

Министерство сельского хозяйства Российской Федерации
Департамент научно-технологической политики и образования
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Красноярский государственный аграрный университет»

СОГЛАСОВАНО:

Директор ИЭиУ АПК
Шапорова З.Е.

« 27 » марта 2025 г.

УТВЕРЖДАЮ:

Ректор
Пыжикова Н.И.

« 28 » марта 2025 г.



ДОКУМЕНТ ПОДПИСАН
УСИЛЕННОЙ КВАЛИФИЦИРОВАННОЙ
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

ВЫДАННОЙ: ФГБОУ ВО КРАСНОЯРСКИЙ ГАУ
ВЛАДЕЛЕЦ: РЕКТОР ПЫЖИКОВА Н.И.
ДЕЙСТВИТЕЛЕН: 15.05.2025 - 08.08.2026

ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
(текущего оценивания, промежуточной аттестации)

Институт Экономики и управления АПК

Кафедра Информационные технологии и математическое обеспечение
информационных систем

Наименование и код ОПОП: 09.04.03 «Прикладная информатика»

Направленность (профиль): Цифровые технологии в агропромышленном
комплексе

Дисциплина: **Технологии обработки больших данных**

Красноярск 2025

Составитель: Миндалёв И.В., доцент

(ФИО, ученая степень, ученое звание)

«21»марта 2025 г.

Эксперт: Чубарь А.В к.т.н., доцент

(ФИО, ученая степень, ученое звание)

«21»марта 2025 г.

ФОС разработан в соответствии с рабочей программой дисциплины.

ФОС обсужден на заседании кафедры Информационные технологии и математическое обеспечение информационных систем протокол

№ 7 «21» марта 2025 г.

Зав. кафедрой Калитина Вера Владимировна, к.п.н., доцент

(ФИО, ученая степень, ученое звание)

«21»марта 2025 г.

ФОС принят методической комиссией института Экономики и управления

АПК протокол № 7 «24» марта 2025 г.

Председатель методической комиссии Рожкова А.В.

«24»марта 2025 г.

Содержание

1	Цель и задачи фонда оценочных средств.....	4
2	Нормативные документы.....	4
3	Перечень компетенций с указанием этапов их формирования в процессе освоения дисциплины. Формы контроля формирования компетенций.	4
4	Показатели и критерии оценивания компетенций	4
5	Фонд оценочных средств.	5
5.1	Фонд оценочных средств для текущего контроля.....	5
5.1.1.	Оценочное средство (тестирование). Критерии оценивания	5
5.1.2.	Оценочное средство: лабораторные работы. Критерии оценивания.....	5
5.2	Фонд оценочных средств для промежуточного контроля.....	6
5.2.1	Оценочное средство (зачет с оценкой). Критерии оценивания	6
6	Учебно-методическое и информационное обеспечение дисциплины	7
6.1	Основная литература	7
6.2	Дополнительная литература	7
6.3	Методические указания, рекомендации и другие материалы к занятиям	7
6.4	Программное обеспечение.....	7
6.5	Перечень ресурсов информационно-телекоммуникационной сети «Интернет»	8

1 Цель и задачи фонда оценочных средств

Целью создания ФОС дисциплины «Технологии обработки больших данных» является установление соответствия учебных достижений запланированным результатам обучения и требованиям образовательных программ и рабочих программ модулей.

ФОС по дисциплине решает задачи:

- контроль и управление процессом приобретения магистрантами необходимых знаний, умений, навыков и уровня сформированности компетенции, определённых в ФГОС ВО по направлению 09.04.03 «Прикладная информатика»;
- контроль и управление достижением целей реализации ОПОП, определенных в виде набора профессиональных компетенций выпускников;
- обеспечение соответствия результатов обучения задачам будущей профессиональной деятельности через совершенствование традиционных и внедрение инновационных методов обучения в образовательный процесс университета.

Назначение фонда оценочных средств:

Используется для оперативного и регулярного управления учебной деятельностью (в том числе самостоятельной) магистрантов. А также предназначен для оценки степени достижения запланированных результатов обучения по завершению изучения дисциплины «Технологии обработки больших данных» в установленной учебным планом форме в 3 семестре – зачет с оценкой.

2 Нормативные документы

ФОС разработан на основе Федерального государственного образовательного стандарта высшего образования по направлению подготовки **09.04.03** «Прикладная информатика», рабочей программы дисциплины «Технологии обработки больших данных».

3 Перечень компетенций с указанием этапов их формирования в процессе освоения дисциплины. Формы контроля формирования компетенций.

Компетенция	Этап формирования компетенции	Образовательные технологии	Тип контроля	Форма контроля
Способность применять современные методы и инструментальные средства прикладной информатики для автоматизации и информатизации решения прикладных задач различных классов и создания ИС (ПК-1)	теоретический (информационный)	лекции, самостоятельная работа	текущий	Тестирование
	практико-ориентированный	лабораторные работы, самостоятельная работа	текущий	лабораторные работы
	оценочный	аттестация	промежуточный	Зачет с оценкой

4 Показатели и критерии оценивания компетенций

Таблица 4.1 – Показатели и критерии оценки результатов обучения

Показатель оценки результатов обучения	Критерий оценки результатов обучения
--	--------------------------------------

ПК–1 -Способность применять современные методы и инструментальные средства прикладной информатики для автоматизации и информатизации решения прикладных задач различных классов и создания ИС	
Пороговый уровень	Знать: методы системного анализа, математические методы формализации, понятие больших данных, основы построения систем обработки и анализа больших данных.
Продвинутый уровень	Уметь: применять системный подход и математические методы в формализации и решении прикладных задач, использовать технологии баз данных, адаптированных для систем с большими данными.
Высокий уровень	Владеть: навыками применения методов на логическом, математическом и алгоритмическом уровнях, методиками автоматизации обработки больших данных.

Таблица 4.2 – Шкала оценивания

Показатель оценки результатов обучения	Шкала оценивания
Пороговый уровень	60-72 баллов (удовлетворительно)
Продвинутый уровень	73-86 баллов (хорошо)
Высокий уровень	87-100 баллов (отлично)

5 Фонд оценочных средств.

5.1 Фонд оценочных средств для текущего контроля

Текущий контроль используется для оперативного и регулярного управления учебной деятельностью (в том числе самостоятельной) магистрантов. В условиях рейтинговой системы контроля результаты текущего оценивания магистранта используются как показатель его текущего рейтинга. Текущий контроль успеваемости магистрантов включает в себя: тестирование, выполнение лабораторных работ.

5.1.1. Оценочное средство (тестирование). Критерии оценивания

Тестирование проводится с целью контроля по окончании каждого тематического модуля, с помощью ДОТ на сайте <https://e.kgau.ru>, каждый студент проходит тестирование (время прохождения теста – не ограничено) в компьютерном классе или на персональном компьютере, тест-билет содержит 40 вопросов. Банк тестовых заданий приведен в приложении.

Критерии оценивания		
Количество правильных ответов	Процент выполнения	Оценка
35-40	более 87 %	10 баллов
29-34	73-86 %	7 баллов
24-28	60-72 %	4 балла
0-23	менее 60%	0 баллов

Итого за тестирование в течение 4 семестра магистрант может набрать 30 баллов.

5.1.2. Оценочное средство: лабораторные работы. Критерии оценивания

По 1 модулю предусмотрено решение 2 лабораторных работ. По 2 модулю предусмотрено решение 9 лабораторных работ.

Максимальное количество баллов за выполненную лабораторную работу – 5 баллов.

Инструкции по выполнению работ представлены в moodle раздел текущие задания.

Критерии оценивания

Процент выполнения	Оценка
Выполнены все пункты задания, результат загружен в moodle	5 баллов / зачтено
Не выполнены все пункты задания, результат не загружен в moodle	0 баллов / незачтено

Итого за выполнение лабораторных работ в течение 4 семестра магистрант может набрать 55 баллов.

5.2 Фонд оценочных средств для промежуточного контроля

ФОС промежуточной аттестации обучающихся по дисциплине предназначен для оценки степени достижения запланированных результатов обучения по завершению изучения дисциплины в установленной учебным планом форме: зачет – в 4 семестре.

В ходе текущего контроля проводится оценивание качества изучения и усвоения магистрантами учебного материала по разделам, темам, модулям (логически завершенной части учебного материала) в соответствии с требованиями программы.

5.2.1 Оценочное средство (зачет с оценкой). Критерии оценивания

Зачет с оценкой по дисциплине "Технологии обработки больших данных" проводится в виде итогового тестирования по вопросам основных тем.

Для допуска к промежуточному контролю магистрант должен набрать необходимое количество баллов по итогам текущей аттестации – 40-80 баллов.

Итоговое тестирование включает создание ментальных карт (mindmap) на основе зачетных вопросов. Каждая карта должна включать как минимум 30 взаимосвязанных терминов.

Вопросы к итоговому контрольному тестированию

1. Основные характеристики больших данных
2. Роль больших данных в АПК
3. Консолидация данных
4. Визуализация данных
5. Основные конструкции языка R
6. HDFS - основы организации
7. Архитектура Hadoop
8. Исполнение Map/Reduce
9. Исполнение программ в Hadoop
10. Основы YARN
11. Аналитика потоковых данных в платформе Storm
12. Архитектура Apache Spark
13. Организация данных в Apache Spark

Критерии оценивания

Оценивание итогового тестирования осуществляется по следующим критериям:

- Магистрант, давший правильные ответы 85-100% терминов, получает максимальное количество баллов – 15.
- Магистрант, давший правильные ответы в пределах 70-85% терминов, получает 10 баллов.
- Магистрант, давший правильные ответы в пределах 60-70% терминов, получает 5 баллов.
- Магистрант, давший правильные ответы на менее чем 60% терминов, не набирает баллов и приходит на контрольное тестирование снова.

Баллы, полученные на итоговом тестировании суммируются с баллами, полученными в течение семестра на текущей аттестации и выводится итоговая оценка по следующим критериям:

- 60 – 100 – минимальное количество баллов – оценка «зачтено».
- 0 – 59 – среднее количество баллов – оценка «не зачтено».

Обучающийся, не сдавший зачет, приходит на передачу в сроки в соответствии с графиком ликвидации академических задолженностей: <http://www.kgau.ru>.

6 Учебно-методическое и информационное обеспечение дисциплины

6.1 Основная литература

1. Железнов, М. М. Методы и технологии обработки больших данных : учебно-методическое пособие / М. М. Железнов. — Москва : МИСИ – МГСУ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/145102> (дата обращения: 29.09.2021).
2. Нурматова, Е. В. Управление большими базами данных и высоконагруженными системами : учебное пособие / Е. В. Нурматова, Р. Ф. Халабия, Л. В. Бунина. — Москва : РТУ МИРЭА, 2019. — 120 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/171496> (дата обращения: 29.09.2021)

6.2 Дополнительная литература

3. Информационные технологии. Базовый курс : учебник для вузов / А. В. Костюк, С. А. Бобонец, А. В. Флегонтов, А. К. Черных. — 3-е изд., стер. — Санкт-Петербург : Лань, 2021. — 604 с. — ISBN 978-5-8114-8776-9. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/180821> (дата обращения: 29.09.2021)

6.3 Методические указания, рекомендации и другие материалы к занятиям

- 1 Аникин Ю., Борисова И., Загоруйко Н., Зырянов А., Павловский Е. «Введение в аналитику больших массивов данных». <https://www.intuit.ru/studies/courses/12385/1181/info>.
- 2 Сузи Р. Язык программирования Python. <https://www.intuit.ru/studies/courses/49/49/info>
- 3 Уайт Т. Nadoor: Подробное руководство. — СПб.: Питер, 2013. — 672 с.: ил. — (Серия «Бестселлеры O'Reilly»).

6.4 Программное обеспечение

Лицензионное ПО Красноярского ГАУ

1. Операционная система Astra Linux (лицензия № 192400033-alse-1.7-client-base_orel-x86_64-0-12913 от 28.08.2023).
2. Офисный пакет приложений Libre Office входит в комплект поставки Astra Linux.
3. Офисный пакет приложений Мой Офис (лицензия № ПР0000-35377 от 24.07.2024).
4. 1С Предприятие 8.2 (акт предоставления прав № Тг059122 от 24.10.2012).
5. Справочная правовая система "Консультант+" (договор № 20175200211 от 22.04.2020).
6. Moodle 3.5.6a (договор № 969.2 от 17.04.2020).

Свободно-распространяемое ПО или бесплатная лицензия с открытым исходным кодом:

1. ГИС Панорама x64 версия 15 мультиплатформенная лицензия (104622 фиксированная лицензия)

2. PostgreSQL; SWI-Prolog, Ramus Educational; StarUML; XMind v3.0; QT Creator, Oracle VM Virtual Box; DBeaver Community; MySQL Community Edition; Gimp; Wireshark; Graphical Network Simulator-3; NASM; SMath Studio; OpenJDK; Notepad++; LibreCad; Yandex (браузер).

6.5 Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Интернет-ресурсы

1. Введение в аналитику больших массивов данных

<https://www.intuit.ru/studies/courses/12385/1181/info>

2. Python: основы и применение <https://stepik.org/course/512/promo>

Электронные библиотечные системы

1. Каталог библиотеки Красноярского ГАУ -- www.kgau.ru/new/biblioteka/ ;
2. Центральная научная сельскохозяйственная библиотека - www.cnsnb.ru/ ;
3. Научная электронная библиотека "eLibrary.ru" – www.elibrary.ru ;
4. Электронная библиотечная система «Лань» - <https://e.lanbook.com/>
5. Электронно-библиотечная система «Юрайт» - <https://urait.ru/>
6. Электронно-библиотечная система «AgriLib» - <http://ebs.rgazu.ru/>
7. Электронная библиотека Сибирского Федерального университета - <https://bik.sfukras.ru/>
8. Национальная электронная библиотека - <https://rusneb.ru/>
9. Электронная библиотечная система «ИРБИС64+» - http://5.159.97.194:8080/cgi-bin/irbis64r_plus/cgiirbis_64_ft.exe?C21COM=F&I21DBN=IBIS_FULLTEXT&P21DBN=IBIS&Z21ID=&S21CNR=5
10. Электронный каталог Государственной универсальной научной библиотеки Красноярского края - <https://www.kraslib.ru/>
11. Научная электронная библиотека «КиберЛенинка». <https://cyberleninka.ru>

Информационно-справочные системы

1. Справочно-правовая система КонсультантПлюс

<http://www.consultant.ru/cons/cgi/online.cgi?req=home;rnd=0.8636296761039928>

2. Информационно-правовой портал «Гарант» <http://www.garant.ru/>

Профессиональные базы данных

1. Проглиб <https://proglib.io>
2. ХабрХабр <https://habr.com/ru/>

Таблица – Тип тестового задания

Тип задания	Наименование
1	Задания закрытого типа на установление соответствия
2	Задания закрытого типа на установление последовательности
3	Задания комбинированного типа, предполагающие выбор одного правильного ответа из предложенных
4	Задания комбинированного типа, предполагающие выбор нескольких ответов из предложенных
5	Задания открытого типа, в том числе с развёрнутым ответом

Тип задания	Задание	
3	1. Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какой из следующих вариантов является ключевым характеристикой больших данных? 1. Низкая скорость обработки. 2. Малый объем данных. 3. Высокая вариативность и разнообразие данных. 4. Низкая ценность информации	3 Большие данные характеризуются тем, что поступают из разных источников и имеют разные форматы (тексты, изображения, видео и т.д.). Это и есть высокая вариативность.
3	2. Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Что такое "5 V" в контексте больших данных? 1. 5 основных типов данных. 2. 5 этапов обработки данных. 3. 5 ключевых характеристик больших данных: Объем, Скорость, Разнообразие, Ценность и Верификация. 4. 5 методов анализа данных.	3 "5 V" — это международно признанная концепция, описывающая пять основных характеристик Big Data: объем (Volume), скорость (Velocity), разнообразие (Variety), ценность (Value) и достоверность (Veracity)
3	3. Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какая из следующих задач является основной при обработке больших данных? 1. Уменьшение объема данных. 2. Увеличение скорости обработки данных. 3. Извлечение ценной информации из огромных объемов данных. 4. Создание новых данных.	3 Основная задача обработки больших данных — находить полезные закономерности, инсайты и ценную информацию в огромных массивах разнородных данных. Это позволяет принимать обоснованные решения и оптимизировать бизнес-процессы
3	4 Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какая из следующих характеристик больших данных подразумевает необходимость обработки данных в режиме реального времени? 1. Объем. 2. Скорость. 3. Разнообразие. 4. Ценность.	2 характеристика "скорость" (Velocity) подразумевает необходимость обработки данных в режиме реального времени, другие варианты не подходят
5	5. Внимательно прочитайте текст задания и понять суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. В контексте больших данных, _____ анализ — это процесс использования сложных алгоритмов и статистических методов для выявления скрытых закономерностей, трендов и взаимосвязей, которые могут быть недоступны при традиционных подходах.	Прогнозный/ предиктивный

5	6. Внимательно прочитайте текст задания и понять суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. _____ вычисления — это подход, при котором задачи распределяются между несколькими процессорами или узлами, что позволяет значительно ускорить обработку больших объемов данных.		параллельные
5	7. Внимательно прочитайте текст задания и понять суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. _____ данные — это информация, которая организована в виде таблиц или баз данных с четко определенными полями, что упрощает их обработку и анализ.		Структурированные
2	8. Прочитайте текст и установите последовательность порядка Установите правильный порядок этапов обработки больших данных. 1. Сбор данных из различных источников. 2. Хранение данных в распределенных системах. 3. Очистка и предварительная обработка данных. 4. Визуализация результатов и принятие решений. 5. Анализ данных с использованием алгоритмов машинного обучения.		1,3,2,5,4
1	9. Прочитайте текст и установите соответствие Соотнесите характеристики больших данных с их определениями.		1 — С, 2 — А, 3 — В, 4 — D
	1. Объем 2. Скорость 3. Разнообразие 4. Ценность	Примерный ответ: Потоковая обработка данных — это непрерывная обработка информации в реальном времени по мере ее поступления, что позволяет мгновенно реагировать на изменения. Пакетная обработка предполагает сбор данных за определенный период и их обработку большими партиями, что подходит для задач, не требующих срочного анализа. Основное отличие — в скорости и подходе к работе с данными.	
5	10. Объясните, что такое «потоковая обработка данных» и чем она отличается от «пакетной обработки данных».		Примерный ответ: Потоковая обработка данных — это непрерывная обработка информации в реальном времени по мере ее поступления, что позволяет мгновенно реагировать на изменения. Пакетная обработка предполагает сбор данных за определенный период и их обработку большими партиями, что подходит для задач, не требующих срочного анализа. Основное отличие — в скорости и подходе к работе с данными.
3	11. Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какой из следующих инструментов чаще всего используется для обработки и анализа больших данных? 1. Microsoft Word 2. Apache Hadoop 3. Adobe Photoshop		2 Hadoop специально создан для работы с большими данными, остальные программы для этого не предназначены.

	4. Microsoft Excel		
3	12. Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какой формат данных часто используется для хранения и обмена большими объемами структурированных данных в распределенных системах обработки данных, таких как Hadoop? 1. CSV 2. JSON 3. XML 4. Avro		4 Avro — это формат, специально разработанный для хранения и обмена большими объемами структурированных данных в распределённых системах, таких как Hadoop
3	13. Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какой метод используется для снижения размера данных и улучшения производительности при обработке больших наборов данных? 1. Дедупликация 2. Сжатие данных 3. Шардирование 4. Все вышеперечисленное		4 Для снижения размера данных и повышения производительности при обработке больших данных применяют дедупликацию, сжатие и шардирование — все эти методы используются совместно для оптимизации хранения и ускорения обработки
3	14. Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какая технология параллельного программирования используется для распараллеливания задач на уровне ядер процессора в системах с общей памятью? 1. MPI (Message Passing Interface) 2. OpenMP (Open Multi-Processing) 3. oneTBB (Intel Threading Building Blocks) 4. CUDA		2 OpenMP предназначен для распараллеливания задач на уровне ядер процессора в системах с общей памятью, используя многопоточность и совместный доступ к памяти
5	15. Внимательно прочитайте текст задания и понять суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. Технология параллельного программирования для распараллеливания задач на уровне ядер процессора называется _____.		OpenMP
5	16. Внимательно прочитайте текст задания и понять суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. Механизм синхронизации потоков, предотвращающий одновременный доступ к общим ресурсам, называется _____.		Mutex
5	17. Внимательно прочитайте текст задания и понять суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. Технология, используемая для распределения задач между несколькими процессорами или узлами, называется _____.		MPI
1	18. Прочитайте текст и установите соответствие Дана задача по параллелизации вычисления суммы элементов массива с использованием нескольких ядер процессора. Установите соответствие между описанием подзадач и инструментами для их исполнения		1D, 2A, 3B, 4C
	1. Распараллеливание цикла для вычисления суммы на нескольких ядрах процессора. 2. Распределение задач между процессами в кластере.	Примерный ответ: MPI лучше подходит для распределенных задач в кластере, а OpenMP — для параллелизации на нескольких ядрах одного	

	3. Вычисления на графическом процессоре. 4. Параллелизация на уровне потоков с использованием высокоуровневых абстракций.	процессора.	
5	19 Внимательно прочитать текст задания и понять суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. Какие задачи лучше решать с помощью MPI, а какие — с помощью OpenMP?		Примерный ответ: MPI лучше подходит для распределенных задач в кластере, а OpenMP — для параллелизации на нескольких ядрах одного процессора.
3	20 Перечислите основные компоненты экосистемы Hadoop		HDFS (Hadoop Distributed File System), YARN (Yet Another Resource Negotiator), MapReduce, Hadoop Common
3	21 Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. В чем заключается ключевая особенность модели MapReduce? 1. Выполнение всех операций на одном сервере 2. Распределение данных между узлами и параллельная обработка 3. Использование SQL-запросов для анализа данных 4. Обработка только структурированных данных		2 Ключевая особенность модели MapReduce — это автоматическое распределение данных между узлами кластера и параллельная обработка больших объемов информации на разных серверах, что позволяет эффективно работать с очень большими данными
3	22 Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какой из следующих методов используется для агрегации данных на этапе Reduce? 1. Группировка по ключу 2. Сортировка данных 3. Объединение данных 4. Все вышеперечисленное		1 Сортировка и объединение — это скорее вспомогательные процессы, а сама агрегация происходит именно после группировки по ключу.
3	23 Прочитайте текст, выберите правильный вариант ответа и запишите аргументы, обосновывающие выбор ответа. Какие устройства используются для сбора больших данных в сельском хозяйстве? 1. Только спутниковые снимки. 2. Только датчики для определения влаги и состояния почвы. 3. Дроны, метеостанции, датчики для почвы и спутниковые снимки. 4. Только беспилотные транспортные средства.		3 В сельском хозяйстве для сбора больших данных применяют разные устройства: дроны для аэрофотосъемки и мониторинга посевов, спутниковые снимки для анализа состояния полей, датчики для измерения влажности и состояния почвы, а также метеостанции для сбора погодных данных
4	24 Прочитайте текст, выберите все правильные варианты ответа и запишите аргументы, обосновывающие выбор ответа. Какие из перечисленных утверждений верны для алгоритма MapReduce? 1. MapReduce позволяет обрабатывать большие объемы данных параллельно на нескольких узлах. 2. На этапе Reduce данные сортируются и группируются по ключу. 3. MapReduce требует предварительной загрузки всех данных в оперативную память для обработки. 4. MapReduce использует HDFS для хранения данных.		1,2,4,6 1. MapReduce обрабатывает большие данные параллельно на кластере 2. На этапе Reduce данные сортируются и группируются по ключу 4. MapReduce использует HDFS для хранения данных 6. Shuffle — это этап

	5.Этап Map отвечает за агрегацию данных по ключам. 6.Shuffle — это этап, где происходит передача данных от Mapper к Reducer.	передачи, сортировки и группировки данных между Mapper и Reducer
5	25 Внимательно прочитайте текст задания и поняте суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. Для оптимизации задач MapReduce в АПК часто используется метод _____ для уменьшения времени обработки данных.	параллелизация
5	26 Внимательно прочитайте текст задания и поняте суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. В АПК для обнаружения аномалий в данных о погодных условиях и урожайности используются методы _____ для анализа временных рядов.	Машинного обучения
5	27 Внимательно прочитайте текст задания и поняте суть вопроса. Продумать логику и полноту ответа. Записать ответ, используя четкие компактные формулировки. В сельскохозяйственном предприятии используются 25 датчиков для мониторинга температуры почвы. Каждый датчик собирает данные 4 раза в день. Если за месяц (30 дней) каждый датчик собирает в среднем 10 килобайт данных, сколько килобайт данных будет собрано в общей сложности за месяц?	3000
1	28 Прочитайте текст и установите соответствие Соотнесите этапы Reduce с их реализацией на коде	
	Этап	8,2,3,1,4,5,6,7
	1. Grouping 2. Aggregation 3. Output	- Суммирование значений: Используется для подсчета общей суммы значений для каждого ключа. - Поиск максимального значения: Позволяет найти наибольшее значение среди всех значений одного ключа. - Группировка по ключу: Необходима для того, чтобы данные с одинаковыми ключами обрабатывались вместе.
2	29 Прочитайте текст и установите последовательность порядка Расположите фрагменты кода в правильном порядке, чтобы реализовать параллельную обработку массива данных с помощью OpenMP. Фрагменты кода: <code>#pragma omp parallel for</code> <code>int main() {</code> <code>int data[100]; // Инициализация массива</code> <code>for (int i = 0; i < 100; i++) {</code> <code>data[i] *= 2; // Обработка данных</code> <code>return 0;</code> <code>}</code> <code>#include <omp.h></code>	8,2,3,1,4,5,6,7
4	30 Прочитайте текст, выберите все правильные варианты ответа	Суммирование

	и запишите аргументы, обосновывающие выбор ответа. Выберите подходящие методы агрегации данных на этапе Reduce и объясните свой выбор. Варианты ответов: 1.Суммирование значений 2.Поиск максимального значения 3.Группировка по ключу 4.Конкатенация строк 5.Сортировка данных	значений: Используется для подсчета общей суммы значений для каждого ключа. • Поиск максимального значения: Позволяет найти наибольшее значение среди всех значений одного ключа. • Группировка по ключу: Необходима для того, чтобы данные с одинаковыми ключами обрабатывались вместе.

ЭКСПЕРТНОЕ ЗАКЛЮЧЕНИЕ

на фонды оценочных средств по дисциплине
«Технологии обработки больших данных»
для подготовки магистров по направлению подготовки
09.04.03 «Прикладная информатика»
профиль «Цифровые технологии в АПК»

Представленные на рецензию фонды оценочных средств оформлены с соблюдением всех требований, предъявляемых к оформлению ФОС по стандартам ФГОС ВО.

Дисциплина **«Технологии обработки больших данных»** является частью учебного плана подготовки по программе магистратуры направления **09.04.03 «Прикладная информатика»** профиль «Цифровые технологии в АПК».

Оценочные средства для контроля успеваемости студентов представлены в полном объеме. При помощи фонда оценочных средств осуществляется контроль и управление процессом приобретения студентами необходимых знаний, умений, практического опыта и компетенций, определенных ФГОС ВО.

Представленные оценочные средства по дисциплине стимулируют познавательную деятельность за счет заданий разного уровня сложности, компетентностного подхода, формируют навыки само- и взаимопонимания.

Фонды оценочных средств соответствуют обязательному минимуму содержания ФГОС ВО, обеспечивают проведение аттестации студентов учреждений ВО, дают возможность определить соответствие студентов конкретной характеристике.

Представленные ФОС для подготовки по программе магистратуры направления **09.04.03 «Прикладная информатика»** профиль «Цифровые технологии в АПК» могут быть использованы в учебном процессе и соответствуют требованиям ФГОС ВО.

Эксперт:
доцент кафедры Систем автоматики,
автоматизированного управления и проектирования
ФГАОУ ВО Сибирский федеральный университет,
Институт космических и информационных технологий, канд. техн. наук, доцент



Алексей
Владимирович
Чубарь