



**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

федеральное государственное бюджетное образовательное учреждение
высшего образования
«ИРКУТСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
(ФГБОУ ВО «ИГУ»)

Институт математики и информационных технологий
Кафедра алгебраических и информационных систем

«УТВЕРЖДАЮ»
Директор ИМИТ ИГУ
М. В. Фалалеев
М. В. Фалалеев
«25» мая 2022 г.



Рабочая программа дисциплины (модуля)

Б1.В.06 Анализ и управление большими данными

Направление подготовки информационные технологии	02.04.02	Фундаментальная	информатика	и
Направленность (профиль) подготовки машинное обучение		Анализ данных научных исследований и		
Квалификация выпускника	магистр			
Форма обучения	очная			

Иркутск 2022 г.

Согласовано с УМК Института математики
и информационных технологий
Протокол № 3 от «04» апреля 2022 г.

Председатель _____
Антоник В.Г.

Рекомендовано кафедрой Алгебраических и
информационных систем ИМИТ ИГУ:
Протокол № 9 От «24» марта 2022 г.

Зав. кафедрой _____
Пантелеев В.И.

1. Цели и задачи дисциплины

Цель

Целями освоения дисциплины «Анализ и управление большими данными» являются ознакомление студентов с организацией таким систем хранения больших данных, как озеро данных и хранилище данных, а также процессами сбора и обработки данных.

Задачи:

Задачами освоения дисциплины являются следующие: сформировать компетенции, необходимые для организации систем хранения данных; выработать навыки проектирования хранилищ данных и организации ETL- и ELT-процессов; познакомить с такими подходами к проектированию хранилищ данных, как многомерное моделирование, якорное моделирование и Data Vault.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

2.1. Учебная дисциплина (модуль) относится к части программы, формируемой участниками образовательных отношений, и изучается на втором курсе.

2.2. Для изучения данной учебной дисциплины (модуля) необходимы знания и навыки, сформированные такими дисциплинами как «Современные технологии баз данных», «Облачные технологии».

2.3. Перечень последующих учебных дисциплин, для которых необходимы знания, умения и навыки, формируемые данной учебной дисциплиной: данная дисциплина дает знания, умения и навыки для выполнения научно-исследовательских работ, для решения прикладных задач и выполнения выпускной квалификационной работы.

3. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Процесс освоения дисциплины направлен на формирование компетенций (элементов следующих компетенций) в соответствии с ФГОС ВО по соответствующему направлению подготовки.

Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с индикаторами достижения компетенций

Компетенция	Индикаторы компетенций	Результаты обучения
ПК-1 Способен осуществлять управление, обработку, визуализацию и анализ данных (включая работу с большими данными), в том числе методами машинного обучения	ИДК _{ПК1.1} Способен пользоваться методами и инструментами получения, хранения, передачи и обработки данных (в том числе больших)	Знать: OLAP и OLTP-системы; слои в двух- и трехслойной структуре хранилища данных; способы хранения исторических данных с помощью различных SCD. Уметь: выбирать подходящий SCD для хранения истории изменения данных (в том числе больших), проектировать соответствующую систему хранения данных. Владеть: навыками организации ETL- и ELT-процессов.
	ИДК _{ПК1.2} Способен разрабатывать системы хранения и	Знать: основные системы хранения больших данных, хранилище данных, озеро

	обработки данных (в том числе больших)	<p>данных; методы проектирования хранилищ данных: многомерное моделирование, Data Vault, якорное моделирование.</p> <p>Уметь: проектировать хранилище данных по Кимбалу и по Инмону; выбирать подходящий под задачу метод организации системы хранения данных.</p> <p>Владеть: навыками проектирования хранилищ данных с использованием многомерного моделирования, Data Vault, якорного моделирования.</p>
--	--	---

4. СОДЕРЖАНИЕ И СТРУКТУРА ДИСЦИПЛИНЫ

Объем дисциплины составляет 5 зачетных единиц, 180 часов, в том числе 44 часа на контроль, практическая подготовка 34 часа.
Форма промежуточной аттестации: 3 семестр — экзамен.

4.1. СОДЕРЖАНИЕ ДИСЦИПЛИНЫ, СТРУКТУРИРОВАННОЕ ПО ТЕМАМ, С УКАЗАНИЕМ ВИДОВ УЧЕБНЫХ ЗАНЯТИЙ И ОТВЕДЕННОГО НА НИХ КОЛИЧЕСТВА АКАДЕМИЧЕСКИХ ЧАСОВ

№ п/п	Раздел дисциплины/темы	Се мес тр	Виды учебной работы, включая самостоятельную работу обучающихся и трудоемкость (в часах)				Формы текущего контроля успеваемости
			Контактная работа преподавателя с обучающимися			Самостоятельная работа + контроль	
			Лекции	Семинарские (практические занятия)	Контроль обучения		
1	Введение в большие данные		2	2			
2	Системы хранения данных		2	4			лаб.
3	Хранилище данных		2	6	2	14	лаб., тест
4	Многомерное моделирование		2	4	2	14	лаб.
5	Data Vault		2	4	2	14	лаб.
6	Якорное моделирование		2	4	2	14	лаб.
7	Сравнение подходов к хранению и обработке больших данных		2	2	2		лаб.
8	Распределенный анализ данных и машинное обучение		2	4		20	лаб.
	Подготовка к экзамену					44	
Итого часов			16	34	10	120	

4.2. План внеаудиторной самостоятельной работы обучающихся по дисциплине

Семестр	Название раздела, темы	Самостоятельная работа обучающихся			Оценочное средство	Учебно-методическое обеспечение самостоятельной работы
		Вид самостоятельной работы	Сроки выполнения	Затраты времени (час.)		
1	Хранилище данных	Выполнение практической работы	01.10	14	Проверка домашней работы	Материалы курса на платформе ИОС DOMIC
1	Многомерное моделирование	Выполнение практической работы	01.11	14	Проверка домашней работы	Материалы курса на платформе ИОС DOMIC
1	Data Vault	Выполнение практической работы	01.11	14	Проверка домашней работы	Материалы курса на платформе ИОС DOMIC
1	Якорное моделирование	Выполнение практической работы	01.12	14	Проверка домашней работы	Материалы курса на платформе ИОС DOMIC
1	Распределенный анализ данных и машинное обучение	Выполнение практической работы	01.12	20	Проверка домашней работы	Материалы курса на платформе ИОС DOMIC
1	Подготовка к экзамену	Изучение теоретического материала	01.01	44		Материалы курса на платформе ИОС DOMIC
Общая трудоемкость самостоятельной работы по дисциплине (час)				120		
Из них объем самостоятельной работы с использованием электронного обучения и дистанционных образовательных технологий (час)				120		

4.3. СОДЕРЖАНИЕ УЧЕБНОГО МАТЕРИАЛА

1. Введение в большие данные.
Предпосылки возникновения больших данных. Свойства больших данных. Применение больших данных.
2. Системы хранения данных.
OLTP- и OLAP-системы. Хранилище данных. Озеро данных. ETL- и ELT-процессы. Поточковая и пакетная обработка. Lambda- и карпа-архитектуры.
3. Хранилище данных (DWH).

- Уровневая архитектура, слои данных. DWH по Кибмаллу и Инмону.
4. Многомерное моделирование.
Реляционные базы данных. Нормальные формы. Витрины данных. Таблицы фактов и измерений. Модели «Звезда», «Снежинка» и «Созвездие». Виды фактов. Хранение истории изменения данных, SCD.
 5. Data Vault
Типы таблиц Data Vault. Хабы. Линки. Сателлиты. Data Vault 2.0. Мосты, PIT, предрассчитанные значения. Проектирование Data Vault.
 6. Якорное моделирование
Типы таблиц для якорной модели. Якорь. Атрибут. Связь. Узел. Проектирование якорной модели.
 7. Сравнение подходов к хранению и обработке больших данных
Трудоемкость развития хранилища данных. Сравнение моделей по сложности эксплуатации и сложности внесения изменений.
 8. Распределенный анализ данных и машинное обучение
Жизненный цикл моделей. Распределенные модели машинного обучения. Распределенные данные. Spark ML.

4.3.1. Перечень семинарских, практических занятий и лабораторных работ

№ п/н	№ раздела и темы	Наименование семинаров, практических и лабораторных работ	Трудоемкость (час.)		Оценочные средства	Формируемые компетенции (индикаторы)*
			Всего часов	Из них практическая подготовка		
1	2	3	4	5	6	7
1	2	Системы хранения данных	4	4	Проверка лабораторной работы	ПК-1 (ИДК _{ПК1.1} ИДК _{ПК1.2})
2	3	Хранилище данных	6	6	Проверка лабораторной работы	ПК-1 (ИДК _{ПК1.2})
3	4	Многомерное моделирование	4	4	Проверка лабораторной работы	ПК-1(ИДК _{ПК1.2})

4	5	Data Vault	4	4	Проверка лабораторной работы	ПК-1(ИДК _{ПК1.2})
5	6	Якорное моделирование	4	4	Проверка лабораторной работы	ПК-1(ИДК _{ПК1.2})
6	7	Сравнение подходов к хранению и обработке больших данных	2	2	Проверка лабораторной работы	ПК-1(ИДК _{ПК1.2})
7	8	Распределенный анализ данных и машинное обучение	4	4	Проверка лабораторной работы	ПК-1(ИДК _{ПК1.1})
		Всего	34	34		

4.3.2. Перечень тем (вопросов), выносимых на самостоятельное изучение студентами в рамках самостоятельной работы (СР)

Тема	Задание	Формируемые компетенции
Хранилище данных	Решить задачи для самостоятельного выполнения	ПК-1
Многомерное моделирование	Решить задачи для самостоятельного выполнения	ПК-1
Data Vault	Решить задачи для самостоятельного выполнения	ПК-1
Якорное моделирование	Решить задачи для самостоятельного выполнения	ПК-1
Распределенный анализ данных и машинное обучение	Решить задачи для самостоятельного выполнения	ПК-1

4.4. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОРГАНИЗАЦИИ САМОСТОЯТЕЛЬНОЙ РАБОТЫ СТУДЕНТОВ

Во время изучения дисциплины студент посещает лекции, практические занятия, выполняет лабораторные задания, готовится к тестам, зачетам и экзаменам. Для каждого вида деятельности необходимо правильно организовать самостоятельную работу.

Лекции. В высшем учебном заведении лекция является важной формой учебного процесса. На лекции студенты получают глубокие и разносторонние знания. Лекция способствует развитию творческих способностей, формирует идейную убежденность, позволяет устанавливать связь учебного материала с производством, новейшими научными достижениями. Лекция требует три вида деятельности: подготовку к лекции, работу на лекции и работу после лекции.

После прослушивания лекции студент должен проработать и осмыслить полученный материал. На каждый пример, приведенный на лекции, желательно, (если это возможно) привести свой. Материал, изложенный в лекции, можно просмотреть в других источниках.

В процессе лекционного занятия студент должен выделять важные моменты, выводы, анализировать основные положения. Недостаточно только «слушать» лекцию. Возможности памяти человека не универсальны. Как бы внимательно студент ни слушал лекцию, большая часть информации вскоре после восприятия будет забыта. Чтобы восстановить лекционный материал, его нужно повторить, а для этого лекцию необходимо конспектировать. Конспект лекций должен быть в отдельной тетради, в которой не должно быть ничего, кроме лекции. Не надо стремиться подробно слово в слово записывать всю лекцию. Конспектируйте только самое важное в рассматриваемой теме: ключевые слова и их значения, примеры использования конструкций, что старается выделить лектор, на чем акцентирует внимание студентов.

Тетрадь для конспекта лекций также требует особого внимания. Ее нужно сделать удобной, практичной и полезной, ведь именно она является основным информативным источником при подготовке к различным отчетным занятиям, зачетам, экзаменам. Конечно, оформление лекционной тетради – это дело вкуса. Но целесообразно отделить поля, где студент мог бы изложить свои мысли, вопросы, появившиеся в ходе лекции. Полезно одну из страниц оставлять свободной. Она потребует потом, при самостоятельной подготовке. Сюда можно будет занести дополнительную информацию по данной теме, полученную из других источников: рисунки, схемы, примеры кода и т.д.

Лабораторное занятие. Лабораторные занятия по решению задач существенно дополняют лекции. В процессе анализа и решения задач студенты расширяют и углубляют знания, полученные из лекционного курса и учебников, приобретают умение применять общие закономерности к конкретным случаям.

Необходимо, чтобы студенты готовили теоретический материал, т.к. именно невыполнение этого требования приводит к неудаче при решении задач.

Несмотря на различие в видах задач, их решение можно проводить по следующему общему плану (некоторые пункты плана могут выпадать в некоторых конкретных случаях): а) прочесть внимательно условие задачи; б) посмотреть, все ли термины в условиях задачи известны и понятны (если что-то неясно, следует обратиться к учебнику, просмотреть решения предыдущих задач, посоветоваться с преподавателем); в) произвести анализ задачи, (нужно чётко понимать, в чем будет заключаться решение задачи); г) решить задачу; д) протестировать полученное решение на данных из примеров к задаче, а также на дополнительных данных.

Если задача не решена или «не решается», то необходимо еще раз вернуться к пунктам а) и б). Сколько раз нужно возвращаться к этим пунктам? Практика показывает, что не более десяти раз. Если и после этого задача «не решается», то можно попытаться найти решение этой или похожей задачи в различных источниках.

Домашнее задание. При выполнении домашнего задания необходимо просмотреть текст лекции, разобраться с новыми определениями, посмотреть задания, которые были выполнены на лабораторной работе и применить полученные знания для выполнения домашней работы.

Тест. В первую очередь постарайтесь узнать чего ждать от теста, какие примерно там будут задания. Если вам доступны образцы теста (как, например, при сдаче ЕГЭ), необходимо этим воспользоваться и ежедневно тренироваться.

Не оставляйте все на самый последний момент. Если будете постоянно готовиться к тесту, вы наверняка улучшите свои знания. Для этого составьте план на каждый день, чтобы правильно распределять свое время.

Делайте небольшие перерывы во время учебы. В промежутках можно дать себе небольшую физическую нагрузку. Мозг лучше всего работает, когда умственный труд сменяется физическим. Прогуляйтесь, побегайте, поиграйте в баскетбол, попинайте мяч – помимо стимуляции умственной деятельности, это снимет стресс.

Отдых и контроль над волнением — одни из главных составляющих успеха при подготовке к тесту. Часто ошибки совершаются только из-за стресса, который мешает сконцентрироваться и собраться. Чтобы быть отдохнувшим и расслабленным, соблюдайте составленный режим и старайтесь высыпаться.

Экзамен. На экзамене оцениваются: 1) понимание и степень усвоения теории; 2) методическая подготовка; 3) знание фактического материала; 4) знакомство с обязательной литературой; 5) умение приложить теорию к практике, решать практические задачи и т. д.; 6) логика, структура и стиль ответа, умение защищать выдвигаемые положения. Но значение экзаменов не ограничивается проверкой знаний. Являясь естественным завершением работы студента, они способствуют обобщению и закреплению знаний и умений, приведению их в строгую систему, а также устранению возникших в процессе занятий пробелов.

Студенты готовятся к экзаменам по-разному. Одни из них прорабатывают лишь некоторые вопросы, выбранные наугад, другие стремятся запомнить весь материал подряд, не вникая глубоко в его суть. Работа при этом концентрируется на одном стремлении – сдать экзамен. Недостатки такой системы очевидны. Очевидно также, что подготовка не должна ограничиваться чтением лекционных записей. Первоначальные необработанные конспекты студента содержат факты, определения, выводы, сделанные преподавателем, но в них, как правило, слабо просматривается связующая идея курса, так как студент, записывая каждую лекцию в отдельности, редко способен сразу и достаточно точно уловить общую направляющую мысль. Поэтому конспект требует дополнительной обработки на основе использования учебников и рекомендованной литературы.

Существенные недостатки имеет и такой способ подготовки к экзаменам, как беглый просмотр всего материала. Он эффективен только на некоторых этапах планирования и закрепляющего повторения. Более надежный и целесообразный путь – это тщательная систематизация материала при вдумчивом повторении, установлении внутриспредметных связей, увязке различных тем и разделов, закреплении путем решения задач.

Перед экзаменом назначается консультация. Цель ее – дать ответы на вопросы, возникшие в ходе самостоятельной подготовки. Хотелось бы обратить особое внимание на важность предэкзаменационных консультаций. Здесь студент имеет полную возможность получить ответ на все неясные ему вопросы. А для этого он должен проработать до консультации весь курс. Кроме того, преподаватель будет отвечать на вопросы других студентов, что будет для вас повторением и закреплением знаний. И еще очень важное обстоятельство: лектор на консультации, как правило, обращает внимание на те разделы, по которым на предыдущих экзаменах ответы были неудовлетворительными, а также фиксирует внимание на наиболее трудных разделах курса. Некоторые студенты не приходят на консультации либо потому, что считают, что у них нет вопросов к лектору, либо полагают, что у них и так мало времени и лучше самому почитать материал по конспекту или в учебнике. Это глубокое заблуждение. Никакая другая работа не сможет принести столь значительного эффекта накануне экзамена, как консультация преподавателя.

Подготовку к экзамену следует начинать с первого дня изучения дисциплины. Как правило, на лекциях подчеркиваются наиболее важные и трудные вопросы или разделы курса, требующие внимательного изучения и обдумывания. Нужно эти вопросы выделить и обязательно постараться разобраться в них, не дожидаясь экзамена, проработать их, готовясь к лабораторным занятиям, попробовать самостоятельно решить несколько типовых задач. И если, несмотря на это, часть материала осталась неувоенной, ни в коем случае нельзя успокаиваться, надеясь на то, что это не попадет на экзамене. Факты говорят об обратном: если те или другие вопросы курса не вошли в экзаменационный билет, преподаватель может их задать (и часто задает) в виде дополнительных вопросов. Точно такое же отношение должно

быть выработано к вопросам и задачам, перечисленным в экзаменационной программе, выдаваемой студентам еще до экзамена. Обычно эти же вопросы и аналогичные задачи содержатся в экзаменационных билетах. Не следует оставлять без внимания ни одного раздела курса; если не удалось в чем-то разобраться самому, нужно обратиться к товарищам; если и это не помогло выяснить какой-либо вопрос до конца, нужно обязательно задать этот вопрос преподавателю на предэкзаменационной консультации. Чрезвычайно важно приучить себя к умению самостоятельно мыслить, учиться думать, понимать суть дела. Очень полезно после проработки каждого раздела восстановить в памяти содержание изученного материала, кратко записав это на листе бумаги. Если этого не сделать, то большая часть материала останется не понятой, а лишь формально заученной, и при первом же вопросе экзаменатора студент убедится в том, насколько поверхностно он усвоил материал.

4.5. ПРИМЕРНАЯ ТЕМАТИКА КУРСОВЫХ РАБОТ (ПРОЕКТОВ)

Не предусмотрено.

5. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

а) перечень литературы

1. Макшанов, А. В. Большие данные. Big Data [Электронный ресурс] / А. В. Макшанов, А. Е. Журавлев, Л. Н. Тындыкарь. - Электрон. текстовые дан. - Санкт-Петербург : Лань, 2021. - 188 с. - ЭБС "Лань". - Неогранич. доступ. - ISBN 978-5-8114-6810-2 — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165835> (дата обращения: 05.05.2022). — Режим доступа: для авториз. пользователей.

2. Грас, Джоэл. Data Science. Наука о данных с нуля [Текст] : учеб. пособие / Д. Грас. - СПб. : БХВ - Петербург, 2019. - 336 с. ; 23 см. - ISBN 978-5-9775-3758-2 (30 экз.).

3. Сенько, А. В. Работа с BigData в облаках. Обработка и хранение данных с примерами из Microsoft Azure [Электронный ресурс] / А. В. Сенько. - Электрон. текстовые дан. - Санкт-Петербург : Питер, 2019. - 448 с. : ил. - ЭБС "Айбукс". - Неогранич. доступ. - ISBN 978-5-4461-0578-6 — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://ibooks.ru/reading.php?short=1&productid=359208> (дата обращения: 05.05.2022). — Режим доступа: для авториз. пользователей.

4. Цильковский, И. А. Методы анализа знаний и данных [Электронный ресурс] : уч.-методич. пособие / И. А. Цильковский. — Новосибирск : Изд-во НГТУ, 2010. — 68 с. — Режим доступа: ЭБС «Руконт».

5. Сапрыкин, О. Н. Интеллектуальный анализ данных : учебное пособие / О. Н. Сапрыкин. — Самара : Самарский университет, 2020. — 80 с. — ISBN 978-5-7883-1563-8. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/188906> (дата обращения: 05.05.2022). — Режим доступа: для авториз. пользователей.

б) периодические издания

в) список авторских методических разработок:

Материалы курса, опубликованные в ИОС «DOMIC».

г) базы данных, информационно-справочные и поисковые системы _____

Google Colaboratory <https://colab.research.google.com/>

Draw.io <https://app.diagrams.net/>

6. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

6.1. УЧЕБНО-ЛАБОРАТОРНОЕ ОБОРУДОВАНИЕ:

Для проведения лекционных занятий необходима аудитория с презентационным оборудованием, для проведения лабораторных занятий необходима аудитория на 25-30 рабочих мест (в зависимости от численности учебной группы), оборудованная доской, презентационной техникой, компьютерами.

6.2. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ:

1. Браузер.
2. Python 3.
3. Anaconda.

6.3. ТЕХНИЧЕСКИЕ И ЭЛЕКТРОННЫЕ СРЕДСТВА:

ИОС DOMIC, презентационное оборудование, персональный компьютер с возможностью демонстрации презентаций в формате pdf.

7. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

При реализации данного курса используются следующие образовательные технологии: технологии традиционного обучения, игровые технологии, технологии проблемного обучения, технологии обучения в сотрудничестве, технологии контекстного обучения,

интерактивные технологии, технологии дистанционного обучения, активные педагогические технологии.

8. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

8.1. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ВХОДНОГО КОНТРОЛЯ

Не предусмотрено.

8.2. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ТЕКУЩЕГО КОНТРОЛЯ

Полный фонд оценочных средств находится в ИОС «Домик» по адресу domic.isu.ru.

8.3. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ПРОМЕЖУТОЧНОГО КОНТРОЛЯ

Контроль за изучением дисциплины осуществляется следующим способом: в течение семестра студенты выполняют лабораторные работы, заключающие в написании программ на языке Python с последующим анализом результатов, и отвечают на устные вопросы по некоторым темам. В течение семестра можно набрать максимум 80 баллов.

Экзамен проходит в устной форме и принимается по билетам. Каждый билет содержит один теоретический вопрос.

Критерии оценки:

1. Промежуточный (текущий) контроль призван определить, насколько глубоко студент изучил материал дисциплины согласно рабочей программы дисциплины, знает и правильно понимает содержание предмета.

2. Промежуточный (текущий) контроль выявить степень умения студента правильно применять полученные знания при анализе процессов и явления согласно предметной области в своей практической деятельности.

3. Промежуточный (текущий) контроль демонстрирует способность студента к самостоятельной работе в предметной области дисциплины по приобретению представлений, знаний и навыков.

Оценка «отлично» выставляется, если ответ студента в полной мере соответствует названным выше принципам. Твердо знает предмет, рекомендованную обязательную литературу, грамотно, аргументировано излагает основные вопросы и проблемы дисциплины, полно отвечает на дополнительные вопросы, умеет применять теоретические знания при анализе явлений, ситуаций и проблем в предметной области дисциплины.

Оценка «хорошо» предполагает наличие прочных знаний в объеме рабочей программы дисциплины, знание минимума рекомендованной литературы, умение анализировать явления, ситуации и проблемы в предметной области дисциплины.

Оценка «удовлетворительно» ставится, если студент показал в целом верные знания по всем вопросам билета, обнаружил понимание основного содержания экзаменационных вопросов, умение в целом применять теоретические знания при анализе явлений, ситуаций и проблем в предметной области дисциплины.

Оценка «неудовлетворительно» ставится в случае, если ответы студента не удовлетворяют требованиям к оценкам «удовлетворительно» или студент не выполнил требования квалификации.

При ответе на билеты установлено следующее соответствие.

При текущем контроле оценке «отлично» соответствует 30 баллов, оценке «хорошо» соответствует 25 баллов, оценке «удовлетворительно» соответствует 20 баллов.

При промежуточном контроле оценке «отлично» соответствует 80 баллов, оценке «хорошо» соответствует 70 баллов, оценке «удовлетворительно» соответствует 60 баллов.

Дополнительные баллы ставятся за ответы на сложные вопросы, активную работу в семестре, участие в олимпиадах по математике и информационным технологиям (до 20

баллов).

Общая оценка за промежуточный контроль выставляется в соответствии с балльно-рейтинговой системой (от 60 до 69 баллов – удовлетворительно, от 70 до 79 баллов – хорошо, от 80 баллов – отлично).

Примерные вопросы к экзамену:

1. Понятие больших данных, их основные свойства и применения.
2. OLTP- и OLAP-системы.
3. ETL-процессы, основные принципы организации. Поточковая и пакетная обработка информации.
4. Системы хранения данных. Хранилище данных. Озеро данных.
5. Хранилище данных. Трехслойная архитектура хранилища. Подходы к проектированию по Кимбаллу и Инмону.
6. Многомерное проектирование.
7. Подходы к хранению истории изменения данных. SCD.
8. Data Vault. Типы таблиц. Пример.
9. Якорное моделирование. Типы таблиц. Пример.
10. Распределенные модели и данные. Spark ML.

Разработчики:


(подпись)

доцент

(занимаемая должность)

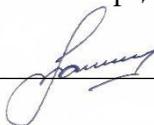
Казимиров А.С.

(Ф.И.О.)

Программа составлена в соответствии с требованиями ФГОС ВО по направлению подготовки 02.04.02 «Фундаментальная информатика и информационные технологии» (уровень магистратуры), утвержденный приказом Министерства образования и науки Российской Федерации от «23» августа 2017 г. № 811, зарегистрированный в Минюсте России «13» сентября 2017 г. № 48168 с изменениями и дополнениями от 26 ноября 2020 г., 8 февраля 2021 г.

Программа рассмотрена на заседании кафедры Алгебраических и информационных систем ИМИТ ИГУ «24» марта 2022 г.

Протокол № 9 Зав. кафедрой



Пантелеев В.И.

Настоящая программа не может быть воспроизведена ни в какой форме без предварительного письменного разрешения кафедры-разработчика программы.