МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования

«Рязанский государственный радиотехнический университет»

Кафедра «Безопасность жизнедеятельности и экология»

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

по дисциплине

«Облачные технологии и большие данные»

Направление

09.03.04 Программная инженерия

Профиль

Программная инженерия систем искусственного интеллекта
Программное обеспечение систем искусственного интеллекта

Квалификация выпускника – бакалавр

Форма обучения – очная

Рязань

1. ОБЩИЕ ПОЛОЖЕНИЯ

Оценочные материалы – это совокупность учебно-методических материалов и процедур, предназначенных для оценки качества освоения обучающимися данной дисциплины как части основной образовательной программы.

Цель – оценить соответствие знаний, умений и уровня приобретённых компетенций обучающихся целям и требованиям основной образовательной программы в ходе текущего контроля и промежуточной аттестации.

Основная задача – обеспечить оценку уровня сформированности компетенций и индикаторов их достижения, приобретаемых обучающимися в соответствии с требованиями ФГОС ВО.

Контроль знаний обучающихся проводится в форме текущего контроля (лабораторные работы) и промежуточной аттестации (зачёт).

2. ПЕРЕЧЕНЬ КОМПЕТЕНЦИЙ С УКАЗАНИЕМ ЭТАПОВ ИХ ФОРМИРОВАНИЯ

При освоении дисциплины формируются следующие компетенции:

- **ПК-9** – способен организовывать хранение данных, выбирая адекватные технологические решения (BD-3);
  - **ПК-9.1** – разрабатывает, отлаживает и тестирует прикладные решения с элементами ИИ с применением различных технологий хранения структурированных данных, оценивает качество построенных прикладных решений (BD-3.1);
  - **ПК-9.2** – разрабатывает, отлаживает и тестирует прикладные решения с элементами ИИ с применением различных технологий хранения неструктурированных данных, оценивает качество (BD-3.2).
- **ПК-10** – способен применять различные модели и (или) технологии обработки данных (BD-4);
  - **ПК-10.1** – осуществляет выбор технологий обработки больших данных, приемлемых для создания прикладной системы ИИ с заданными требованиями (BD-4.1);
  - **ПК-10.2** – разрабатывает и отлаживает прикладные решения с элементами ИИ с применением различных технологий обработки данных (BD-4.2);
  - **ПК-10.3** – тестирует, испытывает и оценивает качество решений с элементами ИИ, реализованных с использованием технологий обработки данных (BD-4.3).
- **ПК-16.2** – разрабатывает и поддерживает системы обработки больших данных различной степени сложности (PL-2.2).
- **ПК-25** – способен применять технологии организации инфраструктуры больших данных (BD-5);

- **ПК-25.1** – осуществляет выбор направления вспомогательных технологических решений для формирования единого стека работы с большими данными (BD-5.1);
- **ПК-25.2** – разрабатывает и отлаживает прикладные решения с элементами ИИ с применением различных технологий организации инфраструктуры БД (BD-5.2).

Указанные компетенции формируются в соответствии со следующими этапами:

- формирование и развитие теоретических знаний (лекционные занятия, самостоятельная работа);
- приобретение и развитие практических умений (лабораторные работы, самостоятельная работа);
- закрепление знаний и навыков (выполнение и защита лабораторных работ, подготовка к зачёту).

3. ПОКАЗАТЕЛИ И КРИТЕРИИ ОЦЕНИВАНИЯ КОМПЕТЕНЦИЙ

Сформированность каждой компетенции оценивается по трёхуровневой шкале:

- **пороговый уровень** – обязательный для всех обучающихся; соответствует освоению более 40% приведённых знаний, умений и навыков;
- **продвинутый уровень** – соответствует освоению более 60% знаний, умений и навыков;
- **эталонный уровень** – соответствует освоению более 80% знаний, умений и навыков.

При освоении менее 40% компетенция считается несформированной.

Таблица 1. Показатели достижения индикаторов компетенций

Компетенция	Индикаторы (краткое содержание)	Этап	Оценочное средство
ПК-9.1	Знать основные принципы организации реляционных систем хранения данных, принципы организации распределённых реляционных систем; уметь создавать базы данных в реляционных СУБД, писать SQL-запросы; владеть навыками создания решений распределённых хранилищ.	7	Зачёт
ПК-9.2	Знать основные технологии NoSQL, популярные хранилища класса ключ-значение, документные, колоночные и	7	Зачёт

Компетенция	Индикаторы (краткое содержание)	Этап	Оценочное средство
	графовые БД; уметь выбирать технологии и механизмы хранения и доступа к большим данным; владеть командами работы с хранилищами NoSQL.		
ПК-10.1	Знать принципы организации распределённых хранилищ данных; уметь выбирать технологии обработки больших данных; владеть пониманием выбора технологий Grid, MapReduce, Hadoop, Spark.	7	Зачёт
ПК-10.2	Знать принципы организации параллельной обработки хранилищ на базе современных технологий больших данных; уметь разрабатывать решения обработки больших данных; владеть навыками создания решений на базе параллельной, потоковой и пакетной обработки.	7	Зачёт
ПК-10.3	Знать принципы тестирования и оценивания качества решений параллельной, потоковой и пакетной обработки больших данных; уметь тестировать и оценивать качество решений; владеть навыками тестирования.	7	Зачёт
ПК-16.2	Знать фреймворки для пакетной обработки данных (Spark, MapReduce), инструменты потоковых вычислений (Kafka, Flink), аналитические БД (ClickHouse, Greenplum); уметь адаптировать алгоритмы под большие данные, разрабатывать ETL; владеть принципами использования фреймворков.	7	Зачёт

Компетенция	Индикаторы (краткое содержание)	Этап	Оценочное средство
ПК-25.1	Знать методологию создания инфраструктуры БД; уметь выполнять отдельные функции в проектах по созданию инфраструктуры БД; владеть навыками выбора технологических решений для формирования единого стека работы с большими данными.	7	Зачёт
ПК-25.2	Знать компоненты инфраструктуры больших данных; уметь применять прикладные решения с элементами ИИ и различными технологиями организации инфраструктуры БД; владеть навыками участия в разработке решений.	7	Зачёт

4. ПАСПОРТ ОЦЕНОЧНЫХ МАТЕРИАЛОВ ПО ДИСЦИПЛИНЕ

Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции (или её части)	Наименование оценочного средства
Тема 1.1. Введение в облачные вычисления (история, принципы, монолитная/сервис-ориентированная/микросервисная архитектуры, методология 12 факторов)	ПК-9.1, ПК-9.2, ПК-10.1, ПК-10.2, ПК-10.3, ПК-16.2, ПК-25.1, ПК-25.2	Зачёт
Тема 1.2. Микросервисы (Gateway API, Rate limiter, API Composition, SAGA, высокая доступность, внешняя конфигурация, обнаружение сервисов)	ПК-9.1, ПК-9.2, ПК-10.1, ПК-10.2, ПК-10.3, ПК-16.2, ПК-25.1, ПК-25.2	Зачёт
Тема 1.3. Сети и протоколы (OSI, TCP/IP, HTTP, WebSocket, сериализация XML/JSON/BSON/ProtoBuf)	ПК-9.1, ПК-9.2, ПК-10.1, ПК-10.2, ПК-10.3, ПК-16.2, ПК-25.1, ПК-25.2	Зачёт

Контролируемые разделы (темы) дисциплины	Код контролируемой компетенции (или её части)	Наименование оценочного средства
Тема 1.4. Виртуализация (Docker, Kubernetes, архитектура, кластеры)	ПК-9.1, ПК-9.2, ПК-10.1, ПК-10.2, ПК-10.3, ПК-16.2, ПК-25.1, ПК-25.2	Зачёт
Тема 2.1. Концепции больших данных (принципы, характеристики, хранение: MongoDB, HBase, Redis, DynamoDB, Cassandra)	ПК-9.1, ПК-9.2, ПК-10.1, ПК-10.2, ПК-10.3, ПК-16.2, ПК-25.1, ПК-25.2	Зачёт
Тема 2.2. Модели программирования для больших данных (MapReduce, рабочие потоки, передача сообщений, массовый синхронный параллелизм)	ПК-9.1, ПК-9.2, ПК-10.1, ПК-10.2, ПК-10.3, ПК-16.2, ПК-25.1, ПК-25.2	Зачёт
Тема 2.3. Инструменты обработки больших данных (Hadoop, Spark, Storm, Hive, Pig, ClickHouse, Greenplum)	ПК-9.1, ПК-9.2, ПК-10.1, ПК-10.2, ПК-10.3, ПК-16.2, ПК-25.1, ПК-25.2	Зачёт
Тема 2.4. Сравнение инструментов программирования (сравнительный анализ характеристик систем)	ПК-9.1, ПК-9.2, ПК-10.1, ПК-10.2, ПК-10.3, ПК-16.2, ПК-25.1, ПК-25.2	Зачёт

5. ТИПОВЫЕ КОНТРОЛЬНЫЕ ЗАДАНИЯ

5.1. Промежуточная аттестация (зачёт)

Зачёт проводится в форме устного ответа по билетам, включающим теоретический вопрос и практическое задание (задачу). Также может использоваться тестирование.

5.1.1. Тестовые вопросы закрытого типа (20 вопросов)

1. Что такое облачные вычисления?
а) Модель предоставления повсеместного и удобного сетевого доступа к общему пулу конфигурируемых вычислительных ресурсов;
б) Локальная сеть организации;
в) Тип базы данных;

г) Метод шифрования.
Правильный ответ: а)

2. Какая архитектура предполагает разбиение приложения на независимо развёртываемые сервисы?
а) Монолитная;
б) Микросервисная;
в) Файловая;
г) Иерархическая.
Правильный ответ: б)

3. Сколько факторов включает методология «12-факторного приложения»?
а) 10;
б) 12;
в) 15;
г) 8.
Правильный ответ: б)

4. Что такое Docker?
а) Технология контейнеризации приложений;
б) База данных;
в) Язык программирования;
г) Сетевой протокол.
Правильный ответ: а)

5. Что такое Kubernetes?
а) Система оркестрации контейнеров;
б) Реляционная СУБД;
в) Фреймворк для машинного обучения;
г) Web-сервер.
Правильный ответ: а)

6. Какой компонент Kubernetes запускает контейнеры?
а) Pod;
б) Service;
в) Volume;
г) Namespace.
Правильный ответ: а)

7. Что такое NoSQL?
а) Класс систем управления базами данных, отличающихся от реляционной модели;
б) Язык запросов;
в) Протокол передачи данных;
г) Архитектура микросервисов.
Правильный ответ: а)

8. Какая из перечисленных СУБД является документной?
а) MongoDB;
б) PostgreSQL;
в) MySQL;
г) Oracle.
Правильный ответ: а)

9. Какая из перечисленных СУБД является колоночной?
а) ClickHouse;
б) MongoDB;
в) Redis;
г) Neo4j.
Правильный ответ: а)

10. Что такое MapReduce?
а) Модель параллельных вычислений для обработки больших данных;
б) Язык программирования;
в) СУБД;
г) Протокол передачи данных.
Правильный ответ: а)

11. Какой фреймворк используется для пакетной обработки больших данных?
а) Apache Spark;
б) Nginx;
в) Flask;
г) Django.
Правильный ответ: а)

12. Какой инструмент используется для потоковой обработки данных?
а) Apache Kafka;
б) Apache Hadoop;
в) Apache Pig;
г) Apache Hive.
Правильный ответ: а)

13. Что такое ETL?
а) Extract, Transform, Load – процесс извлечения, преобразования и загрузки данных;
б) Язык программирования;
в) Сетевой протокол;
г) Тип базы данных.
Правильный ответ: а)

14. Что такое API Gateway?
а) Компонент, обеспечивающий единую точку входа для клиентов к микросервисам;
б) База данных;
в) Фреймворк для ML;

г) Операционная система.
Правильный ответ: а)

15. Что такое Circuit Breaker?
а) Паттерн для предотвращения каскадных отказов в микросервисной архитектуре;
б) Тип базы данных;
в) Протокол передачи данных;
г) Алгоритм шифрования.
Правильный ответ: а)

16. Что такое виртуализация?
а) Технология разделения вычислительных ресурсов между несколькими средами;
б) Тип базы данных;
в) Язык программирования;
г) Сетевой протокол.
Правильный ответ: а)

17. Что такое Docker Image?
а) Шаблон для создания контейнера;
б) Запущенный контейнер;
в) Сетевой интерфейс;
г) Файловая система.
Правильный ответ: а)

18. Какой из перечисленных инструментов используется для SQL-подобного программирования в больших данных?
а) Apache Hive;
б) Docker;
в) Flask;
г) Kubernetes.
Правильный ответ: а)

19. Что такое Greenplum?
а) Аналитическая база данных на основе PostgreSQL;
б) Web-сервер;
в) Язык программирования;
г) Протокол передачи данных.
Правильный ответ: а)

20. Что такое SAGA?
а) Паттерн управления распределёнными транзакциями в микросервисах;
б) Тип базы данных;
в) Язык запросов;
г) Фреймворк для ML.
Правильный ответ: а)

5.1.2. Тестовые вопросы открытого типа (10 вопросов)

1. Опишите методологию «12-факторного приложения» и её значение для разработки облачно-ориентированных приложений.

2. Объясните разницу между монолитной, сервис-ориентированной и микросервисной архитектурами.

3. Опишите архитектуру Docker и основные компоненты (Docker Image, Dockerfile, Docker Container, Docker Daemon).

4. Перечислите основные компоненты архитектуры Kubernetes и их назначение.

5. Сравните реляционные и NoSQL базы данных. Приведите примеры СУБД для каждого типа.

6. Опишите модель программирования MapReduce и её применение для обработки больших данных.

7. Сравните Apache Spark и Apache Hadoop по основным характеристикам.

8. Опишите принципы работы Apache Kafka и её применение для потоковой обработки данных.

9. Что такое аналитические базы данных? Приведите примеры (ClickHouse, Greenplum) и опишите их особенности.

10. Опишите процесс ETL и его роль в пайплайнах обработки больших данных.

5.1.3. Типовые практические задания (задачи) к зачёту

1. Реализовать 12-факторное приложение на Python с использованием фреймворка Flask.
2. Разработать микросервисную архитектуру с API Gateway и Circuit Breaker.
3. Контейнеризовать приложение с использованием Docker (написать Dockerfile, собрать образ, запустить контейнер).
4. Выполнить базовые операции с MongoDB: создание коллекции, вставка, чтение, обновление и удаление документов.
5. Реализовать обработку больших данных с использованием Apache Spark и ClickHouse.
6. Провести сравнительный анализ фреймворков обработки больших данных (Spark, Kafka, ClickHouse, Greenplum).
7. Разработать ETL-пайплайн для обработки данных с использованием Apache Spark.
8. Настроить кластер Kubernetes и развернуть в нём контейнеризированное приложение.
9. Реализовать потоковую обработку данных с использованием Apache Kafka.
10. Выполнить аналитический запрос к ClickHouse на основе большого набора данных.

5.1.4. Вопросы к зачёту (30 вопросов)

1. История и предпосылки появления облачных вычислений. Общие принципы и закономерности.
2. Монолитная, сервис-ориентированная и микросервисная архитектуры.
3. Методология двенадцати факторов: кодовая база, зависимости, конфигурация, внешние сервисы.
4. Методология двенадцати факторов: сборка, релиз, выполнение, процессы, привязка портов.
5. Методология двенадцати факторов: параллелизм, утилизируемость, паритет разработки и производства, журналирование, административные процессы.
6. Микросервисы. Отдельная база данных для каждого сервиса. Gateway API. Rate limiter. API Composition.
7. SAGA. Высокая доступность. Особенности облачно-ориентированных приложений.
8. Внешняя конфигурация. Обнаружение сервисов. Автоматический предохранитель (Circuit Breaker).
9. Модель OSI. Протоколы сети. IP, TCP, установка соединения, подтверждение доставки.
10. Управление потоком передачи. Контроль перегрузки канала. MTU. Протоколы ARP, STP, OSPF, BGP, DNS, GeoDNS.
11. Протокол HTTP. Методы HTTP. Статусы HTTP. Заголовки HTTP. Согласование содержимого. Cookie. Сессия.
12. Real-Time коммуникация. WebSocket. HTTP/2.
13. Протоколы сериализации. Символьные протоколы. Бинарные протоколы. XML. JSON. BSON. ProtoBuf.
14. Виртуализация. Типы виртуализации. Контейнеризация. Docker.
15. Архитектура Docker. Docker Image. Dockerfile. Docker Container. Docker Daemon.
16. Кластеры Kubernetes. Архитектура Kubernetes. Node. Pod. Service. Volumes. Namespace. Secrets.
17. Принципы и характеристики больших данных. Хранение больших данных.
18. MongoDB. HBase. Redis. DynamoDB. Cassandra. Масштабируемый анализ данных.
19. Параллельные вычисления. Модель на основе MapReduce.
20. Модель на основе рабочего потока. Модель на основе передачи сообщений.
21. Модель на основе массового синхронного параллелизма. SQL и NoSQL модели.
22. Модель на основе разделённого глобального адресного пространства.
23. Инструменты программирования на основе модели MapReduce. Apache Hadoop.
24. Инструменты программирования на основе рабочих потоков. Apache Spark. Apache Storm.
25. Инструменты программирования на основе передачи сообщений. Spark GraphX.
26. Инструменты SQL-подобного программирования. Apache Hive. Apache Pig.
27. Аналитические базы данных. Колоночные СУБД. ClickHouse. Greenplum.
28. Сравнительный анализ характеристик систем обработки больших данных: входные данные, класс приложения, инфраструктура.
29. ETL-процессы в пайплайнах обработки больших данных. Инструменты и подходы.
30. Перспективные направления развития облачных технологий и технологий больших данных.

6. КРИТЕРИИ ОЦЕНИВАНИЯ

6.1. Шкала оценивания тестирования

Шкала	Критерий
3 балла (эталонный уровень)	Процент верных ответов от 85 до 100%
2 балла (продвинутый уровень)	Процент верных ответов от 75 до 84%
1 балл (пороговый уровень)	Процент верных ответов от 60 до 74%
0 баллов	Процент верных ответов от 0 до 59%

6.2. Шкала оценивания теоретического вопроса

Шкала	Критерий
3 балла (эталонный уровень)	Дан полный ответ, показаны глубокие систематизированные знания, приведены примеры, даны ответы на дополнительные вопросы.
2 балла (продвинутый уровень)	Дан полный ответ, но на некоторые дополнительные вопросы ответы да с помощью наводящих вопросов.
1 балл (пороговый уровень)	Дан неполный ответ, на дополнительные вопросы ответы даны только с помощью преподавателя.
0 баллов	Студент не смог ответить на вопрос.

6.3. Шкала оценивания практического задания

Шкала	Критерий
3 балла (эталонный уровень)	Задание решено верно, получены корректные результаты, сделан выводы.
2 балла (продвинутый уровень)	Задание решено верно, но имеются незначительные технические неточности.

Шкала	Критерий
1 балл (пороговый уровень)	Задание решено верно с дополнительными наводящими вопроса преподавателя.
0 баллов	Задание не решено.

6.4. Итоговая оценка

На зачёт выносится тестовое задание, практическое задание и теоретический вопрос.
Максимальный балл – 9.
Оценка «зачтено» выставляется при сумме баллов не менее 5 и выполнении всех предусмотренных лабораторных работ.
Оценка «не зачтено» – при сумме менее 5 баллов или наличии несданных лабораторных работ.

Оператор ЭДО ООО "Компания "Тензор"
ДОКУМЕНТ ПОДПИСАН ЭЛЕКТРОННОЙ ПОДПИСЬЮ
СОГЛАСОВАНО **ФГБОУ ВО "РГРТУ", РГРТУ,** Овечкин Геннадий Владимирович, Заведующий кафедрой ВПМ **05.11.26** 14:46 (MSK) Простая подпись