

МИНОБРНАУКИ РОССИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ВОРОНЕЖСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ВГУ»)

УТВЕРЖДАЮ

Заведующий кафедрой
функционального анализа
и операторных уравнений

 Каменский М.И.
подпись, расшифровка подписи
20.03.2025 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Б1.О.13.02 Специальные технологии баз данных и информационных систем

- 1. Код и наименование направления подготовки:** 10.05.04
Информационно-аналитические системы безопасности
- 2. Профиль подготовки/специализация:** Информационная безопасность
финансовых и экономических структур
- 3. Квалификация выпускника:** специалист по защите информации
- 4. Форма обучения:** очная
- 5. Кафедра, отвечающая за реализацию дисциплины:** функционального
анализа и операторных уравнений
- 6. Составители программы:** Болдырева Елена Сергеевна, преподаватель,
математический факультет, кафедра функционального анализа и
операторных уравнений
- 7. Рекомендована:** научно-методическим советом математического
факультета, 18.03.2025 Протокол №0500-03.
- 8. Учебный год:** 2029-2030 **Семестр:** 9

9. Цели и задачи учебной дисциплины:

Целями освоения учебной дисциплины являются:

- обучение студентов принципам разработки, эксплуатации, анализу корректности и устойчивости информационных систем и отдельных их компонентов, обрабатывающих большие потоки данных, включая применение технологии обработки информации в режиме реального времени, в том числе многомерных данных в финансовой и экономической сфере.

Задачи учебной дисциплины:

– изучение методов представления многомерных данных в аналитических системах;
– изучение методов анализа многомерных данных;
– приобретение навыков по разработке и эксплуатации приложений специального анализа данных;
– получение умений и навыков выполнения анализа корректности и устойчивости функционирования отдельных компонентов, подсистем и в целом всей информационной системы;
– получение студентами умений и навыков разработки и применения автоматизированных технологий обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени.

10. Место учебной дисциплины в структуре ООП:

Учебная дисциплина «Специальные технологии баз данных и информационных систем» относится к обязательной части блока Б1, формируемой участниками образовательных отношений.

Предшествующие дисциплины – Информатика, Языки программирования, Технология и методы программирования, Базы данных и экспертные системы.

Дисциплина рассчитана на обучающихся, обладающих знаниями в области технологии программирования и работы на ЭВМ, баз данных.

11. Планируемые результаты обучения по дисциплине/модулю (знания, умения, навыки), соотнесенные с планируемыми результатами освоения образовательной программы (компетенциями) и индикаторами их достижения:

Код	Название компетенции	Коды	Индикаторы	Планируемые результаты обучения
ОПК–2	Способен использовать знания норм права при решении задач профессиональной деятельности	ОПК-2.2	Способен выполнять анализ корректности и устойчивости функционирования отдельных компонентов и подсистем по противодействию легализации доходов, полученных	Знать: способы и методы анализа и классификации материалов финансовых расследований преступлений, связанных с легализацией (отмыванием) доходов, полученных преступным путем, финансированием терроризма, экстремизма и распространения

			преступным путем, и финансированию терроризма;	оружия массового уничтожения; Уметь: использовать современные информационные технологии для поиска информации по противодействию легализации доходов, полученных преступным путем, финансированию терроризма, экстремизма и распространения оружия массового уничтожения; Владеть: навыками использования функциональных возможностей современных информационных систем для поиска информации по противодействию легализации доходов, полученных преступным путем, финансированию терроризма, экстремизма и распространения оружия массового уничтожения.
		ОПК-2.2.2	Использует современные информационные технологии для поиска информации по противодействию легализации доходов, полученных преступным путем, финансированию терроризма, экстремизма и распространения оружия массового уничтожения	
		ОПК-2.2.3	Использует функциональные возможности современных информационных систем для поиска информации по противодействию легализации доходов, полученных преступным путем, финансированию терроризма, экстремизма и распространения оружия массового уничтожения	

		ОПК-2.4	Способен разрабатывать и применять автоматизированные технологии обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени	
		ОПК-2.4.1	Разрабатывает, эксплуатирует и анализирует корректность и устойчивость информационных систем и отдельных их компонентов, обрабатывающих большие потоки данных, включая применение технологии обработки информации в режиме реального времени, в том числе многомерных данных в финансовой и экономической сферах	Знать: способы и методы разработки и применения автоматизированных технологий обработки больших информационных потоков (массивов) финансовой и/или экономической информации в режиме реального времени; Уметь: разрабатывать, эксплуатировать и анализировать корректность и устойчивость информационных систем и отдельных их компонентов, обрабатывающих большие потоки данных, включая применение технологии обработки информации в режиме реального времени, в том числе многомерных данных в финансовой и экономической сферах;
		ОПК-2.4.2	Использует современные компьютерные технологии вычленения и исследования значимой	Владеть: навыками использования современных компьютерных технологий вычленения и исследования значимой

			информации из потоков данных	информации из потоков данных.
--	--	--	------------------------------	-------------------------------

12. Объем дисциплины в зачетных единицах/час. - 4 / 144.

Форма промежуточной аттестации – экзамен.

13. Трудоемкость по видам учебной работы

Вид учебной работы		Трудоемкость	
		Всего	По семестрам
			9 семестр
Аудиторные занятия		72	72
в том числе:	лекции	36	36
	практические	-	-
	лабораторные	36	36
Самостоятельная работа		36	36
в том числе: курсовая работа (проект)		-	-
Форма промежуточной аттестации (экзамен – 36 час.)		36	36
Итого:		144	144

13.1. Содержание дисциплины

п/п	Наименование раздела дисциплины	Содержание раздела дисциплины	Реализация дисциплины с помощью онлайн-курса, ЭУМК
1. Лекции			

1.1	Введение в курс. Системы поддержки принятия решений.	Структура курса, цели и задачи. Контрольные мероприятия по курсу. Понятие системы поддержки принятия решений, круг решаемых задач. Виды аналитических систем. Типичная архитектура системы поддержки принятия решений и ее особенности.	https://edu.vsu.ru/course/view.php?id=29860
1.2	Хранилища данных	Понятие, причины создание, принципиальные отличия от БД в OLTP системе. Виды реализации и перечень хранимых данных. Понятие метаданных.	https://edu.vsu.ru/course/view.php?id=29860
1.3	Преобразование и очистка данных	ETL процесс. Понятие качества данных, основные проблемы с качеством данных. Подходы к количественной оценке качества данных. Методы решения проблем качества данных.	https://edu.vsu.ru/course/view.php?id=29860
1.4	Поисково-аналитические технологии	Понятие запроса и объекта запроса. Классификация поисковых задач. Этапы поиска информации. Особенности прохождения каждого этапа при поиске по структурированным массивам данных.	https://edu.vsu.ru/course/view.php?id=29860
1.5	Оперативный анализ данных (OLAP-системы)	Понятие многомерного анализа данных. Гиперкуб и возможные операции над ним. Понятие OLAP. 12 правил Кодда, 6 дополнительных правил Кодда, 4 группы правил. FASMI. Подходы к реализации OLAP.	https://edu.vsu.ru/course/view.php?id=29860
1.6	Нормативно-справочная информация (НСИ) и ее использование в	Понятие НСИ. Роль НСИ в информационной системе и ее задачи. Типовые ошибки и заблуждения в отношении НСИ.	https://edu.vsu.ru/course/view.php?id=29860

	ИС		hp?id=29860
2. Практические занятия			
3. Лабораторные работы			
3.1	Введение в курс. Системы поддержки принятия решений.	Структура курса, цели и задачи. Контрольные мероприятия по курсу. Понятие системы поддержки принятия решений, круг решаемых задач. Виды аналитических систем. Типичная архитектура системы поддержки принятия решений и ее особенности.	https://edu.vsu.ru/course/view.php?id=29860
3.2	Хранилища данных	Понятие, причины создание, принципиальные отличия от БД в OLTP системе. Виды реализации и перечень хранимых данных. Понятие метаданных.	https://edu.vsu.ru/course/view.php?id=29860
3.3	Преобразование и очистка данных	ETL процесс. Понятие качества данных, основные проблемы с качеством данных. Подходы к количественной оценке качества данных. Методы решения проблем качества данных.	https://edu.vsu.ru/course/view.php?id=29860
3.4	Поисково-аналитические технологии	Понятие запроса и объекта запроса. Классификация поисковых задач. Этапы поиска информации. Особенности прохождения каждого этапа при поиске по структурированным массивам данных.	https://edu.vsu.ru/course/view.php?id=29860

3.5	Оперативный анализ данных (OLAP-системы)	Понятие многомерного анализа данных. Гиперкуб и возможные операции над ним. Понятие OLAP. 12 правил Кодда, 6 дополнительных правил Кодда, 4 группы правил. FASMI. Подходы к реализации OLAP.	https://edu.vsu.ru/course/view.php?id=29860
3.6	Нормативно-справочная информация (НСИ) и ее использование в ИС	Понятие НСИ. Роль НСИ в информационной системе и ее задачи. Типовые ошибки и заблуждения в отношении НСИ.	https://edu.vsu.ru/course/view.php?id=29860

13.2. Темы (разделы) дисциплины и виды занятий

№ п/п	Наименование темы (раздела) дисциплины	Виды занятий (часов)				
		Лекции	Практические	Лабораторные	Самостоятельная работа	Всего
1	Введение в курс. Системы поддержки принятия решений.	2	-	2	2	6
2	Хранилища данных	6	-	6	6	18
3	Преобразование и очистка данных	6	-	6	6	18
4	Поисково-аналитические технологии	8	-	8	8	24
5	Оперативный анализ данных (OLAP-системы)	8	-	8	8	24
6	Нормативно-справочная информация (НСИ) и ее использование в ИС	6	-	6	6	18
	Итого:	36	-	36	36	108

14. Методические указания для обучающихся по освоению дисциплины

В процессе преподавания дисциплины используются такие виды учебной работы, как лекции, лабораторные занятия, а также различные виды самостоятельной работы обучающихся. На лекциях рассказывается теоретический материал, на лабораторных занятиях решаются примеры по теоретическому материалу, прочитанному на лекциях.

При изучении курса «Специальные технологии баз данных и информационных систем» обучающимся следует внимательно слушать и конспектировать материал, излагаемый на аудиторных занятиях. Для его понимания и качественного усвоения рекомендуется следующая последовательность действий.

1. После каждой лекции студентам рекомендуется подробно разобрать прочитанный теоретический материал, разобрать примеры, решенные на лекции. Перед следующей лекцией обязательно повторить материал предыдущей лекции.

2. Перед лабораторным занятием обязательно повторить лекционный материал. После лабораторного занятия еще раз разобрать решенные на этом занятии примеры, после чего приступить к выполнению домашнего задания. Если при решении примеров, заданных на дом, возникнут вопросы, обязательно задать на следующем практическом занятии или в часы консультаций преподавателю.

3. При подготовке к лабораторным занятиям повторить основные понятия по темам, изучить примеры. Решая задачи, предварительно понять, какой теоретический материал нужно использовать. Наметить план решения, попробовать на его основе решить практические задачи.

4. Выбрать время для работы с литературой по дисциплине в библиотеке

Освоение дисциплины предполагает не только обязательное посещение обучающимся аудиторных занятий (лекций и лабораторных занятий) и активную работу на них, но и самостоятельную учебную деятельность, на которую отводится 36 часов.

Самостоятельная учебная деятельность студентов по дисциплине «Специальные технологии баз данных и информационных систем» предполагает изучение рекомендуемой преподавателем литературы по вопросам лекционных и лабораторных занятий (приведены выше), самостоятельное освоение понятийного аппарата и подготовку к текущим аттестациям (выполнению лабораторных заданий) (примеры см. ниже).

Вопросы лекционных и практических занятий обсуждаются на занятиях в виде устного опроса – индивидуального и фронтального. При подготовке к лекционным и практическим занятиям, обучающимся важно помнить, что их задача, отвечая на основные вопросы плана занятия и дополнительные вопросы преподавателя, показать свои знания и кругозор, умение логически построить ответ, владение математическим аппаратом и иные коммуникативные навыки, умение отстаивать свою профессиональную позицию. В ходе устного опроса выявляются детали, которые по каким-то причинам оказались недостаточно осмысленными студентами в ходе учебных занятий. Тем самым опрос выполняет важнейшие обучающую, развивающую и корректирующую функции, позволяет студентам учесть недоработки и избежать их при подготовке к промежуточным аттестациям.

Все выполняемые студентами самостоятельно задания (выполнение лабораторных заданий) подлежат последующей проверке преподавателем. Результаты текущих аттестаций учитываются преподавателем при проведении промежуточной аттестации (9 семестр – экзамен).

15. Перечень основной и дополнительной литературы, ресурсов интернет, необходимых для освоения дисциплины

а) основная литература:

№ п/п	Источник
1	Дейт К. Дж. Введение в системы баз данных / К. Дж. Дейт. – Москва : 8-е издание, Вильямс, 2017. – 1328 с.
2	Астахова, И. Ф.а. SQL в примерах и задачах : учебное пособие / Астахова И.Ф., Толстобров А.П., Мельников В.М. — 2-е изд., испр. и доп. — Воронеж : Изд-во Воронеж.гос.ун-та, 2000. — 161 с.
3	Стасышин, В. М. Базы данных: технологии доступа : учебное пособие для академического бакалавриата / В.М. Стасышин, Т.Л. Стасышина ; Новосибир. гос. техн. ун-т. — 2-е изд., испр. и доп. — Москва : Юрайт, 2018. — 176, [2] с. : ил., табл. — (Университеты России). — Библиогр.: с. 169-170.
4	Базы данных : учебник для вузов / А.Д. Хомоненко, В.М. Цыганков, М.Г. Мальцев ; под ред. А.Д. Хомоненко. — 4-е изд., доп. и перераб. — Санкт-Петербург : КОРОНА принт, 2004. — 736 с.

б) дополнительная литература:

№ п/п	Источник
5	Фуфаев, Э. В. Базы данных / Э.В. Фуфаев, Д.Э. Фуфаев. — 7-е изд., стер. — М. : Академия, 2012. — 319, [1] с.
6	Голицына О. Л. Системы управления базами данных / О.Л. Голицына, - М. :Форум, 2011. — 432 с.
7	Федоров, А. Г. Введение в OLAP-технологии Microsoft : практическое пособие : [16+] / А. Г. Федоров, Н. З. Елманова. – Москва : Диалог-МИФИ, 2002. – 268 с.
8	Кузнецов, С. Д. СУБД (системы управления базами данных) и файловые системы / Сергей Кузнецов. — М. : Майор, 2001. — 175 с. : ил.

в) информационные электронно-образовательные ресурсы:

№ п/п	Ресурс
9	Ипатова, Э. Р. Методологии и технологии системного проектирования информационных систем : учебник / Э. Р. Ипатова, Ю. В. Ипатов. — 3-е изд., стер. — Москва : ФЛИНТА, 2021. — 256 с. : табл., схем. — (Информационные технологии). — Режим доступа: по подписке. — URL: https://biblioclub.ru/index.php?page=book&id=79551
10	Лопушанский, В. А. Информационные системы. Системы управления базами данных : теория и практика : учебное пособие / В. А. Лопушанский, С. В. Макеев, Е. С. Бунин. - Воронеж : ВГУИТ, 2021. - 107 с. - ISBN 978-5-00032-519-3.

16. Перечень учебно-методического обеспечения для самостоятельной работы

№ п/п	Источник
1	<i>Груздев, Д.В. Базы данных: SQL, DELPHI, ORACLE [Электронный ресурс] : учебно-методическое пособие для вузов : [для студ.3 к. мат. фак.] / Д.В. Груздев ; Воронеж. гос. ун-т .— Электрон. текстовые и граф. дан. — Воронеж : ИПЦ ВГУ, 2010 .— Загл. с титул. экрана .— Свободный доступ из интрасети ВГУ.— http://www.lib.vsu.ru/elib/texts/method/vsu/m10-79.pdf.</i>
2	<i>Коннолли, Томас. Базы данных : Проектирование, реализация и сопровождение. Теория и практика / Томас Коннолли, Каролин Бегг, Анна Страчан ; Пер. с англ. Ю. Г. Гордиенко, А. В. Слепцова; Под ред. А. В. Слепцова .— 2-е изд.,испр. и доп. — М.; СПб.; Киев : Вильямс, 2000 .— 1111 с. : ил., табл. — Парал. тит. л. англ</i>

17. Образовательные технологии, используемые при реализации учебной дисциплины, включая дистанционные образовательные технологии (ДОТ), электронное обучение (ЭО), смешанное обучение):

При реализации дисциплины используются следующие образовательные технологии: логическое построение дисциплины, установление межпредметных связей, обозначение теоретического и практического компонентов в учебном материале, актуализация личного и учебно-профессионального опыта обучающихся, включение элементов дистанционных образовательных технологий.

В практической части курса используется стандартное современное программное обеспечение персонального компьютера.

В части освоения материала лекционных и лабораторных занятий, самостоятельной работы по отдельным разделам дисциплины, прохождения текущей и промежуточной аттестации может применяться электронное обучение и дистанционные образовательные технологии на портале «Электронный университет ВГУ».

18. Материально-техническое обеспечение дисциплины:

Для проведения лекционных и лабораторных занятий используются учебные аудитории. Для самостоятельной работы используется класс с компьютерной техникой, оснащенный необходимым программным обеспечением, электронными учебными пособиями и законодательно - правовой и нормативной поисковой системой, имеющий выход в глобальную сеть.

Компьютерный класс: специализированная мебель, маркерная доска, персональные компьютеры.

Ubuntu (бесплатное и/или свободное ПО, лицензия: <https://ubuntu.com/download/desktop>)

Visual Studio Community (бесплатное и/или свободное ПО, лицензия <https://visualstudio.microsoft.com/ru/vs/community/>)

LibreOffice (GNU Lesser General Public License (LGPL), бесплатное и/или свободное ПО, лицензия: <https://ru.libreoffice.org/about-us/license/>)

В самостоятельной работе обучающиеся используют ресурсы Зональной научной библиотеки ВГУ (электронный каталог: <http://www.lib.vsu.ru>)

19. Оценочные средства для проведения текущей и промежуточной аттестаций

Порядок оценки освоения обучающимися учебного материала определяется содержанием следующих разделов дисциплины:

№ п/п	Наименование раздела дисциплины (модуля)	Компетенция(и)	Индикатор(ы) достижения компетенции	Оценочные средства
1.	Введение в курс. Системы поддержки принятия решений.	ОПК–2	ОПК-2.2.2 ОПК-2.2.3 ОПК-2.4.1 ОПК-2.4.2	
2.	Хранилища данных	ОПК–2	ОПК-2.2.2 ОПК-2.2.3 ОПК-2.4.1 ОПК-2.4.2	Лабораторная работа
3.	Преобразование и очистка данных	ОПК–2	ОПК-2.2.2 ОПК-2.2.3 ОПК-2.4.1 ОПК-2.4.2	Лабораторная работа
4.	Поисково-аналитические технологии	ОПК–2	ОПК-2.2.2 ОПК-2.2.3 ОПК-2.4.1 ОПК-2.4.2	Лабораторная работа
5	Оперативный анализ данных (OLAP-системы)	ОПК–2	ОПК-2.2.2 ОПК-2.2.3 ОПК-2.4.1 ОПК-2.4.2	Лабораторная работа

№ п/п	Наименование раздела дисциплины (модуля)	Компетенция(и)	Индикатор(ы) достижения компетенции	Оценочные средства
6	Нормативно-справочная информация (НСИ) и ее использование в ИС	ОПК–2	ОПК-2.2.2 ОПК-2.2.3 ОПК-2.4.1 ОПК-2.4.2	Лабораторная работа
Промежуточная аттестация форма контроля – экзамен				Перечень вопросов к экзамену

20. Типовые оценочные средства и методические материалы, определяющие процедуры оценивания

20.1. Текущий контроль успеваемости

Контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств: лабораторная работа.

Текущая аттестация проводится в соответствии с Положением о текущей аттестации обучающихся по программам высшего образования Воронежского государственного университета.

Текущая аттестация заключается в защите лабораторных работ, содержание которых приведено ниже.

При текущем контроле уровень освоения учебной дисциплины и степень сформированности компетенции определяются оценками «зачтено», «не зачтено».

Перечень лабораторных работ

1. Развертывание OLAP-куба в MS SQL Server Analysis Services.
2. Изменения мер, атрибутов и иерархий OLAP-кубов в Microsoft SQL Server Analysis Services
3. Определение расширенных свойств атрибутов и измерений в OLAP-кубах.
4. Инструменты анализа данных в MS SQL server.
5. Анализ данных в BIDevStudio.

Критерии оценивания результатов обучения при текущей аттестации.

Для оценивания результатов контрольной работы используется **шкала**: «зачтено», «не зачтено».

Соотношение показателей, критериев и шкалы оценивания результатов обучения показаны в следующей таблице:

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
При выполнении лабораторной работы студент продемонстрировал в достаточной мере: знание основ составления компьютерных программ для решения типовых математических задач, имеющихся ресурсов для решения прикладных математических задач, умение использовать стандартные пакеты программного обеспечения для решения типовых математических задач, владение навыками хранения, поиска, сбора, систематизации, обработки и использования информации.	Достаточный уровень	Зачтено
При выполнении лабораторной работы студент не продемонстрировал в достаточной мере: знание основ составления компьютерных программ для решения типовых математических задач, имеющихся ресурсов для решения прикладных математических задач, умение использовать стандартные пакеты программного обеспечения для решения типовых математических задач, владение навыками хранения, поиска, сбора, систематизации, обработки и использования информации.	–	Не зачтено

20.2 Промежуточная аттестация

Промежуточная аттестация по дисциплине осуществляется с помощью следующих оценочных средств:

- Собеседование по экзаменационным билетам

Промежуточная аттестация предназначена для определения уровня освоения всего объема учебной дисциплины. Промежуточная аттестация по дисциплине «Специальные технологии баз данных и информационных систем» проводится в форме экзамена (9 семестр).

Промежуточная аттестация, как правило, осуществляется в конце семестра. Результаты текущей аттестации обучающегося по решению кафедры могут быть учтены при проведении промежуточной аттестации. При несогласии студента, ему дается

возможность пройти промежуточную аттестацию (без учета его текущих аттестаций) на общих основаниях.

В ходе промежуточной аттестации обучающемуся выдается КИМ. В ходе выполнения заданий нельзя пользоваться ресурсами глобальной сети Интернет.

Промежуточная аттестация в 9 семестре по дисциплине осуществляется в форме собеседования по экзаменационным билетам с помощью ниже приведенных оценочных средств (перечень вопросов к экзамену).

Перечень вопросов к экзамену

1. Что такое хранилище данных? Зачем нужны хранилища данных? Какие задачи решает? Области применения хранилищ данных.
2. Виды моделей хранилища данных
3. Общие свойства и структура хранилищ данных
4. Общие свойства хранилищ данных. Данные хранилища.
5. Компоненты хранилища
6. Расскажите о OLTP и OLAP подходах. Какие между ними отличия?
7. Что такое ETL-процесс?
8. Что называется гиперкубом? Какие операции над гиперкубом вы знаете?
9. Какую ER-диаграмму называют диаграмма-снежинка? Что такое star-схема?
10. Расскажите о киосках данных (data mart). Зачем они нужны?
11. Какой жизненный цикл имеет разработка хранилища данных? Какие отличия от разработки обычного ИТ проекта?
12. Какие задачи решает интеллектуальный анализ данных. Приведите примеры систем, позволяющих его проводить.
13. Аппроксимация данных. Определение. Приведите примеры систем, позволяющих ее проводить.
14. Классификация данных. Определение. Приведите примеры систем, позволяющих ее проводить.
15. Кластеризация данных. Определение. Приведите примеры систем, позволяющих ее проводить.
16. Классификация и кластеризация данных, назовите главные отличия данных подходов к обработке данных и различия в их назначении.
17. Большие данные. Назовите главные отличия больших данных от данных обрабатываемых в классических OLAP системах.
18. Поточковые данные. Назовите главные отличия больших данных от данных обрабатываемых в классических OLAP системах.
19. Назовите основные сложности при обработке данных в реальном времени.
20. Самообучающиеся системы анализа и самообучаемый, приведите основные отличия.
21. Многопоточная обработка данных в облачных системах, построенных на принципах Hadoop. Опишите схему работы некоторых представителей таких систем. Назовите преимущества и недостатки этого принципа.
22. Сравните схему хранения данных в виде структуры «Звезда» и «Снежинка». Назовите основные преимущества и недостатки каждой из структур.
23. Что такое сущность, атрибут, связи?
24. Какие типы ключей вы знаете?
25. Перечислите основные этапы проектирования реляционных баз данных.
26. Концептуальная модель.
27. Физическая модель.
28. Многомерное моделирование и звездообразные схемы
29. Какие типовые архитектуры хранилищ данных вам известны?
30. Чем обусловлен выбор модели хранилища данных?

31. В чем особенность звездообразных схем структуры хранилищ данных?
32. Выбор метода реализации хранилищ данных
33. Проблема интеграции данных.
34. Возможности SQL Server 2008 Integration Services.
35. Планирование ETL проекта для хранилища данных
36. Технология (процесс) добычи знаний. Решаемые задачи.
37. Математические основы (РАД).
38. Data Mining в MSSAS
39. Что такое сводные таблицы? Для чего они применяются?
40. Из каких элементов состоят сводные таблицы?
41. На основе каких источников возможно создание сводных таблиц?
42. Как создать сводную таблицу в MS Excel при наличии готовой таблицы? И как можно изменить структуру полученной сводной таблицы (имеется ввиду положение полей)?
43. Можно ли, и если можно, то как, создать диаграмму для сводной таблицы?

Для оценивания результатов обучения на экзамене используется 4-балльная шкала: «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Критерии оценивания компетенций	Уровень сформированности компетенций	Шкала оценок
Полное соответствие обучающимся всем перечисленным показателям по каждому из вопросов контрольно-измерительного материала. Обучающийся в полной мере владеет понятийным аппаратом данной области науки (теоретическими основами дисциплины), способен иллюстрировать ответ примерами, применять теоретические знания для решения практических задач в области курса, студент умеет работать с различными источниками научной информации, грамотно и правильно представляет свои результаты, правильно отвечает на вопросы КИМ	Повышенный уровень	Отлично
Несоответствие ответа обучающегося одному из перечисленных выше показателей (к одному из вопросов контрольно-измерительного материала) и правильный ответ на дополнительный вопрос в пределах программы. ИЛИ Несоответствие ответа обучающегося любым двум из перечисленных показателей (либо двум к одному	Достаточный уровень	Хорошо

вопросу, либо по одному к каждому вопросу контрольно-измерительного материала) и правильные ответы на два дополнительных вопроса в пределах программы.		
Несоответствие ответа обучающегося любым двум из перечисленных показателей и неправильный ответ на дополнительный вопрос в пределах программы. ИЛИ Несоответствие ответа обучающегося любым трем из перечисленных показателей (в различных комбинациях по отношению к вопросам контрольно-измерительного материала).	Пороговый уровень	Удовлетворительно
Несоответствие ответа обучающегося любым из перечисленных показателей (в различных комбинациях по отношению к вопросам контрольно-измерительного материала).	—	Неудовлетворительно

20.3 Фонд оценочных средств сформированности компетенций студентов, рекомендуемый для проведения диагностических работ

Задания открытого типа (тестовые, повышенный уровень сложности):

1. Для добавления нумерации страниц в отчет необходимо добавить данные в блок ... (вставьте два слова)

Ответ: Page Footer

2. Для добавления всех данных из определенного поля базы данных в отчет необходимо переместить поле в блок (вставьте слово)

Ответ: Details

3. Для формирования наименования отчета в SAP Crystal Reports необходимо вставить данные в блок ... (вставьте два слова)

Ответ: Report Header

4. Под интеллектуальным анализом данных подразумевается технология ... (вставьте два слова)

Ответ: Data Mining

5. Под оперативным анализом данных подразумевается технология ... (вставьте слово)

Ответ: OLAP

6. Инструмент MS SQL Server, который используется для создания новых OLAP-кубов (вставьте три слова)

Ответ: SS Data Tools

7. Инструмент MS SQL Server «SS Data Tools» используется для создания новых ...

Ответ: OLAP-кубов

8. Компонент MS SQL Server, который используется для операций ETL ...

(вставьте два слова)

Ответ: Integration Services

9. Компонент MS SQL Server «Integration Services» используется для операций ...

(вставьте слово)

Ответ: ETL

10. Поддержка модели хранилища данных в актуальном состоянии в CA ERwin Data Modeler осуществляется с помощью инструмента ... (вставьте два слова)

Ответ: Complete Compare

11. Для таблиц, связанных неидентифицирующей связью в модели ERwin Data Modeler, справедливо утверждение: «.... *родительской таблицы автоматически мигрирует в дочернюю таблицу в область атрибутов*»

(вставьте два слова)

Ответ: Первичный ключ

12. Таблица фактов содержит , являющиеся предметом анализа (вставьте слово)

Ответ: данные

13. Таблица измерений содержит ... , по которым можно анализировать данные (вставьте слово)

Ответ: категории

14. Вставьте два слова

Альтернативный подход к созданию аналитических решений по сравнению с традиционными промышленными системами бизнес-анализа. Решения типа предлагают пользователям интерактивный графический пользовательский интерфейс, базирующийся на архитектуре in-memory, что соответствует запросу бизнеса на простые и быстрые в работе BI-системы.

Ответ: data discovery

15. Вставьте слово

Процесс управления данными, который включает в себя:

- извлечение данных из внешних источников;
- трансформацию и очистка извлеченных данных, чтобы они соответствовали потребностям бизнес-модели;
- загрузку преобразованных данных в хранилище данных.

Как называется данный процесс?

Ответ: ETL

16. Вставьте слово

Технологии интерактивной аналитической обработки данных для поддержки принятия решений, ориентированные главным образом на нерегламентированные интерактивные запросы, как правило, имеет дело с историческими данными и часто используется в хранилищах данных, называется ...

Ответ: OLAP

17. Вставьте слово

Технологии OLAP, использующие представление куба данных, которое поддерживается над реляционной базой данных, называются ...

Ответ: ROLAP

18. Вставьте слово

..... – скалярное значение, вычисляемое по некоторому множеству детализированных записей

Ответ: агрегат

19. Вставьте слово

.... – степень детализации фактов в кубе данных. Определяется количеством измерений, связанных с таблицей фактов, и иерархиями измерений.

Ответ: гранулярность

20. Вставьте слово

Процесс, противоположный нормализации. Процесс преобразования отношений базы данных к виду, допускающему избыточность данных, ориентированных в первую очередь на высокую производительность при выполнении аналитических запросов. Основное назначение данного процесса состоит в ограничении межтабличных соединений в запросах. Назовите этот процесс

Ответ: денормализация

21. Вставьте слово

... – это способ ранжирования, используемый для разделения агрегированных фактов на составляющие их части

Ответ: измерение

22. Вставьте два слова

..... – это основная структура данных в многомерных моделях данных. Имеет несколько независимых измерений – систему координат представляемого пространства данных. Комбинации значений координат по всем измерениям определяют точки куба,

называемые ячейками. С ячейками ассоциируются значения переменных, называемых показателями и имеющих, как правило, числовые типы.

Ответ: куб данных

23. Вставьте слово

... – упрощенное представление реальности, разработанное с целью донести информацию до определенной аудитории для обеспечения анализа, взаимодействия и понимания

Ответ: модель

24. Вставьте слово

... – процесс преобразования отношений базы данных к виду, отвечающему нормальным формам

Ответ: нормализация

25. Вставьте слово

.... куба данных – это операция над кубом данных, которая строит куб меньшей размерности при фиксированных значениях для каких-либо измерений. Этот термин обозначает также результат указанной операции.

Ответ: Проекция

26. Вставьте слово

Операция перемещения вниз по уровням иерархии измерения гиперкуба называется ...

Ответ: развертка/развёртка

27. Вставьте слово

Операция перемещения вверх по уровням иерархии измерения гиперкуба называется ...

Ответ: свертка/свёртка

28. Вставьте слово

Схема «...» – это схема реляционной базы данных, служащая для поддержки многомерного представления содержащихся в ней данных. Описываемая база данных включает таблицу фактов и ряд связанных с ней при помощи внешних ключей таблиц измерений. Таблицы измерений могут иметь составные первичные ключи и являются денормализованными. Благодаря этому упрощается восприятие структуры данных пользователем и формулировка запросов, уменьшается количество операций соединения таблиц при обработке запросов. Однако в связи с избыточностью данных возрастает требуемый для их хранения объем памяти.

Ответ: Звезда

29. Вставьте слово

Схема «....» – разновидность схемы типа звезды, предусматривающая нормализацию таблиц измерений. Первичные ключи в них состоят из единственного атрибута (соответствуют единственному элементу измерения). Это позволяет минимизировать избыточность данных и более эффективно выполнять запросы, связанные со структурой значений измерений.

Ответ: Снежинка

30. Вставьте слово

... – это структурный элемент куба данных, соответствующий набору значений по всем измерениям.

Ответ: ячейка

Задания закрытого типа (тестовые, средний уровень сложности):

1. Таблица фактов содержит

- a) множество метаданных модели
- b) данные, являющиеся предметом анализа
- c) схему «звезда»
- d) список трансформаций данных, необходимых для процесса ETL
- e) категории, по которым производится анализ данных

Ответ: b

2. Таблица измерений содержит

- a) множество метаданных модели
- b) данные, являющиеся предметом анализа
- c) схему «снежинка»
- d) категории, по которым можно анализировать данные
- e) список трансформаций данных, необходимых для процесса ETL

Ответ: d

3. Консольная таблица характеризуется следующими утверждениями

- a) Является обязательной частью схемы «звезда»
- b) Содержит данные, являющиеся предметом анализа
- c) Является частью схемы «снежинка»
- d) Повышает производительность аналитических запросов
- e) Не может быть связана непосредственно с таблицей фактов
- f) Увеличивает степень нормализации модели данных

Ответ: c, e, f

4. Требования к транзакционным системам включают

- a) Быструю вставку и обновление
- b) Быстрое объединение и выборки (joins and selects)
- c) Прямой доступ к текущей информации
- d) Максимум ссылочной целостности
- e) Доступ к агрегированной исторической информации

Ответ: a, c, d

5. Требования к аналитическим системам включают

- a) Быструю вставку и обновление
- b) Быстрое объединение и выборки (joins and selects)
- c) Прямой доступ к текущей информации
- d) Максимум ссылочной целостности
- e) Доступ к агрегированной исторической информации

Ответ: b, e

6. Схема «звезда» характеризуется следующими утверждениями:

- a) Включает одну таблицу фактов
- b) Может включать несколько таблиц фактов
- c) Включает только одну таблицу измерений
- d) Может включать несколько таблиц измерений

- e) Все сведения об иерархиях измерения хранятся в одной таблице измерения
 - f) Сведения об иерархиях измерения хранятся в нескольких связанных таблицах
- Ответ: a, d, e

7. Схема «снежинка» характеризуется следующими утверждениями:

- a) Включает одну таблицу фактов
 - b) Может включать несколько таблиц фактов
 - c) Включает только одну таблицу измерений
 - d) Может включать несколько таблиц измерений
 - e) Все сведения об иерархиях измерения хранятся в одной таблице измерения
 - f) Сведения об иерархиях измерения хранятся в нескольких связанных таблицах
- Ответ: a, d, f

8. Для таблиц, связанных идентифицирующей связью в модели ERwin Data Modeler, справедливы утверждения:

- a) Первичный ключ родительской таблицы не мигрирует по связи в дочернюю таблицу;
- b) Первичный ключ родительской таблицы автоматически мигрирует в дочернюю таблицу в область ключевых колонок;
- c) Первичный ключ родительской таблицы автоматически мигрирует в дочернюю таблицу в область атрибутов;
- d) Идентификация объектов дочерней таблицы невозможна без объекта родительской таблицы.

Ответ: b, d

9. Для таблиц, связанных неидентифицирующей связью в модели ERwin Data Modeler, справедливы утверждения:

- a) Первичный ключ родительской таблицы не мигрирует по связи в дочернюю таблицу;
- b) Первичный ключ родительской таблицы автоматически мигрирует в дочернюю таблицу в область ключевых колонок;
- c) Первичный ключ родительской таблицы автоматически мигрирует в дочернюю таблицу в область атрибутов;
- d) Идентификация объектов дочерней таблицы невозможна без объекта родительской таблицы.

Ответ: c

10. Для связи между таблицей-измерением и таблицей-фактов справедливы утверждения:

- a) Таблицы должны быть связаны идентифицирующей связью;
- b) Таблицы могут быть связаны связью многие-ко-многим;
- c) Таблицы могут быть связаны неидентифицирующей связью;
- d) Таблица-измерение должна находиться на родительском конце связи;

Ответ: a, d

11. Для связи между таблицей-измерением и консольной таблицей справедливы утверждения:

- a) Таблицы должны быть связаны идентифицирующей связью;
- b) Таблицы могут быть связаны связью многие-ко-многим;
- c) Таблицы могут быть связаны неидентифицирующей связью;
- d) Таблица-измерение должна находиться на родительском конце связи;

Ответ: c

12. Для прямой генерации каталога витрины данных в CA ERwin Data Modeler требуется:

- a) Переключиться на логический уровень модели данных;
- b) Переключиться на физический уровень модели данных;
- c) Выполнить подключение к базе данных;
- d) Заменить все неидентифицирующие связи идентифицирующими.

Ответ: b, c

13. Поддержка модели хранилища данных в актуальном состоянии в CA ERwin Data Modeler осуществляется с помощью инструмента:

- a) Forward Engineering;
- b) Reverse Engineering;
- c) Complete Compare;
- d) Report Designer.

Ответ: c

14. Компонент MS SQL Server, который используется для операций ETL:

- a) Database Engine;
- b) Analysis Services;
- c) Integration Services;
- d) Reporting Service;

Ответ: c

15. Инструмент MS SQL Server, который используется для создания новых OLAP-кубов:

- a) SS Data Tools;
- b) SS Management Studio;
- c) Report Designer;
- d) Database Engine.

Ответ: a

16. Использование надстроек Excel для анализа таблиц требует:

- a) Использование данных из хранилища данных;
- b) Использование данных на страницах Excel, форматированных произвольным способом;
- c) Использование данных на страницах Excel, форматированных как таблица;
- d) Подключения к аналитическим службам MS SQL Server;
- e) Автономной работы Excel.

Ответ: c, d

17. Под оперативным анализом данных подразумеваются следующие технологии:

- a) OLAP;
- b) Data Mining;
- c) Имитационное моделирование.

Ответ: a

18. Под интеллектуальным анализом данных подразумеваются следующие технологии:

- a) OLAP;
- b) Data Mining;
- c) Имитационное моделирование.

Ответ: b

19. Особенностью систем Data Discovery являются:

- a) Использование технологий in-memory;
- b) Простота формирования отчетов;
- c) Расширенные возможности использования методов глубокой аналитики;
- d) Возможности формирования регламентных отчетов.

Ответ: a, b

20. Для того, чтобы выбрать определенные данные в Power BI необходимо:

- a) Написать SQL-запрос для выделения нужных данных;
- b) Интерактивно перенести поля для фильтрации в область «Фильтры»;
- c) В комбобоксе области «Фильтры» выбрать поля и прописать условия для отбора данных;
- d) Настроить фильтры в системе невозможно.

Ответ: b, c

21. Для того, чтобы выбрать определенные данные в SAP Crystal Reports необходимо:

- a) Написать SQL-запрос для выделения нужных данных;
- b) Настроить фильтры при формировании отчета в окне «Фильтры»;
- c) Интерактивно перенести поля для фильтрации в область «Фильтры»;
- d) Настроить фильтры в области «Проекты».
- e) Настроить фильтры в системе невозможно.

Ответ: b, d

22. Для публикации отчетов, сформированных в системе Power BI необходимо:

- a) Зарегистрироваться на powerbi.com и загрузить отчет через портал;
- b) Выбрать пункт меню: файл-сохранить как-сохранить как публикацию.
- c) Выбрать пункт меню: файл-опубликовать в интернете-создать код публикации.

Ответ: c

23. В системе Power BI существуют следующие возможности управления данными:

- a) Можно формировать хранилища данных по схеме «звезда».
- b) Можно переименовывать поля и менять их формат.
- c) Можно изменять структуру источника данных.
- d) Можно менять наименование и формат поля, разделять одно поле на несколько.
- e) Нет возможности настраивать источники данных.

Ответ: e

24. В системе Tableau существуют следующие возможности управления данными:

- a) Можно формировать хранилища данных по схеме «звезда».
- b) Можно переименовывать поля и менять их формат.
- c) Можно изменять структуру источника данных.
- d) Можно менять наименование и формат поля, разделять одно поле на несколько.
- e) Нет возможности настраивать источники данных.

Ответ: d

25. В системе Tableau существуют следующие возможности управления данными:

- a) Можно загружать несколько источников, но использовать для формирования отчетов только один.
- b) Можно загружать и одновременно использовать несколько источников.

с) Можно загружать и использовать только один источник данных.

Ответ: b

26. В системе Power BI существуют следующие возможности управления данными:

а) Можно загружать несколько источников, но использовать для формирования отчетов только один.

б) Можно загружать и одновременно использовать несколько источников.

с) Можно загружать и использовать только один источник данных.

Ответ: b

27. В системе SAP Crystal Reports существуют следующие возможности управления данными:

а) Можно загружать несколько источников, но использовать для формирования отчетов только один.

б) Можно загружать и одновременно использовать несколько источников.

с) Можно загружать и использовать только один источник данных.

Ответ: a

28. Для формирования наименования отчета в SAP Crystal Reports необходимо вставить данные в блок:

а) Report Footer

б) Page Footer

с) Page Header

д) Report Header

е) Details

Ответ: d

29. Для добавления всех данных из определенного поля базы данных в отчет необходимо переместить поле в блок:

а) Report Footer

б) Page Footer

с) Page Header

д) Report Header

е) Details

Ответ: e

30. Для добавления нумерации страниц в отчет необходимо добавить данные в блок:

а) Report Footer

б) Page Footer

с) Page Header

д) Report Header

е) Details

Ответ: b

Критерии и шкалы оценивания заданий ФОС:

1) Задания закрытого типа (выбор одного варианта ответа, верно/неверно):

- 1 балл – указан верный ответ;
- 0 баллов – указан неверный ответ.

2) Задания закрытого типа (множественный выбор):

- 2 балла – указаны все верные ответы;
- 0 баллов — указан хотя бы один неверный ответ.

3) Задания закрытого типа (на соответствие):

- 2 балла – все соответствия определены верно;
- 0 баллов – хотя бы одно сопоставление определено неверно.

4) Задания открытого типа (короткий текст):

- 2 балла – указан верный ответ;
- 0 баллов – указан неверный ответ.

5) Задания открытого типа (число):

- 2 балла – указан верный ответ;
- 0 баллов – указан неверный ответ.

Задания раздела 20.3 рекомендуются к использованию при проведении диагностических работ с целью оценки остаточных результатов освоения данной дисциплины (знаний, умений, навыков).