комплекс рекомендаций и разъяснений, позволяющий обучающимся оптимальным образом организовать процесс изучения как теоретического учебного материала дисциплины, так и подготовки к семинарским, практическим занятиям, в том числе проводимым с использованием **активных и интерактивных технологий обучения**.

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

В разделе указываются используемые при изучении дисциплины специализированные лаборатории и кабинеты с оборудованием, компьютерные классы, лекционные аудитории, оснащенные мультимедийным оборудованием и т.п., имеющиеся в ГОУ ВПО «ДонАУиГС».