



Федеральное государственное бюджетное образовательное учреждение высшего образования  
«ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ ДВАЖДЫ ГЕРОЯ  
СОВЕТСКОГО СОЮЗА, ЛЕТЧИКА-КОСМОНАВТА А.А. ЛЕОНОВА»

**УТВЕРЖДАЮ**

**и.о. проректора**

**А.В. Троицкий**

«\_\_\_» \_\_\_\_\_ 2023 г.

**ИНСТИТУТ ИНФОКОММУНИКАЦИОННЫХ СИСТЕМ И  
ТЕХНОЛОГИЙ**

**КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И  
УПРАВЛЯЮЩИХ СИСТЕМ**

**РАБОЧАЯ ПРОГРАММА**

**ДИСЦИПЛИНЫ**

**«АНАЛИЗ БОЛЬШИХ ДАННЫХ»**

**Направление подготовки:** 09.03.02. Информационные системы и технологии

**Профиль:** высокопроизводительные вычислительные и телекоммуникационные интеллектуальные системы и комплексы

**Уровень высшего образования:** бакалавриат

**Форма обучения:** очная

Королев 2023

Рабочая программа является составной частью основной профессиональной образовательной программы и проходит рецензирование со стороны работодателей в составе основной профессиональной образовательной программы. Рабочая программа актуализируется и корректируется ежегодно.

**Автор:** к.т.н., доц. Аббасова Т.С. Рабочая программа дисциплины: Анализ больших данных. – Королев МО: «Технологический университет», 2023.

**Рецензент:** д.т.н. профессор Артюшенко В.М.

Рабочая программа составлена в соответствии с требованиями федерального государственного образовательного стандарта высшего профессионального образования (ФГОС ВО) по направлению подготовки бакалавров 09.03.02 «Информационные системы и технологии» и Учебного плана, утвержденного Ученым советом Технологического университета.

Протокол № 9 от 11.04. 2023 г.

**Рабочая программа рассмотрена и одобрена на заседании кафедры:**

Заведующий кафедрой (ФИО, ученая степень, звание, подпись)	Артюшенко В.М., д.т.н. профессор 			
Год утверждения (переутверждения)	2023	2024	2025	2026
Номер и дата протокола заседания кафедры	№12 от 05.04 .2023			

**Рабочая программа согласована:**

Руководитель ОПОП



к.т.н., доцент Е.Г. Макарова

**Рабочая программа рекомендована на заседании УМС:**

Год утверждения (переутверждения)	2023	2024	2025	2026
Номер и дата протокола заседания УМС	№5 от 11.04.2023			

## **1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОПОП**

**Целью изучения** дисциплины «Анализ больших данных» является теоретическая и практическая подготовка студентов к работе с большими данными. Знания, полученные в результате освоения дисциплины, помогут при сборе и анализе огромных объемов структурированной или неструктурированной информации, при разработке моделей данных и получении новых знаний. Все это необходимо выпускнику, освоившему программу бакалавриата, для решения различных задач практической и научно-исследовательской деятельности.

Дисциплина направлена на частичное формирование следующих компетенций:

### **универсальные компетенции (УК):**

- УК-1. Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач;
- УК-2. Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений

**Основными задачами дисциплины** являются:

- приобретение студентами знаний о технологиях подготовки, хранения, обработки и анализа больших данных;
- применение статистических и математических методов для анализа больших объемов информации;
- статистический, инженерный, интуитивный анализ информации;
- приобретение практических навыков работы с программами анализа больших данных.

Показатель освоения компетенции отражают следующие индикаторы:

### **Трудовые действия:**

- Осуществляет поиск информации для решения поставленной задачи по различным типам запросов;
- При обработке информации отличает факты от мнений, интерпретаций, оценок, формирует собственные мнения и суждения, аргументирует свои выводы, в том числе с применением философского понятийного аппарата
- Оценивает решение поставленных задач в зоне своей ответственности в соответствии с запланированными результатами контроля, при необходимости корректирует способы решения задач

### **Необходимые умения:**

- Определяет и ранжирует информацию, требуемую для решения поставленной задачи; Определяет связи между поставленными задачами и ожидаемые результаты их решения;
- В рамках поставленных задач определяет имеющиеся ресурсы и ограничения, действующие правовые нормы

### **Необходимые знания:**

- Анализирует задачу, выделяя ее базовые составляющие;
- Анализирует пути решения проблем мировоззренческого, нравственного и личностного характера на основе использования основных философских идей и категорий в их историческом развитии и социально-культурном контексте
- Формулирует проблему, решение которой напрямую связано с достижением цели проекта;
- Анализирует план-график реализации проекта в целом и выбирает способ решения

## **2. Место дисциплины в структуре ОПОП ВО**

Дисциплина «**Анализ больших данных**» относится к дисциплинам части Б1.В, формируемой участниками образовательных отношений, основной профессиональной образовательной программы подготовки бакалавров по направлению подготовки 09.03.02 «Информационные системы и технологии».

Дисциплина реализуется кафедрой информационных технологий и управляющих систем.

Изучение данной дисциплины базируется на ранее изученных дисциплинах «Информатика», «Вычислительные сети, системы и телекоммуникации», «Технологии обработки информации» и компетенциях ОПК-2, ОПК-6, ОПК-7, ПК-4, ПК-13.

Знания и компетенции, полученные при освоении дисциплины, являются базовыми при изучении дисциплины «Автоматизированные информационные системы» и выполнении выпускной квалификационной работы бакалавра.

### 3. Объем дисциплины и виды учебной работы

Общая трудоемкость дисциплины составляет 3 зачетных единиц, 108 часов.

Таблица 1

Виды занятий	Всего часов	Семестр	Семестр	Семестр	Семестр
		пятый	шестой	седьмой	восьмой
<b>Общая трудоемкость</b>	<b>108</b>	<b>108</b>			
<b>ОЧНАЯ ФОРМА ОБУЧЕНИЯ</b>					
<b>Аудиторные занятия</b>	<b>48</b>	<b>48</b>			
Лекции (Л)	16	16			
Практические занятия (ПЗ)	32	32			
Лабораторные работы (ЛР)	-	-			
Практическая подготовка	-	-			
<b>Самостоятельная работа</b>	<b>60</b>	<b>60</b>			
Курсовые работы (проекты)	-	-			
Расчетно-графические работы	-	-			
Контрольная работа, домашнее задание	+ -	+ -			
Текущий контроль знаний	Тест	+			
Вид итогового контроля	зачет / экзамен	экзамен			
<b>ЗАОЧНАЯ ФОРМА ОБУЧЕНИЯ</b>					
<b>Аудиторные занятия</b>	<b>12</b>		<b>12</b>		
Лекции (Л)	4		4		
Практические занятия (ПЗ)	8		8		
Лабораторные работы (ЛР)	-		-		
Практическая подготовка	-		-		
Курсовые работы (проекты)	-		-		
<b>Самостоятельная работа</b>	<b>96</b>		<b>96</b>		
Контрольная работа, домашнее задание					
Вид итогового контроля	зачет / экзамен		экзамен		
<b>ОЧНО-ЗАОЧНАЯ ФОРМА ОБУЧЕНИЯ: ОТСУТСТВУЕТ</b>					

## 4. Содержание дисциплины

### 4.1. Темы дисциплины и виды занятий

Таблица 2

Наименование разделов и тем	Лекции, час очн/заоч/ очн-заочн	Лаб. занятия, час очн/заоч /очн-заочн	Занятия в интерактивной форме, час очн/заоч/ очн-заочн	Практическая подготовка очн/заоч/ очн-заочн	Код компетенций
Тема 1. Теоретические основы и источники больших данных	4/1/-	8/2/-	4/1/-	-	ПК-4 ПК-13
Тема 2. Методы статистического анализа больших данных	4/1/-	6/2/-	4/1/-	-	ПК-4 ПК-13
Тема 3. Организация хранилищ данных	4/1/-	8/2/-	4/1/-	-	ПК-4 ПК-13
Тема 4. Технологии больших данных в автоматизации	4/1/-	10/2/-	4/1/-	-	ПК-4 ПК-13
<b>Итого по первому семестру:</b>	<b>16/4/-</b>	<b>32/8/-</b>	<b>16/4/-</b>	-	

### 4.2. Содержание тем дисциплины

#### **Тема 1. Теоретические основы и источники больших данных**

Теоретические основы больших данных. Обзор источников больших данных и особенностей работы с каждым из них (открытые источники информации: статистические сборники, опубликованные отчеты и результаты исследований; доступ к закрытой информации). Технологии хранения больших данных (базы данных, хранилища данных, многомерные хранилища данных). Когнитивный анализ данных. Алгоритмы для решения задач добычи и обработки больших данных. Технология MapReduce для распараллеливания алгоритмов. Алгоритмы хэширования с учетом близости и потоковой обработки данных для тщательного анализа больших данных. Показатель PageRank. Методики сбора данных. Правовые нормы обработки больших данных.

#### **Тема 2. Методы статистического анализа больших данных**

Методы статистического анализа данных. Нахождение частых предметных наборов и кластеризация. Понижение размерности данных. Подготовка исходных данных для анализа: первичная обработка и визуализация имеющихся данных. Визуализация аналитических данных – представление информации в виде рисунков, диаграмм, с использованием

интерактивных возможностей и анимации как для получения результатов, так и для использования в качестве исходных данных для дальнейшего анализа.

### **Тема 3. Организация хранилищ данных**

Организация хранилища данных, оперативный (OLAP) и интеллектуальный (Data Mining) анализ данных. Примеры формулировок задач при использовании технологий OLAP и Data Mining. Процесс извлечения знаний в Data Mining. Data Mining и Business Intelligence. Операции интеллектуального анализа данных в Data Mining: классификация, кластеризация, регрессионный анализ, выявление ассоциаций, прогнозирование временных последовательностей, агрегирование, обнаружение отклонений. Многомерные модели для анализа данных. Куб данных. Гиперкуб данных. Многомерные хранилища данных. Реализация OLAP-функциональности средствами анализа данных в офисных приложениях.

### **Тема 4. Технологии больших данных в автоматизации**

Автоматизация бизнес-процессов и технологических процессов на базе технологий больших данных. Определение основных этапов внедрения технологий Big Data в автоматизацию технологических и бизнес-процессов. Обзор современных популярных программных средства анализа данных: Statistica, SPSS, Excel, R-Studio, Power-BI и другие. Загрузка, анализ и преобразование больших данных в различных программных средах. Набор программных сервисов Microsoft, которые работают вместе, превращая несвязанные источники данных компании в целостные интерактивные отчеты. Базы данных, файлы Excel, данные из облачных источников и интернета, текстовые файлы как источники данных. Чат-боты для обработки больших данных. Инфографика в интерактивных сервисах. Power BI для выполнения задач самостоятельной и корпоративной бизнес-аналитики. Улучшение возможностей работы с цифровыми технологиями обработки больших данных

## **5. Перечень учебно-методического обеспечения для самостоятельной работы по дисциплине**

1. Методические указания для обучающихся по освоению дисциплины.

## **6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине**

Структура фонда оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине приведена в Приложении 1.

## **7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины**

### **Основная литература:**

1. Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н. Большие данные. Big Data: Учебник для вузов / Макшанов А. В.; – Москва: изд-во «Лань». 2023. 188 с.
2. Лесковец Ю., Раджараман А., Ульман Джеффри Д. Анализ больших наборов данных / Лесковец Ю.; Пер. с англ. Слинкина А. – Москва: ДМК Пресс. 2023. 500 с.
3. Аббасова Т. С., Артюшенко В. М., Акимкина Э. Э. Повышение эффективности систем поддержки принятия решений на основе многомерных хранилищ данных: монография. – Москва, Берлин: Директ-Медиа, 2021. 128 с. Режим доступа: по подписке. – URL: [http://biblioclub.ru/index.php?page=book\\_red&id=615177](http://biblioclub.ru/index.php?page=book_red&id=615177). – Текст: электронный.

### **Дополнительная литература:**

1. Кравченко, Ю.А. Тенденции развития компьютерных технологий: учебное пособие / Ю.А. Кравченко, Э.В. Кулиев, Д.В. Заруба; Южный федеральный университет, Инженерно-технологическая академия. – Таганрог: Южный федеральный университет, 2017. – 107 с.: схем., ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=493214> (дата обращения: 29.09.2020). – Библиогр. в кн. – ISBN 978-5-9275-2360-3. – Текст: электронный.
2. Форман, Д. Много цифр. Анализ больших данных при помощи Excel / Форман Д.; Пер. с англ. Соколовой А. - Москва: Альпина Пабли., 2016. - 461 с. ISBN 978-5-9614-5032-3. – Текст: электронный. – URL: <https://znanium.com/catalog/product/551044> (дата обращения: 14.04.2022). – Режим доступа: по подписке.

## **8. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины**

1. Дорожные карты по сквозным технологиям. Режим доступа: <https://d-russia.ru/opublikovany-dorozhnye-karty-po-skvoznym-tehnologiyam-na-ih-realizatsiyu-potrebuetsya-bolshe-850-mlrd-rub.html>
2. The Big Data Value Association (BDVA) <http://www.bdva.eu>
3. Институт развития Интернета <http://ири.рф/>



## **9. Методические указания для обучающихся по освоению дисциплины**

Методические указания для обучающихся по освоению дисциплины (модуля) приведены в Приложении 2.

## **10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине**

**Перечень программного обеспечения:** Libre Office (для создания отчетов), Power BI (пробная пользовательская версия как неуправляемая служба).

**Информационные справочные системы:** не предусмотрено курсом данной дисциплины

### **Ресурсы информационно-образовательной среды Технологического университета:**

Рабочая программа и методическое обеспечение по дисциплине «Анализ больших данных».

## **11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине**

### **Лекционные занятия:**

- аудитория, оснащенная презентационной техникой (проектор, экран);
- комплект электронных презентаций / слайдов на темы:
  - 1) Задачи и методы анализа больших данных
  - 2) Архитектуры систем обработки больших данных
  - 3) Технологии статистической обработки больших данных
  - 4) Корреляционный и регрессионный анализ больших данных.
  - 5) Технологии OLAP
  - 6) Технологии Data Mining
  - 7) Технологии больших данных в автоматизации
  - 8) Принципы обработки больших данных в Power BI

### **Практические занятия:**

- учебный класс, оснащенный вычислительной техникой (ПК);
- рабочее место преподавателя, оснащенное компьютером с доступом в Интернет;
- рабочие места студентов, оснащенные компьютерами с доступом в Интернет.

**ИНСТИТУТ ИНФОКОММУНИКАЦИОННЫХ СИСТЕМ И  
ТЕХНОЛОГИЙ**

**КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ  
И УПРАВЛЯЮЩИХ СИСТЕМ**

**ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ  
ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ОБУЧАЮЩИХСЯ ПО  
ДИСЦИПЛИНЕ  
«АНАЛИЗ БОЛЬШИХ ДАННЫХ»  
(Приложение 1 к рабочей программе)**

**Направление подготовки:** 09.03.02. Информационные системы и технологии

**Профиль:** Информационные технологии в технических системах

**Уровень высшего образования:** бакалавриат

**Форма обучения:** очная, заочная

Королев 2023

## 1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

№ п/п	Индекс компетенции	Содержание компетенции (или ее части)*	Раздел дисциплины, обеспечивающий формирование компетенции (или ее части)	В результате изучения раздела дисциплины, обеспечивающего формирование компетенции (или ее части), обучающийся приобретает:		
				Трудовые действия	Необходимые умения	Необходимые знания
1	УК-1.	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	<b>Тема 1.</b>	Осуществляет поиск информации для решения поставленной задачи по различным типам запросов; При обработке информации отличает факты от мнений, интерпретаций, оценок, формирует собственные мнения и суждения, аргументирует свои выводы, в том числе с применением философского понятийного аппарата	Определяет и ранжирует информацию, требуемую для решения поставленной задачи;	Анализирует задачу, выделяя ее базовые составляющие; Анализирует пути решения проблем мировоззренческого, нравственного и личностного характера на основе использования основных философских идей и категорий в их историческом развитии и социально-культурном контексте
2	УК-2	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	<b>Тема 2. Тема 3. Тема 4.</b>	Оценивает решение поставленных задач в зоне своей ответственности и в соответствии с запланированными результатами контроля, при необходимости корректирует способы решения задач	Определяет связи между поставленными задачами и ожидаемые результаты их решения; В рамках поставленных задач определяет имеющиеся ресурсы и ограничения, действующие правовые нормы	Формулирует проблему, решение которой напрямую связано с достижением цели проекта; Анализирует план-график реализации проекта в целом и выбирает способ решения

## 2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания

Код компетенции	Инструменты, оценивающие сформированность компетенции	Показатель оценивания компетенции	Критерии оценки
УК1, УК2	Доклад в форме презентации	<p>А) полностью сформирована (компетенция освоена на высоком уровне) – 5 баллов</p> <p>Б) частично сформирована:</p> <ul style="list-style-type: none"> <li>•компетенция освоена на продвинутом уровне – 4 балла;</li> <li>•компетенция освоена на базовом уровне – 3 балла;</li> </ul> <p>В) не сформирована (компетенция не освоена) – 2 и менее баллов</p>	<p>Проводится устно с использованием мультимедийных систем, а также с использованием технических средств</p> <p>Время, отведенное на процедуру – 10 - 15 мин.</p> <p>Неявка – 0.</p> <p>Критерии оценки:</p> <ol style="list-style-type: none"> <li>1.Соответствие представленной презентации заявленной тематике (1 балл).</li> <li>2.Качество источников и их количество при подготовке доклада и разработке презентации (1 балл).</li> <li>3.Владение информацией и способность отвечать на вопросы аудитории (1 балл).</li> <li>4.Качество самой представленной презентации (1 балл).</li> <li>5.Оригинальность подхода и всестороннее раскрытие выбранной тематики (1 балл).</li> </ol> <p>Максимальная сумма баллов - 5 баллов.</p> <p>Результаты оценочной процедуры представляются обучающимся в срок не позднее 1 недели после проведения процедуры – для текущего контроля. Оценка проставляется в электронный журнал.</p>
УК1, УК2	Контрольная работа	<p>А) полностью сформирована (компетенция освоена на высоком уровне) – 5 баллов</p> <p>Б) частично сформирована:</p> <ul style="list-style-type: none"> <li>•компетенция освоена на продвинутом уровне – 4 балла;</li> <li>•компетенция освоена на базовом уровне – 3 балла;</li> </ul> <p>В) не сформирована (компетенция не освоена) – 2 и менее баллов</p>	<ol style="list-style-type: none"> <li>1. Проводится в форме письменной работы</li> <li>2.Время, отведенное на процедуру – семестр.</li> </ol> <p>Неявка на защиту контрольной работы – 0.</p> <p>Критерии оценки:</p> <ol style="list-style-type: none"> <li>1.Соответствие содержания контрольной работы заявленной тематике (1 балл).</li> <li>2.Качество источников и их количество при подготовке работы (1 балл).</li> <li>3.Владение информацией и способность отвечать на вопросы аудитории (1 балл).</li> <li>4.Качество самой представленной работы (1 балл).</li> <li>5. Использование специализированного программного обеспечения (1 балл).</li> <li>6.Оригинальность подхода и всестороннее раскрытие выбранной тематики (1 балл).</li> </ol> <p>Максимальная сумма баллов - 6 баллов.</p> <p>Результаты оценочной процедуры представляются обучающимся в срок не позднее 1 недели после проведения</p>

			процедуры – для текущего контроля. Оценка проставляется в электронный журнал.
УК1, УК2	Письменное задание	<p>А) полностью сформирована (компетенция освоена на высоком уровне) – 5 баллов</p> <p>Б) частично сформирована:</p> <ul style="list-style-type: none"> <li>•компетенция освоена на продвинутом уровне – 4 балла;</li> <li>•компетенция освоена на базовом уровне – 3 балла;</li> </ul> <p>В) не сформирована (компетенция не освоена) – 2 и менее баллов</p>	<p>1. Проводится в форме письменной работы</p> <p>2.Время, отведенное на процедуру – 10 - 15 мин.</p> <p>Неявка – 0.</p> <p>Критерии оценки:</p> <p>1.Соответствие ответа заявленной тематике (0-5 баллов).</p> <p>Максимальная сумма баллов - 5 баллов.</p> <p>Результаты оценочной процедуры представляются обучающимся в срок не позднее 1 недели после проведения процедуры – для текущего контроля. Оценка проставляется в электронный журнал.</p>
УК1, УК2	Реферат	<p>А) полностью сформирована (компетенция освоена на высоком уровне) – 5 баллов</p> <p>Б) частично сформирована:</p> <ul style="list-style-type: none"> <li>•компетенция освоена на продвинутом уровне – 4 балла;</li> <li>•компетенция освоена на базовом уровне – 3 балла;</li> </ul> <p>В) не сформирована (компетенция не освоена) – 2 и менее баллов</p>	<p>Проводится в письменной форме</p> <p>Критерии оценки:</p> <p>1.Соответствие содержания реферата заявленной тематике (1 балл).</p> <p>2.Качество источников и их количество при подготовке работы (1 балл).</p> <p>3.Владение информацией и способность отвечать на вопросы аудитории (1 балл).</p> <p>4.Качество самой представленной работы (1 балл).</p> <p>5.Оригинальность подхода и всестороннее раскрытие выбранной тематики (1 балл).</p> <p>Максимальная сумма баллов - 5 баллов.</p> <p>Результаты оценочной процедуры представляются обучающимся в срок не позднее 1 недели после проведения процедуры – для текущего контроля. Оценка проставляется в электронный журнал.</p>

### **3. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы**

#### **Тематика докладов в презентационной форме**

1. Дополненная аналитика: практический подход к автоматизации внутренней аналитики
2. Формулировки задач при использовании технологий Data Mining.

3. Построение модели и кластеризация данных с неожиданными характеристиками.
4. Цифровой помощник для маневрового диспетчера.
5. Предсказание цен на сырье и конечные продукты на основе машинного обучения и продвинутой аналитики
6. Способы добычи больших данных.
7. Технологии обработки больших данных.
8. Статистический анализ больших данных.
9. Методы анализа неструктурированных данных текстов.
10. Методы анализа неструктурированных данных сетей.
11. Методы анализа неструктурированных данных изображений.
12. Имитационное моделирование сложных социальных систем.
13. Проблемы машинного обучения.
14. Предиктивные модели.
15. Платформа геоаналитики.
16. Big data на службе семейного бюджета.
17. Скоринговые модели.
18. Учет новых слов в языковых моделях классификации.
19. «Прикладные» нейросети в браузере и на десктопе клиента.
20. Разработка скрининговых тестов.
21. Технологии анализа финансовых потоков и прогнозов.
22. Внедрение управления на основе данных в нефтегазовой отрасли.
23. Искусственный интеллект для снижения вредных выбросов в атмосферу на производстве.
24. Построение системы realtime-аналитики.
25. Границы применимости при создании аналитических решений.

### **Тематика рефератов**

1. Средства анализа данных в офисных приложениях с помощью OLAP-технологий.
2. Методы Data Mining для обучения ассоциативным правилам (англ. association rule learning).
3. Методы классификации в Data Mining.
4. Методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным в Data Mining.
5. Кластерный анализ в Data Mining.
6. Регрессионный анализ в Data Mining.
7. Техники смещения и интеграции разнородных данных для возможности глубинного анализа.
8. Цифровая обработка сигналов для интеграции разнородных данных.
9. Обработка естественного языка для интеграции разнородных данных.
10. Машинное обучение с учителем и без учителя,

11. Использование моделей, построенных на базе статистического анализа или машинного обучения, для получения комплексных прогнозов на основе базовых моделей.
12. Искусственные нейронные сети для анализа больших данных.
13. Прогнозная аналитика на космическом производстве.
14. Имитационное моделирование больших данных.
15. Пространственный анализ с использованием топологической информации в данных.
16. Пространственный анализ с использованием геометрической информации в данных.
17. Пространственный анализ с использованием географической информации в данных.
18. Статистический анализ больших данных с помощью A/B-тестирования.
19. Статистический анализ больших данных с помощью временных рядов.
20. Адаптивные и мультипликативные методы прогнозирования.
21. Дисперсионный анализ влияния качественных факторов.
22. Применение программного обеспечения для решения прикладных задач прогнозирования.
23. Большие данные в информационных технологиях.
24. Подходы, инструменты и методы обработки структурированных данных огромных объёмов.
25. Подходы, инструменты и методы обработки неструктурированных данных огромных объёмов.

### **Тематика письменных заданий**

1. Перечислите методы многомерного анализа больших данных.
2. Опишите принципы факторного анализа.
3. Опишите принципы кластерного анализа.
4. Опишите принципы таксономии.
5. Опишите принципы распознавания образов.
6. Перечислите достоинства факторного анализа.
7. Перечислите достоинства кластерного анализа.
8. Перечислите достоинства таксономии.
9. Перечислите достоинства распознавания образов.
10. Перечислите недостатки факторного анализа.
11. Перечислите недостатки кластерного анализа.
12. Перечислите недостатки таксономии.
13. Перечислите недостатки распознавания образов.
14. Сформулируйте задачу кластеризации.
15. Опишите алгоритм последовательной кластеризации.
16. Опишите метод представления матрицы расстояний с помощью

дендограммы.

17. Дайте характеристику Big Data на мировом рынке.
18. Охарактеризуйте Big Data в России.
19. Определите понятие Data Mining.
20. Определите понятие OLAP.
21. В чем состоит когнитивный анализ данных?
22. Какие модели данных вы знаете?
23. Опишите основную идею корреляционного анализа.
24. Опишите основную идею регрессионного анализа.
25. Опишите основную идею дисперсионного анализа.

#### **4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций**

Формой контроля знаний по дисциплине «Анализ больших данных» являются две текущие аттестации в виде тестов и итоговая аттестация в форме экзамена.

Неделя текущего контроля	Вид оценочного средства	Код компетенций, оценивающий знания, умения, навыки	Содержание оценочного средства	Требования к выполнению	Срок сдачи (неделя семестра)	Критерии оценки по содержанию и качеству с указанием баллов
Согласно графика учебного процесса	Тестирование 1,2	УК1, УК2	25 вопросов	Компьютерное тестирование ; время отведенное на процедуру - 30 минут	Результаты тестирования предоставляются в день проведения процедуры	Критерии оценки определяются процентным соотношением. Не явка -0 Удовлетворительно - от 51% правильных ответов. Хорошо - от 70%. Отлично – от 90%.
Согласно графика учебного процесса	экзамен	УК1, УК2	2 вопроса и 1 практическое задание	Экзамен проводится в письменной форме, путем ответа на вопросы и решения практического задания. время отведенное	Результаты предоставляются в день проведения экзамена	Критерии оценки: <b>«Отлично»:</b> <ul style="list-style-type: none"> <li>• знание основных понятий предмета;</li> <li>• умение использовать и применять полученные знания на практике;</li> <li>• работа на практических занятиях;</li> <li>• знание основных научных теорий, изучаемых предметов;</li> </ul>



				на процедуру – 0,35 часа на студента.		<ul style="list-style-type: none"> <li>• ответ на вопросы билета.</li> <li>«Хорошо»:</li> <li>• знание основных понятий предмета;</li> <li>• умение использовать и применять полученные знания на практике;</li> <li>• работа на практических занятиях;</li> <li>• знание основных научных теорий, изучаемых предметов;</li> <li>• ответы на вопросы билета</li> <li>• неправильно решено практическое задание</li> <li>«Удовлетворительно»:</li> <li>• демонстрирует частичные знания по темам дисциплин;</li> <li>• незнание неумение использовать и применять полученные знания на практике;</li> <li>• не работал на практических занятиях;</li> <li>«Неудовлетворительно»:</li> <li>• демонстрирует частичные знания по темам дисциплин;</li> <li>• незнание основных понятий предмета;</li> <li>• неумение использовать и применять полученные знания на практике;</li> <li>• не работал на практических занятиях;</li> <li>• не отвечает на вопросы.</li> </ul>
--	--	--	--	--	--	--

\*Итоговое начисление баллов по дисциплине осуществляется в соответствии с разработанной и внедренной балльно-рейтинговой системой контроля и оценивания уровня знаний и внеучебной созидательной активности обучающихся, согласно приказу «О внедрении новой балльно-рейтинговой системы контроля и оценивания уровня знаний и внеучебной

#### **4.1. Типовые вопросы, выносимые на тестирование**

Тесты используются в режиме промежуточного контроля. По форме заданий выбраны закрытые тесты (с выборочным ответом). Каждому вопросу соответствует один вариант ответа.

##### **Тест 1**

Тесты используются в режиме промежуточного контроля. По форме заданий выбраны закрытые тесты (с выборочным ответом). Каждому вопросу соответствует один вариант ответа.

- 1. Объём накопленных человечеством цифровых данных измеряется:**
- 2. Укажите фактор, способствовавший появлению тренда больших данных:**
- 3. Сколько Петабайт в Зеттабайте?**
- 4. Отметьте значимое событие, повлиявшее на формирование тренда больших данных.**
- 5. Отметьте вариант, в котором данные структурированы.**
- 6. Что относится к персональным данным, на обработку которых требуется получить согласие?**
- 7. Число, которое является серединой множества чисел, то есть половина чисел имеют значения большие, чем это число, а половина чисел имеют значения меньшие, чем это число, называется...**
- 8. Количественное значение исследуемого признака, наиболее часто встречающееся в выборке, называется...**
- 9. Значение, которое как статистический показатель представляет собой среднюю оценку изучаемого в эксперименте значения, называется...**
- 10. Среднее арифметическое квадратов отклонений значений переменной от её среднего значения.**
- 11. Как называется хранилище, в котором данные размещаются и сохраняются на многочисленных распределенных в сети серверах, предоставляемых в пользование клиентам?**
- 12. Аналитическая панель с понятным интерфейсом, на которой в режиме реального времени на экране отображаются необходимые данные – это...**
- 13. Какие риски из перечисленных характерны для работы с данными?**
- 14. Что обязан сделать оператор, если персональные данные обрабатываются в целях продвижения товаров, работ, услуг на рынке путем осуществления прямых контактов с потребителем или же в целях политической агитации?**

15. Какие технологии используют для сбора данных?
16. Раздел математической статистики, изучающий статистические методы, основанные на последовательной выборке, формируемой в ходе статистического эксперимента – это...
17. Комплекс, состоящий из методов анализа данных и способов их интерпретации, позволяющий принимать успешные решения в будущем на основе результатов прошлых событий – это...
18. Привлечение к решению тех или иных проблем инновационной производственной деятельности широкого круга лиц для использования их творческих способностей, знаний и опыта по типу субподрядной работы на добровольных началах с применением информационных технологий – это...
19. Процесс, используемый компаниями для превращения необработанных больших данных в полезную информацию – это...
20. С чего надо начинать работу с большими данными?

## 2 тестирование

1. Какие технологии используют для передачи и хранения данных?
2. Какие технологии используют для сбора данных?
3. Что НЕ относится к сквозным цифровым технологиям?
4. К какой сквозной цифровой технологии относятся программно-определяемые хранилища данных?
5. Специализированное здание или помещение, в котором компания размещает серверное и сетевое оборудование с последующим подключением клиентов к сети Интернет – это...
6. Централизованное место в архитектуре облака, включающие огромные объемы необработанных данных в их собственном формате – это...
7. Централизованное место в архитектуре облака, включающие огромные объемы необработанных данных в их собственном формате, НЕ доступное пользователям, – это...
8. Предметно-ориентированная информационная база данных, специально разработанная и предназначенная для подготовки отчетов и бизнес-анализа с целью поддержки принятия решений в организации, – это...
9. Представление данных с различными измерениями и фактами реализует
10. Неструктурированные данные чаще всего хранятся в...
11. Для чего нужны озера данных?
12. Для чего нужны хранилища данных?
13. Для чего нужны центры облачные хранилища данных?
14. Для чего нужны центры обработки данных?
15. Какой из поставщиков облачных услуг является отечественным?

**16. Комплекс методов и процедур, направленных на извлечение данных из различных источников, обеспечение необходимого уровня их информативности и качества, преобразование в единый формат, в котором они могут быть загружены в хранилище данных или аналитическую систему.**

**17. Процесс выявления и исправления ошибок, несоответствий данных с целью улучшения их качества, составная часть интеллектуального анализа данных.**

**18. Процесс дополнения сырых данных той информацией, которая в исходном виде в них отсутствует, но необходима для качественного анализа.**

**19. Совокупность алгоритмов и компьютерных методов обработки информации, используемых для выборки данных из специальных хранилищ или баз знаний.**

**20. Сбор информации из баз данных с целью подготовки комбинированных наборов данных для обработки данных.**

#### **4.2. Типовые вопросы, выносимые на зачет**

1. Источники больших данных.
2. Основные определения, термины, задачи анализа больших данных.
3. Понятие Data Mining.
4. Когнитивный анализ данных.
5. Методики сбора данных.
6. Алгоритмы для решения задач добычи данных и очень больших данных.
7. Распараллеливание алгоритмов при обработке больших данных
8. Задачи технологии MapReduce.
9. Алгоритмы хэширования с учетом близости и потоковой обработки данных для тщательного анализа больших данных.
10. Показатель PageRank при обработке больших данных.
11. Принципы обработки больших данных в базах данных.
12. Принципы обработки больших данных в хранилищах данных.
13. Принципы обработки больших данных в многомерных хранилищах данных.
14. Первичная обработка больших данных.
15. Методы статистического анализа данных. Дескриптивная статистика, параметрические, непараметрические, номинальные методы.
16. Корреляционный анализ.
17. Регрессионный анализ.
18. Дисперсионный анализ.
19. Кластерный анализ.

20. Дискриминантный анализ.
21. Факторный анализ.
22. Метод анализа больших данных drill down/up.
23. Классификация данных. Методы решения задачи классификации данных. Примеры классификации данных.
24. Кластеризация данных. Методы решения задачи кластеризации данных. Примеры кластеризации данных.
25. Методы прогнозирования данных при обработке информации.
26. Статистика и отчеты. Визуализация справочной информации. Программные и аппаратные средства поддержки.
27. Визуализация аналитических данных – представление информации в виде рисунков, диаграмм, с использованием интерактивных возможностей и анимации. Программные средства поддержки.
28. Инфографика в интерактивных сервисах. Программные средства поддержки.
29. Улучшение возможностей работы с цифровыми технологиями обработки больших данных.
30. Power BI для выполнения задач самостоятельной и корпоративной бизнес-аналитики.

**ИНСТИТУТ ИНФОКОММУНИКАЦИОННЫХ СИСТЕМ  
И ТЕХНОЛОГИЙ**

**КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ  
И УПРАВЛЯЮЩИХ СИСТЕМ**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО  
ОСВОЕНИЮ ДИСЦИПЛИНЫ  
«АНАЛИЗ БОЛЬШИХ ДАННЫХ»  
(Приложение 2 к рабочей программе)**

**Направление подготовки:** 09.03.02. Информационные системы и технологии

**Профиль:** Информационные технологии в технических системах

**Уровень высшего образования:** бакалавриат

**Форма обучения:** очная, заочная

Королев 2023

## 1. Общие положения

**Цель дисциплины:** теоретическая и практическая подготовка студентов к работе с большими данными. Знания, полученные в результате освоения дисциплины, помогут при сборе и анализе огромных объемов структурированной или неструктурированной информации, при разработке моделей данных и получении новых знаний. Все это необходимо выпускнику, освоившему программу бакалавриата, для решения различных задач практической и научно-исследовательской деятельности.

**Задачи дисциплины:**

- приобретение студентами знаний о технологиях подготовки, хранения, обработки и анализа больших данных;
- применение статистических и математических методов для анализа больших объемов информации;
- статистический, инженерный, интуитивный анализ информации;
- приобретение практических навыков работы с программами анализа больших данных.

## 2. Указания по проведению практических занятий

### Тема 1. Теоретические основы и источники больших данных

#### Практическое занятие 1

##### **Быстрый поиск нормативных документов в системе «Консультант Плюс»**

**Вид практического занятия:** практическая работа в группе

**Образовательные технологии:** технологии компьютерного обучения.

Работа с источником больших данных – системой КонсультантПлюс (некоммерческая версия <https://www.consultant.ru/online/>). Настройка системы КонсультантПлюс под задачи специалистов. Готовые решения и аналитические материалы КонсультантПлюс. Подборка наиболее важных документов по сквозной цифровой технологии «Большие данные» (нормативно-правовых актов, статей консультаций экспертов).

Продолжительность практического занятия 2/0,25/- часа.

#### Практическое занятие 2

##### **Национальная информационно-аналитическая система в электронной библиотеке eLibrary.ru**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

Электронная библиотека как источник больших данных. Исследование структуры и навигации информационных ресурсов электронной библиотеки eLibrary.ru, анализ показателей российского индекса научного цитирования.

Публикация данных о статьях в соответствии с требованиями ГОСТ к оформлению.

Продолжительность практического занятия 2/0,25/- часа.

### **Практическое занятие 3**

#### **Анализ больших данных метеорологических станций**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

**Источник больших данных о погоде.** Форматы данных о погоде. Объем выборки данных о погоде. Инструмента «Статистика погоды». Математические модели и методы для расчета статистических показателей Автоматический расчет показателей погоды. Определение резко выделяющихся данных по критерию Смирнова-Граббса.

Продолжительность практического занятия 2/0,25/- часа.

### **Практическое занятие 4**

#### **Параллельные вычисления**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

**Отличия параллельных вычислений от распределённых.** Выявление (указание) операций или множеств операций последовательного алгоритма, которые могут выполняться одновременно. Статическое и динамическое распараллеливание. Способы параллелизации для решения задач.

Продолжительность практического занятия 2/0,25/- часа.

## **Тема 2. Методы статистического анализа больших данных**

### **Практическое занятие 5**

#### **Интернет-сервисы для анализа и визуализации данных**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

**Выбор интернет-сервиса для обработки данных (Гугл-диск, Яндекс-диск и др.).** Выбор и настройка визуальных элементов в зависимости от цели визуализации и вида данных. Типы графиков и возможности форматирования. Принципы выбора диаграмм. Основные ошибки при визуализации данных. Условное форматирование и форматирование таблиц. Настройка взаимодействия между визуальными элементами. Настройка фильтров и управляющих элементов.

Продолжительность практического занятия 2/0,25/- часа.

### **Практическое занятие 6**

#### **Статистический анализ больших данных**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.



Удобная универсальная вычислительная среда для решения задач обработки экспериментальных данных, а также данных из различных источников. Показатели математической статистики. Обработка статистических данных с помощью встроенных функций табличного процессора. Работа с источником больших данных.

Продолжительность практического занятия 2/0,25/- часа.

### **Практическое занятие 7**

#### **Построение сводных таблиц для анализа больших данных**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

Анализ данных в сводных таблицах с помощью табличного процессора. Динамический итог данных, содержащихся в базе данных. Мастер сводных таблиц. Вставка сводной таблицы в MS Excel. Вставка сводной таблицы в ОнлиОфис. Вставка сводной таблицы в МойОфис. Работа с источником больших данных.

Продолжительность практического занятия 2/0,25/- часа.

## **Тема 3. Организация хранилищ данных**

### **Практическое занятие 8**

#### **Кластерный анализ данных**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

Кластеризация как логическое продолжение идеи классификации. Изначально не predetermined классы объектов. Разбиение объектов на группы. Проведение кластерного анализа методом «ближайшего соседа» и определение расстояния между полученными кластерами в табличном процессоре. Использование теоремы Пифагора для вычисления евклидова расстояния на плоскости.

Продолжительность практического занятия 2/0,25/- часа.

### **Практическое занятие 9**

#### **Регрессионный и корреляционный анализ данных**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

Методы регрессионного и корреляционного анализа и результаты объединения этих методов с помощью табличного процессора. Прогнозирование, планирование, развитие приоритетных направлений, принятие управленческих решений.

Продолжительность практического занятия 2/0,25/- часа.

## **Практическое занятие 10**

### **Прогнозирование временных последовательностей**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

Методы прогнозирования временных рядов, включая методы экстраполяции, методы машинного обучения (регрессия, нейронные сети). Примеры использования временных рядов. Прогнозирование изменения погоды с помощью Python и Excel. Прогнозирование температуры, осадков и других факторов, которые могут повлиять на жизнь людей и на экономику.

Продолжительность практического занятия 2/0,25/- часа.

## **Практическое занятие 11**

### **Реализация OLAP-функциональности средствами анализа данных в офисных приложениях**

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии компьютерного обучения.

Многомерное представление данных. Многомерная обработка. Многомерное хранение. OLAP-средства, встроенные в настольные приложения. Операции над многомерными данными.

Продолжительность практического занятия 2/0,25/-часа.

## **Тема 4. Технологии больших данных в автоматизации**

## **Практическое занятие 12**

### **Создание чат-бота в Telegram**

**Вид практического занятия:** кейс.

**Образовательные технологии:** технологии проектной деятельности.

Разработка чат-бота в Telegram с помощью бессерверных вычислений. Регистрация приложения в Telegram. Создание сервисного аккаунта. Создание функции и связь ее с чат-ботом. Обучающиеся разбиваются на группы по 3 человека, каждая группа реализует свою функцию, которую должен автоматизировать чат-бот. Предлагаются возможные решения и выбирается лучшее из них.

Продолжительность практического занятия 4/0,5/- часа.

## **Практическое занятие 13**

### **Изучение интерфейса и возможностей Power BI**

**Вид практического занятия:** компьютерное моделирование.

**Образовательные технологии:** технологии компьютерного обучения.

Регистрация в Power BI и системные требования. Знакомство с инструментами Power BI и системой четырех этапов решения аналитических задач. Работа с Power BI Service. Фильтры и интерактивность. Фокусировка, сортировка, подсказки.

Продолжительность практического занятия 2/0,25/-часа.

## Практическое занятие 14

### Загрузка и преобразование больших данных в Power BI

**Вид практического занятия:** компьютерное моделирование.

**Образовательные технологии:** технологии проблемного обучения.

Подготовка исходных данных к дальнейшему анализу. Работа со сложными запросами и функциями. Экспорт данных. Создание отчетов, копий и их кастомизация. Подготовка данных к работе. Импорт данных из файлов Calc, CSV. Обработка, преобразования и очистка данных.

Продолжительность практического занятия 2/0,25/- часа.

## Практическое занятие 15

### Построение идентифицирующих кодов при обработке больших данных

**Вид практического занятия:** практическая работа в группах.

**Образовательные технологии:** технологии проблемного обучения.

Технология штрихового кодирования для автоматической идентификации и сбора данных. Идентификация объекта после присвоения ему цифрового, буквенного или буквенно-цифрового кода. Представление двоичного кода в виде штрихов и пробелов с использованием определённой символики. Применение онлайн-генератора штрих-кодов.

Продолжительность практического занятия 2/0,25/- часа.

## 3. Указания по проведению лабораторного практикума

Не предусмотрен учебным планом.

## 4. Указания по проведению самостоятельной работы студентов

№ п/п	Наименование блока (раздела) дисциплины	Виды СРС
1.	Тема 1. Ведение в анализ больших данных	1. Способ представления больших данных с помощью диаграммы разброса 2. Способ представления больших данных с помощью лепестковой диаграммы 3. Способ представления больших данных с помощью тепловой диаграммы 4. Способ представления больших данных с помощью облака тегов
2.	Тема 2.	1. Корреляционный анализ.

	Методы анализа и визуализации больших данных	<ol style="list-style-type: none"> <li>2. Регрессионный анализ.</li> <li>3. Дисперсионный анализ.</li> <li>4. Кластерный анализ.</li> <li>5. Дискриминантный анализ.</li> <li>6. Факторный анализ.</li> </ol>
3	Тема 3. Технологии добычи, хранения и обработки больших данных	<ol style="list-style-type: none"> <li>1. Обработка полевых сейсморазведочных данных в модификациях 2D.</li> <li>2. Обработка полевых сейсморазведочных данных в модификациях 3D.</li> <li>3. Обработка данных в системах дистанционного зондирования земли.</li> </ol>
4	Тема 4. Программные сервисы для анализа больших данных	<ol style="list-style-type: none"> <li>1. Средства информационного описки в сети Интернет</li> <li>2. Информационно-поисковые каталоги</li> <li>3. Мета-поисковые системы</li> <li>4. Оценка показателей эффективности поисковых систем</li> </ol>

## 5. Указания по проведению контрольных работ

### 5.1. Требования к структуре

Структура контрольной работы должна способствовать раскрытию темы: иметь титульный лист, содержание, введение, основную часть, заключение, список литературы.

### 5.2. Требования к содержанию

5.2.1. Во введении обосновывается актуальность темы, определяется цель работы, задачи и методы исследования.

5.2.2. При определении целей и задач исследования необходимо правильно их формулировать. Так, в качестве цели не следует употреблять глагол «сделать». Правильно будет использовать глаголы: «раскрыть», «определить», «установить», «показать», «выявить» и т.д.

5.2.3. Основная часть работы включает вопросы, каждый из которых посвящается решению задач, сформулированных во введении, и заканчивается констатацией итогов.

5.2.4. Приветствуется иллюстрация содержания работы таблицами, графическим материалом (рисунками, схемами и т.п.).

5.2.5. Необходимо давать ссылки на используемую Вами литературу.

5.2.6. Заключение должно содержать сделанные автором работы выводы, итоги исследования.

5.2.7. Вслед за заключением идет список литературы, который должен быть составлен в соответствии с установленными требованиями. Если в работе имеются приложения, они оформляются на отдельных листах, и должны быть соответственно пронумерованы.

### **5.3. Требования к оформлению.**

Объем контрольной работы – 5...10 страниц формата А4, напечатанного с одной стороны текста (1,5 интервал, шрифт Times New Roman). Контрольная работа должна быть также представлена в электронном виде.

## **Тематика контрольных работ**

1. Использование специализированных расчетных пакетов для анализа и обработки многомерных данных.
2. Использование технологий OLAP для анализа и обработки больших данных.
3. Использование технологий Data Mining для анализа и обработки больших данных.
4. Использование CALS-технологий для анализа и обработки больших данных.
5. Использование программной среды Calc для анализа и обработки больших данных.
6. Использование программной среды Power BI для анализа и обработки больших данных.
7. Создание отчета и визуализации в программной среде Power BI.
8. Создание информационной панели в программной среде Power BI.
9. Подготовка и анализ дашборда в программной среде Power BI.
10. Создание фильтров в программной среде Power BI.
11. Фиксирование больших данных в исследуемой области.
12. Выбор инструментов для сбора и хранения данных.
13. Исследование источников больших данных.
14. Исследование способов добычи больших данных.
15. Показатели и мониторинг качества больших данных.

## **6. Перечень основной и дополнительной учебной литературы**

### **Основная литература:**

1. Макшанов А. В., Журавлев А. Е., Тындыкарь Л. Н. Большие данные. Big Data: Учебник для вузов / Макшанов А. В.; – Москва: изд-во «Лань». 2023. 188 с.

2. Лесковец Ю., Раджараман А., Ульман Джеффри Д. Анализ больших наборов данных / Лесковец Ю.; Пер. с англ. Слинкина А. – Москва: ДМК Пресс. 2023. 500 с.

3. Аббасова Т. С., Артюшенко В. М., Акимкина Э. Э. Повышение эффективности систем поддержки принятия решений на основе многомерных хранилищ данных: монография. – Москва, Берлин: Директ-Медиа, 2021. 128 с. Режим доступа: по подписке. – URL: [http://biblioclub.ru/index.php?page=book\\_red&id=615177](http://biblioclub.ru/index.php?page=book_red&id=615177). – Текст: электронный.

#### **Дополнительная литература:**

1. Кравченко, Ю.А. Тенденции развития компьютерных технологий: учебное пособие / Ю.А. Кравченко, Э.В. Кулиев, Д.В. Заруба; Южный федеральный университет, Инженерно-технологическая академия. – Таганрог: Южный федеральный университет, 2017. – 107 с.: схем., ил. – Режим доступа: по подписке. – URL: <http://biblioclub.ru/index.php?page=book&id=493214> (дата обращения: 29.09.2020). – Библиогр. в кн. – ISBN 978-5-9275-2360-3. – Текст: электронный.

2. Форман, Д. Много цифр. Анализ больших данных при помощи Excel / Форман Д.; Пер. с англ. Соколовой А. - Москва: Альпина Пабли., 2016. - 461 с. ISBN 978-5-9614-5032-3. – Текст: электронный. – URL: <https://znanium.com/catalog/product/551044> (дата обращения: 14.04.2022). – Режим доступа: по подписке.

### **7. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

1. Дорожные карты по сквозным технологиям. Режим доступа: <https://d-russia.ru/opublikovany-dorozhnye-karty-po-skvoznym-tehnologiyam-na-ih-realizatsiyu-potrebuetsya-bolshe-850-mlrd-rub.html>
2. The Big Data Value Association (BDVA) <http://www.bdva.eu>
3. Институт развития Интернета <http://ири.рф/>

### **8. Перечень информационных технологий**

**Перечень программного обеспечения:** LibreOffice (для создания отчетов), Power BI (пробная пользовательская версия как неуправляемая служба).

**Информационные справочные системы:** не предусмотрено курсом данной дисциплины

**Ресурсы информационно-образовательной среды Технологического университета**

Рабочая программа и методическое обеспечение по дисциплине «Анализ больших данных».