МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ

«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»
(РУТ (МИИТ)

Рабочая программа дисциплины (модуля),
как компонент образовательной программы
базового высшего образования
по направлению подготовки
09.03.01 Информатика и вычислительная техника,
утвержденной первым проректором ФГАОУ ВО
РУТ (МИИТ) Тимониным В.С.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Качество данных в интеллектуальных транспортных системах

Направление подготовки: 09.03.01 Информатика и вычислительная техника

Направленность (профиль): Цифровая инженерия транспортных процессов

Форма обучения: Очная

Рабочая программа дисциплины (модуля) в виде электронного документа выгружена из единой корпоративной информационной системы управления университетом и соответствует оригиналу

Простая электронная подпись, выданная ФГАОУ ВО РУТ (МИИТ)
ID подписи: 937226
Подписал: руководитель образовательной программы Проневич Ольга Борисовна
Дата: 07.09.2026

1. Общие сведения о дисциплине (модуле).

Целями освоения учебной дисциплины являются формирование у обучающихся системных знаний и практических навыков в области оценки, контроля и управления качеством данных в интеллектуальных транспортных системах (ИТС). В рамках курса рассматриваются стандарты качества данных, методы их оценки (полнота, уникальность, точность, своевременность), подходы к проектированию автоматизированных проверок и расчету индекса доверия к данным. Особое внимание уделяется практическим аспектам внедрения систем мониторинга качества данных в транспортной отрасли.

Задачами освоения дисциплины являются:

- формирование теоретических знаний о стандартах и критериях качества данных, применяемых в интеллектуальных транспортных системах;
- освоение методов оценки качества данных по показателям полноты, уникальности, точности и своевременности;
- развитие практических навыков проектирования чек-листов для автоматизированных проверок качества данных;
- формирование компетенций по расчету индекса доверия к данным (Data Trust Score);
- освоение подходов к проектированию и внедрению систем мониторинга качества данных в транспортной инфраструктуре.

Дисциплина предназначена для формирования компетенций, необходимых для обеспечения высокого качества данных, используемых в интеллектуальных транспортных системах, и принятия обоснованных управленческих решений на основе достоверной информации.

2. Планируемые результаты обучения по дисциплине (модулю).

Перечень формируемых результатов освоения образовательной программы (компетенций) в результате обучения по дисциплине (модулю):

ППК-ДА2 - Способен подготавливать и обрабатывать данные для аналитических исследований.

Обучение по дисциплине (модулю) предполагает, что по его результатам обучающийся будет:

Знать:

- Стандарты качества данных (полнота, уникальность, точность, своевременность) и методологии их оценки в интеллектуальных транспортных системах.

- Подходы к проектированию систем мониторинга качества данных, включая архитектуру, компоненты и этапы внедрения.

Уметь:

- Проектировать чек-листы для автоматизированных проверок качества данных по критериям полноты, уникальности, точности и своевременности.

- Рассчитывать и интерпретировать индекс доверия к данным (Data Trust Score) на основе взвешенной оценки показателей качества.

Владеть:

- Навыками проектирования и внедрения систем мониторинга качества данных в транспортной инфраструктуре.

- Навыками автоматизированной проверки качества данных и визуализации показателей качества на дашбордах.

3. Объем дисциплины (модуля).

3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 3 з.е. (108 академических часа(ов).

3.2. Объем дисциплины (модуля) в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Тип учебных занятий	Количество часов	
	Всего	Семестр №3
Контактная работа при проведении учебных занятий (всего):	48	48
В том числе:		
Занятия лекционного типа	16	16
Занятия семинарского типа	32	32

3.3. Объем дисциплины (модуля) в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации составляет 60 академических часа (ов).

3.4. При обучении по индивидуальному учебному плану, в том числе при ускоренном обучении, объем дисциплины (модуля) может быть реализован полностью в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации.

4. Содержание дисциплины (модуля).

4.1. Занятия лекционного типа.

№ п/п	Тематика лекционных занятий / краткое содержание
1	Тема 1. Введение в управление качеством данных в ИТС. Рассматриваемые вопросы: - Понятие качества данных и его роль в интеллектуальных транспортных системах. - Источники данных в ИТС (сенсоры, камеры, GPS, системы оплаты). - Последствия использования некачественных данных. - Обзор стандартов качества данных (ISO 8000, DAMA DMBOK). - Методологии оценки качества данных (DAMA DMBOK, TDQM).
2	Тема 2. Стандарты качества данных: полнота и уникальность. Рассматриваемые вопросы: - Полнота данных: определение, метрики оценки (пропуски, null-значения). - Уникальность данных: дублирование записей, методы выявления дублей. - Стандарты качества данных (полнота, уникальность). - Влияние неполноты и дублирования на работу ИТС. - Инструменты проверки полноты и уникальности.
3	Тема 3. Стандарты качества данных: точность и своевременность. Рассматриваемые вопросы: - Точность данных: определение, метрики оценки (отклонения, выбросы). - Своевременность данных: актуальность, частота обновления, задержки. - Стандарты качества данных (точность, своевременность). - Влияние неточности и устаревания данных на транспортные процессы. - Методы контроля точности и своевременности.
4	Тема 4. Профилирование данных и выявление аномалий. Рассматриваемые вопросы: - Профилирование данных как инструмент оценки качества. - Статистические методы выявления аномалий и выбросов. - Использование визуализации для анализа распределения данных. - Формирование профиля данных для транспортной системы.
5	Тема 5. Проектирование систем мониторинга качества данных. Рассматриваемые вопросы: - Архитектура системы мониторинга качества данных. - Ключевые компоненты: сбор, валидация, оповещение, визуализация. - Этапы внедрения системы мониторинга. - Интеграция с существующими транспортными системами.
6	Тема 6. Автоматизированные проверки качества данных. Рассматриваемые вопросы:

№ п/п	Тематика лекционных занятий / краткое содержание
	- Типы автоматизированных проверок (синтаксические, семантические, бизнес-правила). - Проектирование чек-листов для автоматизированных проверок . - Обзор инструментов для валидации данных (обзорно, без углубления). - Настройка частоты и триггеров проверок. - Обработка результатов проверок и алертинг.
7	Тема 7. Индекс доверия к данным. Рассматриваемые вопросы: - Понятие индекса доверия к данным. - Методики расчета Data Trust Score. - Составляющие индекса: полнота, точность, своевременность, уникальность. - Взвешенная оценка показателей качества. - Применение Data Trust Score в транспортных системах.
8	Тема 8. Визуализация и управление качеством данных. Рассматриваемые вопросы: - Принципы визуализации показателей качества данных. - Создание дашбордов для мониторинга качества. - Визуализация динамики показателей и индекса доверия. - Кейсы внедрения систем контроля качества в транспортных компаниях. - Тенденции развития Data Quality.

4.2. Занятия семинарского типа.

Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Тема 1. Знакомство с источниками данных в интеллектуальных транспортных системах. Рассматриваемые вопросы: - Анализ типов данных, генерируемых в ИТС (GPS-треки, данные с камер, сенсоры). - Загрузка и первичный анализ транспортных данных. - Выявление потенциальных проблем качества данных.
2	Тема 2. Оценка полноты данных в транспортных системах. Рассматриваемые вопросы: - Анализ пропусков в данных. - Расчет процента полноты для различных источников. - Визуализация пропусков (матрицы, тепловые карты). - Методы обработки пропусков (удаление, заполнение).
3	Тема 3. Оценка уникальности данных и выявление дублей. Рассматриваемые вопросы: - Выявление дублирующихся записей в данных. - Использование методов для поиска дублей. - Оценка влияния дублей на аналитические модели. - Разработка чек-листа для проверки уникальности.
4	Тема 4. Оценка точности данных в транспортных системах. Рассматриваемые вопросы: - Выявление выбросов в данных с использованием статистических методов. - Анализ отклонений и погрешностей в сенсорных данных. - Сопоставление данных с эталонными значениями. - Разработка чек-листа для проверки точности.

№ п/п	Тематика практических занятий/краткое содержание
5	Тема 5. Оценка своевременности данных. Рассматриваемые вопросы: - Анализ временных задержек в поступлении данных. - Оценка актуальности данных. - Анализ частоты обновления данных. - Разработка чек-листа для проверки своевременности.
6	Тема 6. Профилирование данных. Рассматриваемые вопросы: - Применение инструментов профилирования данных. - Анализ статистических характеристик данных. - Выявление аномалий и несоответствий. - Формирование отчета о профиле данных.
7	Тема 7. Проектирование чек-листов для автоматизированных проверок. Рассматриваемые вопросы: - Разработка чек-листов для различных типов данных в ИТС. - Определение пороговых значений для проверок. - Валидация чек-листов на реальных данных. - Документирование чек-листов.
8	Тема 8. Автоматизированные проверки качества данных на Python. Рассматриваемые вопросы: - Разработка скриптов для автоматической проверки полноты и уникальности. - Реализация проверок точности и своевременности. - Формирование отчетов о результатах проверок. - Настройка логирования и алертов.
9	Тема 9. Внедрение автоматизированных проверок с использованием Great Expectations. Рассматриваемые вопросы: - Установка и настройка Great Expectations (обзорно). - Создание ожиданий (Expectations) для транспортных данных. - Проверка данных на соответствие ожиданиям. - Визуализация результатов проверок.
10	Тема 10. Расчет индекса доверия к данным (Data Trust Score). Рассматриваемые вопросы: - Определение весовых коэффициентов для показателей качества. - Расчет компонентов индекса (полнота, уникальность, точность, своевременность). - Нормализация показателей. - Автоматизация расчета индекса.
11	Тема 11. Интерпретация и применение индекса доверия к данным. Рассматриваемые вопросы: - Интерпретация полученных значений индекса. - Применение Data Trust Score в управленческих решениях. - Визуализация динамики индекса доверия. - Формирование рекомендаций по повышению качества данных.
12	Тема 12. Проектирование системы мониторинга качества данных. Рассматриваемые вопросы: - Проектирование архитектуры системы мониторинга для транспортной компании. - Определение ключевых метрик и источников. - Выбор инструментов для реализации системы. - Разработка технического задания.

№ п/п	Тематика практических занятий/краткое содержание
13	Тема 13. Реализация пайплайна мониторинга качества данных. Рассматриваемые вопросы: - Реализация пайплайна сбора и валидации данных. - Настройка периодических проверок. - Обработка и логирование результатов проверок. - Настройка оповещений при отклонениях.
14	Тема 14. Визуализация показателей качества данных. Рассматриваемые вопросы: - Создание дашборда для мониторинга качества данных. - Визуализация динамики показателей полноты, точности, уникальности, своевременности.
15	Тема 15. Комплексный проект: оценка качества данных транспортной компании. Часть 1. Рассматриваемые вопросы: - Анализ реальных данных транспортной компании. - Проведение полной оценки качества данных по всем критериям. - Разработка чек-листов для автоматизированных проверок. - Расчет индекса доверия к данным.
16	Тема 16. Комплексный проект: оценка качества данных транспортной компании. Часть 2. Рассматриваемые вопросы: - Разработка системы мониторинга качества данных. - Создание дашборда для визуализации показателей качества. - Формулировка рекомендаций по повышению качества данных. - Защита проекта и презентация результатов.

4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Изучение литературы и стандартов по качеству данных (ISO 8000, DAMA DMBOK).
2	Изучение документации Great Expectations для валидации данных.
3	Выполнение домашних заданий по оценке качества данных (полнота, уникальность, точность, своевременность).
4	Самостоятельное изучение кейсов внедрения систем контроля качества в транспортных компаниях.
5	Выполнение комплексного проекта по оценке качества данных транспортной компании.
6	Подготовка к промежуточной аттестации.
7	Подготовка к промежуточной аттестации.
8	Подготовка к текущему контролю.

5. Перечень изданий, которые рекомендуется использовать при освоении дисциплины (модуля).

№ п/п	Библиографическое описание	Место доступа
1	Егоров, И. И. Управление данными на этапах жизненного цикла систем : учебно-методическое пособие / И. И. Егоров, А. А. Мишкина. — Москва : РТУ МИРЭА, 2023. — 78 с. — ISBN 978-5-7339-1912-6.	https://e.lanbook.com/book/382628

6. Перечень современных профессиональных баз данных и информационных справочных систем, которые могут использоваться при освоении дисциплины (модуля).

https://habr.com/ru - база знаний в виде статей, обзоров

http://library.miit.ru/ - электронно-библиотечная система научно-технической библиотеки МИИТ

https://superset.apache.org/user-docs/intro/#documentation – документация Apache Superset

https://www.postgresql.org/ - официальная страница PostgreSQL

https://yandex.cloud/ru/docs/datalens/ - официальная документация Yandex Datalens

https://docs.docker.com/ - официальная документация Docker

https://www.metabase.com/docs/latest/ - официальная документация Metabase

https://tableau.pro/ - официальный сайт Tableau с лучшими практиками проектирования дашбордов

https://revealthedata.com/blog/ - Блог о визуализации данных и BI-системах

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, необходимого для освоения дисциплины (модуля).

MS Office (Excel, Word, PowerPoint)
Apache Superset
Yandex Datalens
Metabase
PostgreSQL
Docker
Apache Airflow
Браузер Chrome / Яндекс браузер
Draw.io

8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю).

Компьютер преподавателя с доступом к сети Интернет

Компьютеры студентов с установленным Python-окружением и табличным процессором

Доступ к учебной базе событийных данных и экспериментальным выборкам

Доступ к PostgreSQL и BI-системе для визуализации результатов

Доступ к облачной виртуальной машине или Docker-контейнерам для воспроизводимого анализа

Проектор и экран для проектора

Маркерная доска

9. Форма промежуточной аттестации:

Экзамен в 3 семестре.

10. Оценочные материалы.

Оценочные материалы, применяемые при проведении промежуточной аттестации, разрабатываются в соответствии с локальным нормативным актом РУТ (МИИТ).

Авторы:

руководитель образовательной программы О.Б. Проневич

Согласовано:

Директор Д.В. Паринов

Руководитель образовательной программы О.Б. Проневич

Председатель учебно-методической комиссии Д.В. Паринов