

Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение высшего образования
«ИРКУТСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Структурное подразделение «Институт информационных технологий и анализа данных»

УТВЕРЖДЕНА:

на заседании Совета института ИТиАД им. Е.И.Попова

Протокол №8 от 24 февраля 2025 г.

Рабочая программа дисциплины

«ИНЖЕНЕРИЯ ДАННЫХ»

Направление: 09.04.02 Информационные системы и технологии

Цифровизация промышленных предприятий

Квалификация: Магистр

Форма обучения: очная

Документ подписан простой
электронной подписью
Составитель программы:
Кононенко Роман
Владимирович
Дата подписания: 17.06.2025

Документ подписан простой
электронной подписью
Утвердил: Говорков Алексей
Сергеевич
Дата подписания: 18.06.2025

Документ подписан простой
электронной подписью
Согласовал: Кононенко Роман
Владимирович
Дата подписания: 17.06.2025

Год набора – 2025

Иркутск, 2025 г.

1 Перечень планируемых результатов обучения по дисциплине, соотнесённых с планируемыми результатами освоения образовательной программы

1.1 Дисциплина «Инженерия данных» обеспечивает формирование следующих компетенций с учётом индикаторов их достижения

Код, наименование компетенции	Код индикатора компетенции
ПК-1 Способен вести разработку и проектирование информационных систем по требованиям заказчика, в том числе разрабатывать функциональные прототипы информационных систем	ПК-1.2

1.2 В результате освоения дисциплины у обучающихся должны быть сформированы

Код индикатора	Содержание индикатора	Результат обучения
ПК-1.2	Способен разрабатывать архитектуры информационных систем, структуры баз данных информационных систем в соответствии с особенностями архитектурной спецификации промышленного интернета вещей (IIoT)	<p>Знать 1. основы архитектуры данных: понимание принципов проектирования и структурирования данных в рамках различных архитектур хранения, таких как oltp, olap, и хранилищ данных. 2. типы и структуры данных: знание различных типов данных (реляционные, документные, графовые и пр.) и их характеристик, а также форматов хранения данных (json, xml, parquet и т.д.). 3. etl и elt процессы: понимание процессов извлечения, трансформации и загрузки данных (etl) и их трансформации на этапе загрузки (elt). 4. языки и инструменты для обработки данных: знание языков программирования и технологий для обработки данных, таких как sql, python, r, apache spark и других.</p> <p>Уметь 1. проектировать архитектуры для хранения данных: уметь разрабатывать модели данных и архитектуры для различных сценариев, учитывая требования к производительности и масштабированию. 2. реализовывать etl-процессы: уметь разрабатывать и вести процессы извлечения, трансформации и загрузки данных, обеспечивая их высокое качество. 3. работать с реляционными и нереляционными субд: уметь эффективно управлять, извлекать и</p>

		<p>анализировать данные как из реляционных, так и из нереляционных систем. 4. обрабатывать и анализировать данные: уметь использовать инструменты и языки программирования для анализа и визуализации данных, выявления инсайтов и формирования отчетов. 5. оценивать качество данных и проводить чистку: уметь тестировать и очищать данные, выявляя ошибки и аномалии, а также проводить процедуры по повышению их качества.</p> <p>Владеть 1. инструментами для обработки больших данных: владеть современными инструментами для работы с большими данными, такими как apache hadoop, spark, kafka и другими. 2. технологиями хранения данных в облаке: владеть знаниями о работе с облачными платформами (aws, google cloud, azure) и их сервисами для хранения и обработки данных. 3. инструментами для визуализации данных: владеть инструментами для визуализации данных, такими как tableau, power bi, и средствами программирования, такими как matplotlib и seaborn. 4. методами машинного обучения: владеть основами машинного обучения и уметь применять технологии для извлечения информации из данных. 5. методологиями devops: владеть основами devops в контексте управления данными, включая автоматизацию процессов разработки и внедрения решений по работе с данными.</p>
--	--	--

2 Место дисциплины в структуре ООП

Изучение дисциплины «Инженерия данных» базируется на результатах освоения следующих дисциплин/практик: «Прикладное применение машинного зрения»

Дисциплина является предшествующей для дисциплин/практик: «Машинное обучение»

3 Объем дисциплины

Объем дисциплины составляет – 3 ЗЕТ

Вид учебной работы	Трудоемкость в академических часах (Один академический час соответствует 45 минутам астрономического часа)	
	Всего	Семестр № 3
Общая трудоемкость дисциплины	108	108
Аудиторные занятия, в том числе:	28	28
лекции	14	14
лабораторные работы	0	0
практические/семинарские занятия	14	14
Самостоятельная работа (в т.ч. курсовое проектирование)	44	44
Трудоемкость промежуточной аттестации	36	36
Вид промежуточной аттестации (итогового контроля по дисциплине)	Экзамен	Экзамен

4 Структура и содержание дисциплины

4.1 Сводные данные по содержанию дисциплины

Семестр № 3

№ п/п	Наименование раздела и темы дисциплины	Виды контактной работы						СРС		Форма текущего контроля
		Лекции		ЛР		ПЗ(СЕМ)		№	Кол. Час.	
		№	Кол. Час.	№	Кол. Час.	№	Кол. Час.			
1	2	3	4	5	6	7	8	9	10	11
1	Введение в инженериию данных	1	2							Тест
2	Архитектуры хранения данных	2	2			1	7			Тест
3	ETL и ELT процессы	3	2			2	7			Тест
4	Моделирование данных	4	2							Тест
5	Работа с большими данными	5	2							Тест
6	Качество данных и их обработка	6	2							Тест
7	Инструменты визуализации данных и аналитика	7	2					1	44	Тест
	Промежуточная аттестация								36	Экзамен
	Всего		14				14		80	

4.2 Краткое содержание разделов и тем занятий

Семестр № 3

№	Тема	Краткое содержание
---	------	--------------------

1	Введение в инженерию данных	Обзор понятия инженерии данных, ее основных целей и задач. Рассмотрение ключевых компонентов, таких как сбор, обработка и хранение данных. Обсуждение роли инженеров данных в современных организациях и их взаимодействия с другими участниками разработки.
2	Архитектуры хранения данных	Изучение различных архитектур систем хранения данных, включая реляционные и нереляционные базы данных, хранилища данных и решения Big Data. Рассмотрение принципов проектирования и выбора технологий в зависимости от бизнес-требований.
3	ETL и ELT процессы	Подробное изучение методов извлечения, трансформации и загрузки (ETL) данных, а также их альтернативы, такие как ELT. Практические примеры реализации ETL-процессов, включая инструменты и технологии для их автоматизации.
4	Моделирование данных	Основы концептуального, логического и физического моделирования данных. Рассмотрение различных методологий моделирования, таких как ER-диаграммы и нормализация. Обсуждение принципов документирования моделей данных.
5	Работа с большими данными	Введение в концепцию больших данных, объемы данных и технологии, которые позволяют работать с ними. Изучение фреймворков, таких как Apache Hadoop и Apache Spark, их архитектур и применения в обработке больших данных.
6	Качество данных и их обработка	Основы управления качеством данных. Изучение методов проверки, очистки и трансформации данных для обеспечения их качества. Рассмотрение инструментов и подходов для мониторинга и улучшения качества данных.
7	Инструменты визуализации данных и аналитика	Инструменты визуализации данных и аналитика- Изучение методов визуализации данных и аналитики для извлечения бизнес-инсайтов. Рассмотрение популярных инструментов для визуализации, таких как Tableau и Power BI, а также программных библиотек, таких как Matplotlib и Seaborn для Python.

4.3 Перечень лабораторных работ

Лабораторных работ не предусмотрено

4.4 Перечень практических занятий

Семестр № 3

№	Темы практических (семинарских) занятий	Кол-во академических часов
---	---	----------------------------

1	Конфигурация системы для хранения данных	7
2	Инструменты извлечения данных	7

4.5 Самостоятельная работа

Семестр № 3

№	Вид СРС	Кол-во академических часов
1	Подготовка к практическим занятиям (лабораторным работам)	44

В ходе проведения занятий по дисциплине используются следующие интерактивные методы обучения: Онлайн квиз по теме каждой лекции, вебинар

5 Перечень учебно-методического обеспечения дисциплины

5.1 Методические указания для обучающихся по освоению дисциплины

5.1.1 Методические указания для обучающихся по практическим занятиям

Находятся на электронном образовательном ресурсе el.istu.edu

5.1.2 Методические указания для обучающихся по самостоятельной работе:

Находятся на электронном образовательном ресурсе el.istu.edu

6 Фонд оценочных средств для контроля текущей успеваемости и проведения промежуточной аттестации по дисциплине

6.1 Оценочные средства для проведения текущего контроля

6.1.1 семестр 3 | Тест

Описание процедуры.

Экзамен проходит в устном виде по билетам составленным по контрольным вопросам, на подготовку дается 10-15 минут

Критерии оценивания.

Отлично, Хорошо, Удовлетворительно, Неудовлетворительно

6.2 Оценочные средства для проведения промежуточной аттестации

6.2.1 Критерии и средства (методы) оценивания индикаторов достижения компетенции в рамках промежуточной аттестации

Индикатор достижения компетенции	Критерии оценивания	Средства (методы) оценивания промежуточной аттестации
ПК-1.2	Способен разрабатывать архитектуры информационных систем, структуры баз данных информационных систем в	Выполнение индивидуального задания и

	соответствии с особенностями архитектурной спецификации промышленного интернета вещей (IIoT)	практических работ.
--	--	---------------------

6.2.2 Типовые оценочные средства промежуточной аттестации

6.2.2.1 Семестр 3, Типовые оценочные средства для проведения экзамена по дисциплине

6.2.2.1.1 Описание процедуры

Экзамен проходит в устном виде по билетам составленным по контрольным вопросам, на подготовку дается 10-15 минут

6.2.2.1.2 Критерии оценивания

Отлично	Хорошо	Удовлетворительно	Неудовлетворительно
для получения оценки «отлично» необходимо полностью раскрыть теоретическую часть выполненной работы, правильно выполнить все расчеты согласно с выданным заданием и верно ответить на все уточняющие вопросы	для получения оценки «хорошо» необходимо правильно выполнить все расчеты согласно с выданным заданием и верно ответить на уточняющие вопросы	Для получения оценки «удовлетворительно» достаточно правильно выполнить все расчеты согласно с выданным заданием.	Студент получает оценку «не удовлетворительно» если не выполнил расчет или расчеты выполнены не верно.

7 Основная учебная литература

1. Агальцов Базы данных : учебник для вузов специальности 230100 "Информатика и вычислительная техника". Кн. 2 : Распределенные и удаленные базы данных, 2011. - 270.
2. Голицына О. Л. Базы данных : учебное пособие для вузов по направлению 230700 "Прикладная информатика" / О. Л. Голицына, Н. В. Максимов, И. И. Попов, 2012. - 399.
3. Дорофеев А. С. Базы данных : учебное пособие для специальности "Вычислительные машины, комплексы, системы и сети" / А. С. Дорофеев, 2008. - 99.
4. Кузин А. В. Базы данных : учебное пособие по направлению подготовки дипломированных специалистов "Информатика и вычислительная техника" / А. В. Кузин, С. В. Левонисова, 2010. - 311.
5. Гришин Г. Г. Базы данных на предприятиях автомобильного транспорта : учебное пособие для студентов направления 190600.62 - Эксплуатация транспортно-

технологических машин и комплексов, профиль "Автомобили и автомобильное хозяйство" / Г. Г. Гришин, 2013. - 26.

6. Дударева. Информационное обеспечение, базы данных [Электронный ресурс] : учебное пособие: в 2 ч. Ч. 1, 2010. - 54.

7. Дударева. Информационное обеспечение, базы данных [Электронный ресурс] : учебное пособие: в 2 ч. Ч. 2, 2010. - 64.

8 Дополнительная учебная литература и справочная

1. . Змитрович Анатолий Иосифович. Базы данных : учеб. пособие для вузов по спец. 01.02 "Прикл. математика" и 22.04 "Прогр. обеспечение вычисл. техники и автоматизир. систем" / Анатолий Иосифович Змитрович, 1991. - 270.

2. . Мартишин С. А. Базы данных. Практическое применение СУБД SQL и NoSQL-типа для проектирования информационных систем : учебное пособие / С. А. Мартишин, В. Л. Симонов, М. В. Храпченко, 2024. - 368.

3. Максимов Е. М. Базы данных в системах управления производственными процессами : учебное пособие / Е. М. Максимов, Н. Н. Бахтадзе; под ред. А. Б. Путилина, 2011. - 158.

4. Латыпова Р. Р. Базы данных : курс лекций / Р. Р. Латыпова, 2016. - 95.

5. Агальцов. Базы данных. В 2-х кн. Локальные базы данных, 2014. - 349.

6. Агальцов. Базы данных. В 2-х кн. Распределенные и удаленные базы данных, 2014. - 270.

9 Ресурсы сети Интернет

1. <http://library.istu.edu/>
2. <https://e.lanbook.com/>

10 Профессиональные базы данных

1. <http://new.fips.ru/>
2. <http://www1.fips.ru/>

11 Перечень информационных технологий, лицензионных и свободно распространяемых специализированных программных средств, информационных справочных систем

1. Свободно распространяемое программное обеспечение . Microsoft Windows (Подписка DreamSpark Premium Electronic Software Delivery (3 years). Сублицензионный договор №14527/МОС2957 от 18.08.16г.)

12 Материально-техническое обеспечение дисциплины

1. Компьютер "Intel Core i3/DDR 4Gb/HDD 1Tb/GF 1Gb/LCD23' /ИБП"
2. Компьютер "Intel Core i3/DDR 4Gb/HDD 1Tb/GF 1Gb/LCD23' /ИБП"
3. Компьютер "Intel Core i3/DDR 4Gb/HDD 1Tb/GF 1Gb/LCD23' /ИБП"

4. Компьютер "Intel Core i3/DDR 4Gb/HDD 1Tb/GF 1Gb/LCD23' /ИБП"