**МИНИСТЕРСТВО науки и высшего образования Российской Федерации**

федеральное Государственное бюджетное образовательное учреждение высшего образования

**«ТЮМЕНСКИЙ ИНДУСТРИАЛЬНЫЙ университет»**

|  |
| --- |
| **Утверждаю:**  \_\_\_\_\_\_\_\_\_\_\_\_\_  «\_\_» \_\_\_\_2023г. |

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ**

|  |  |
| --- | --- |
| дисциплины: | **Системы хранения и обработки данных** |
| направление подготовки: | **01.04.02 Прикладная математика и информатика** |
| направленность (профиль): | **Машинное обучение и анализ данных** |
| форма обучения: | **очная, очно-заочная, заочная** |

Фонд оценочных средств рассмотрен на заседании кафедры математики и прикладных информационных технологий

Протокол № \_\_\_ от «\_\_\_» \_\_\_ \_\_\_\_\_\_\_ 2023 г.

**ПЕРЕЧЕНЬ КОМПЕТЕНЦИЙ И ИНДИКАТОРОВ ИХ ДОСТИЖЕНИЯ ПО**

**УЧЕБНОЙ ДИСЦИПЛИНЕ (МОДУЛЮ)**

|  |  |
| --- | --- |
| Код и наименование  компетенции | Код и наименование индикатора  достижения компетенции (ИДК) |
| ПКС-1 Способен управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных | ПКС-1.1. Способен проводить сравнительный анализ и выбор источников данных, средств хранения и обработки данных, оценку условий их приобретения и использования |
|
|
| ПКС-1.2. Способен оценивать экономические показатели (затраты на внедрение/модификацию, эффективность и пр.) технологий больших данных |
|
|
| ПКС-2 Способен разрабатывать (совершенствовать) и внедрять новые методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными | ПКС-2.1. Способен разрабатывать методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными |
|
|
| ПКС-2.2. Способен оптимизировать и внедрять методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными в профессиональной деятельности |
|
|

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ ДЛЯ ПРОВЕРКИ СФОРМИРОВАННОСТИ КОМПЕТЕНЦИЙ**

***Для оценки*** ПКС-1 *Способен управлять этапами жизненного цикла методологической и технологической инфраструктуры анализа больших данных*

**Тестовые вопросы открытого типа по дисциплине "Системы хранения и обработки данных"**

1. Какой тип базы данных использует таблицы с четко определенной схемой?

Ответ запишите все с большой буквы АБВГД.

**Ответ: РЕЛЯЦИОННАЯ**

1. Какой принцип RAID обеспечивает отказоустойчивость через зеркалирование?

Ответ запишите все с большой буквы ABC 5 на английском языке номер через пробел.

**Ответ: RAID 1**

1. Как называется масштабирование путем добавления новых серверов?

Ответ запишите все с большой буквы АБВГД.

**Ответ: ГОРИЗОНТАЛЬНОЕ**

1. Какой набор свойств гарантирует надежность транзакций в реляционных базах данных?

Ответ запишите все с большой буквы ABC на английском языке.

**Ответ: ACID**

***Для оценки ПКС-2*** *Способен разрабатывать (совершенствовать) и внедрять новые методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными*

1. Какая характеристика Big Data описывает высокие темпы поступления данных?

Ответ запишите все с большой буквы на русском языке АБВГД.

**Ответ: СКОРОСТЬ**

1. Согласно CAP-теореме, сколько свойств из трех может одновременно обеспечивать распределенная система?

Ответ запишите все с большой буквы АБВГД.

**Ответ: ДВА**

1. Какой механизм в базах данных ускоряет выполнение запросов?

Ответ запишите все с большой буквы АБВГД.

**Ответ: ИНДЕКСИРОВАНИЕ**

1. Какая модель NoSQL базы данных использует формат "ключ-значение"?

Ответ запишите все с большой буквы ABC- ABC на английском языке номер через девиз.

**Ответ: KEY-VALUE**

1. Как называется система хранения, где данные распределены по нескольким серверам?

Ответ запишите все с большой буквы АБВГД

**Ответ: РАСПРЕДЕЛЕННАЯ**

1. Как называется процесс временного хранения часто запрашиваемых данных для ускорения доступа?

Ответ запишите все с большой буквы АБВГД

**Ответ: КЕШИРОВАНИЕ**

**10 тестовых вопросов закрытого типа**

**(Выберите один правильный ответ)**

1. Какое свойство транзакции из модели ACID гарантирует, что данные останутся сохраненными после сбоя системы?

a) Атомарность

b) Долговечность \*(Верный ответ)\*

c) Согласованность

d) Изолированность

2. Какая из систем управления базами данных относится к категории нереляционных (NoSQL)?

a) MySQL

b) MongoDB \*(Верный ответ)\*

c) Oracle

d) PostgreSQL

3. Какой из этих уровней RAID используется для зеркалирования данных?

a) RAID 0

b) RAID 1 \*(Верный ответ)\*

c) RAID 5

d) RAID 6

4. Что из перечисленного чаще всего используется для работы с большими объемами данных (Big Data)?

a) Microsoft Access

b) Cassandra \*(Верный ответ)\*

c) SQLite

d) MariaDB

5. Какая архитектура чаще всего применяется в облачных системах хранения данных?

a) P2P

b) Централизованная

c) Распределенная \*(Верный ответ)\*

d) Модульная

**Для оценки ПКС-2** *Способен разрабатывать (совершенствовать) и внедрять новые методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными*

6. Какой уровень RAID не обеспечивает отказоустойчивость?

a) RAID 1

b) RAID 5

c) RAID 6

d) RAID 0 \*(Верный ответ)\*

7. В чем основная цель использования индексов в СУБД?

a) Шифрование данных

b) Уменьшение объемов хранимой информации

c) Ускорение поиска данных \*(Верный ответ)\*

d) Оптимизация транзакций

8. Что такое консистентность в контексте CAP-теоремы?

a) Доступность данных для каждого запроса

b) Устойчивость системы к сбоям

c) Единое состояние данных на всех узлах системы \*(Верный ответ)\*

d) Скорость обработки данных

9. Какой тип файловой системы чаще всего используется для распределенных систем хранения?

a) FAT32

b) NTFS

c) HDFS \*(Верный ответ)\*

d) ext4

10. Что означает термин "горизонтальное масштабирование"?

a) Увеличение вычислительной мощности одного сервера

b) Добавление новых серверов или устройств хранения \*(Верный ответ)\*

c) Снижение объема обрабатываемых данных

d) Оптимизация структур данных

**10 тестовых вопросов комбинированного типа**

**(Выберите один или несколько правильных ответов и/или дайте краткий ответ)**

1. Какие из перечисленных СУБД являются NoSQL-базами данных?

a) MongoDB \*(Верный ответ)\*

b) MySQL

c) Cassandra \*(Верный ответ)\*

d) PostgreSQL

2. Какие методы обеспечения отказоустойчивости используются в системах хранения данных?

a) Репликация \*(Верный ответ)\*

b) Индексация

c) Зеркалирование \*(Верный ответ)\*

d) Архивирование

3. Какие основные модели данных поддерживаются NoSQL базами?

Ответ: Ключ-значение, документная, графовая, столбцовая.

4. Что такое HDFS и для чего она используется?

a) Это распределенная файловая система \*(Верный ответ)\*

b) Это система управления базами данных

c) Это облачная платформа

d) Это система для работы с транзакциями

5. Какие преимущества предоставляет кеширование?

a) Ускорение доступа к данным \*(Верный ответ)\*

b) Снижение нагрузки на основное хранилище \*(Верный ответ)\*

c) Обеспечение избыточности данных

d) Сжатие данных

**Для оценки ПКС-2** *Способен разрабатывать (совершенствовать) и внедрять новые методы, модели, алгоритмы, технологии и инструментальные средства работы с большими данными*

6. Для каких задач чаще всего применяются графовые базы данных?

a) Хранение документов

b) Хранение сетевых связей \*(Верный ответ)\*

c) Хранение временных рядов

d) Анализ социальных сетей \*(Верный ответ)\*

7. Укажите, какие из перечисленных технологий предназначены для работы с большими данными:

a) Apache Hadoop \*(Верный ответ)\*

b) Oracle

c) Apache Spark \*(Верный ответ)\*

d) Microsoft Access

8. Что такое модель BASE и как она отличается от ACID?

Ответ:

BASE (Basic Availability, Soft state, Eventual consistency) — модель, применяемая в NoSQL базах данных. Она ориентирована на доступность и масштабируемость, в отличие от ACID, которая предоставляет жесткие гарантии согласованности.

9. Что из перечисленного относится к облачным платформам хранения данных?

a) Amazon S3 \*(Верный ответ)\*

b) HDFS

c) Microsoft Azure Storage \*(Верный ответ)\*

d) Dropbox \*(Верный ответ)\*

10. Какие методы используются для защиты данных в системах хранения?

a) Шифрование \*(Верный ответ)\*

b) Репликация \*(Верный ответ)\*

c) Индексация

d) Резервное копирование \*(Верный ответ)\*

**Критерии оценивания уровня сформированности компетенций**

Оценка «отлично» выставляется обучающемуся, если им даны правильные ответы более, чем на 91% вопросов.

Оценка «хорошо» выставляется обучающемуся, если им даны правильные ответы на 76-90% вопросов.

Оценка «удовлетворительно» выставляется обучающемуся, если им даны правильные ответы на 61-75% вопросов.

Оценка «неудовлетворительно» выставляется обучающемуся, если им даны правильные ответы менее чем на 60% вопросов.