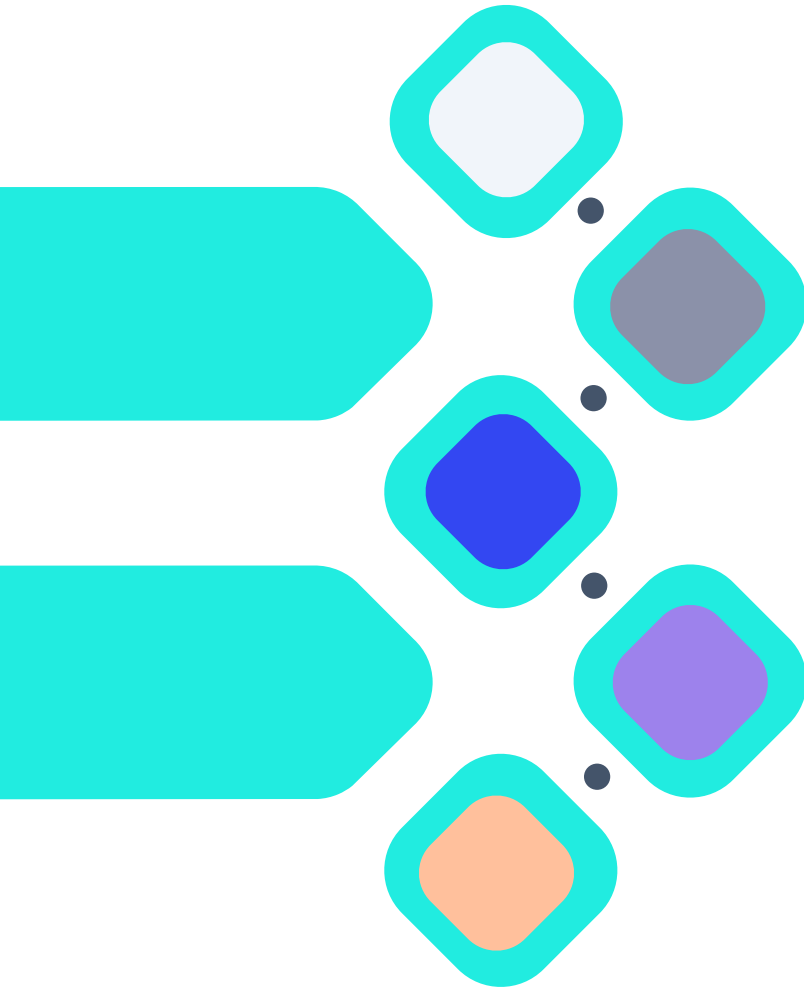


2024

Программа подготовки специалистов по управлению данными в Счетной палате Российской Федерации

Валерий Сохоров
Заместитель директора департамента цифровой
трансформации
Счетная палата Российской Федерации





Счетная палата Российской Федерации в цифрах:

- ◆ > 300 контрольных и экспертно-аналитических мероприятий в год
- ◆ > 150 объектов контроля в год
- ◆ > 600 информационных систем на объектах аудита
- ◆ > 700 постоянно работающих инспекторов

Данные как основа для осуществления государственного аудита (контроля)

- ✓ Данные в государственном секторе требуют комплексного механизма проверки
- ✓ Качественные государственные данные - **основа для реализации эффективного** государственного управления
- ✓ Возрастает роль органов власти как носителей государственных данных в реализации программ и проектов цифровой трансформации
- ✓ Стремительно растет объем межведомственного обмена данными для принятия быстрых и эффективных решений



ДАТА-АНАЛИТИКИ

Дата-аналитики в
департаментах
аудита



ДАТА-СТЮАРДЫ

Дата-стюарды в
департаментах
аудита,
департаменте
цифровой
трансформации и
в ФКУ «ЦЭАИТ СП
РФ»



ДАТА-ИНЖЕНЕР

Дата-инженеры в
ФКУ «ЦЭАИТ СП
РФ»



УПРАВЛЕНИЕ И МЕТОДОЛОГИЯ

Сотрудники
Департамента
цифровой
трансформации

Управление данными

КОМАНДЫ И ЗАДАЧИ

Управление данными в Счетной палате строится на взаимодействии четырех команд по описанию необходимых данных, организации доступа к источникам данных, загрузки, обработки данных и их использованию

● Картирование данных

Создание карт данных осуществляется дата-стюардами при поддержке технических экспертов и заключается в описании источника, набора данных, и их связей

● Загрузка данных

Загрузка и обработка данных осуществляется дата-инженерами на основе созданных карт данных и заключается в разработке ETL/ELT-процессов по получению данных из «озера данных» СП РФ

● Доступ к данным

Доступ к данным обеспечивается дата-стюардом при поддержке дата-инженеров и заключается в организации заключения соглашений, порядков информационного взаимодействия с объектом аудита)

● Использование данных

Использование данных осуществляется дата-стюардами, инспекторским составом в различных аналитических системах, также оценивается уровень надежности и качества данных



Управление данными

«КЛИЕНТСКИЙ ПУТЬ»

Загрузка данных

Загрузка и обработка данных в СУБД (собрать, стандартизировать, сохранить, предоставить, проанализировать)

Витрина данных

Формируется витрина данных, с которой работает инспекторский состав

Создание карты данных

Описание источников и наборов данных



Формирование отчета о качестве данных

Формируется отчет о наличии несоответствий заданным параметрам для дата-стюарда

Доступ к данным

Соглашение об информационном взаимодействии

Проверка качества данных

Проверка данных по заданным параметрам качества

Управление данными

МОДЕЛЬ КОМПЕТЕНЦИЙ

Модель компетенций «Дата-стюард»

Описание уровней проявления

Профессиональные знания, навыки и умения (компетенции)

1 - Базовый уровень (6 баллов)

2 - Средний уровень (от 6 до 12 баллов)

3- Высокий уровень (от 12 до 18 баллов)

1. ТЕХНОЛОГИЧЕСКИЕ ЗНАНИЯ, НАВЫКИ, УМЕНИЯ

Умение работать со структурами данных

- Понимает простые структуры и типы данных: категориальные, количественные, текстовые, временные данных.

- Умеет работать с csv, xml, Excel-форматами данных: импортировать данные в системы анализа данных (PowerQuery, Мастер Импорта Текстов)

- Обладает базовыми знаниями SQL;
- Имеет представление о базах данных и умеет работать с реляционными базами данных (PostgreSQL)

Использование корпоративных инструментов обработки данных

- Обладает теоретическими знаниями о возможностях использования применять не умеет;

- Умеет работать с табличными-форматами данных: преобразовать типы данных в файлах, провести базовые операции со строками, колонками и таблицами (конкатенация и объединение).

- Обладает навыками работы в Полианалисте (и аналогах): умеет агрегировать, фильтровать, визуализировать данные;

- Обладает теоретическими знаниями о принципах функционирования Системы управления данными Счетной палаты.

- Умеет формировать и редактировать в Системе управления данными паспорта карт данных, источников данных, наборов данных, витрин данных.

- Умеет разрабатывать правила надежности/качества данных, подготавливаемых для проведения внешнего государственного аудита.

Модель компетенций «Дата-стюард»

Описание уровней проявления

1 - Базовый уровень
(6 баллов)

2 - Средний уровень
(от 6 до 12 баллов)

3- Высокий уровень
(от 12 до 18 баллов)

2. АНАЛИТИЧЕСКИЕ ЗНАНИЯ, НАВЫКИ, УМЕНИЯ

Профессиональные знания, навыки и умения (компетенции)

Работа с источниками данных

- Способен разобраться в ГИС (умеет работать с графическим интерфейсом).

- Умеет работать с документацией информационных систем и сервисов, включая формирование выгрузки данных внутри ИС

- Способен сформулировать рекомендации по составу и порядку формирования данных при создании и модернизации ими федеральных государственных информационных систем.

Работа с наборами данных

- Умеет выгружать данные в файловом формате (csv, xls,xlsx и др.) из интерфейса ИС

- Умеет анализировать срез данных в формате API(json), SOAP(xml) с использованием специализированных инструментов и/или документации

- Умеет сформировать задание на загрузку данных дата инженеру из любого источника данных руководствуясь только документацией (СМЭВЗ, НСУД)

Знание теории управления данными

- Имеет представление о базовых принципах управления данными.
- Прошел подготовку на дистанционном курсе «Основы управления данными» (Stepik) с сертифицированным результатом менее 100 баллов*.

- Имеет представление о политике и методах организации работы с государственными данными.
- Прошел подготовку на дистанционном курсе 'управления данными» (Stepik) с сертифицированным результатом 100 баллов*.

- Имеет углубленные знания в области работы с данными, подтвержденные дипломами о наличии профильного или дополнительного профессионального образования.

Данные по бухгалтерской финансовой отчетности

Автоматизирована обработка поступающих от Федеральной налоговой службы файлов объемом более 350ГБ данных (более 14 млн строк), необходимых для целей аудита. Данные используются в цифровых решениях департаментов аудита.

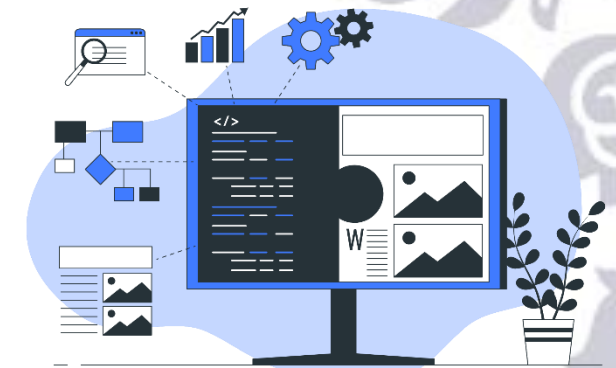


Данные по системе обеспечения законодательной деятельности (СОЗД)

Автоматизирована обработка поступающих 28 тыс. файлов, более 500 тысяч строк, более 5 ГБ данных, необходимых для экспертизы нормативно-правовых актов. Данные используются внутренними информационными системами Счетной палаты

Данные по национальным проектам

Автоматизирована обработка поступающих 25 тыс. файлов, более 10 млн. строк., более 15Гб. Данные необходимы для оценки и анализа планирования и реализации национальных, федеральных и региональных проектов и формирования аналитических срезов, используемых в рамках мониторинга и контрольной и экспертно-аналитической деятельности.



ПЕРЕХОД ОТ ГИС К ДАННЫМ



Доступ к государственным информационным системам*

*ч.4 п.1 статьи 14 Закона о Счетной палате Российской Федерации



Работа с данными объекта аудита**

*ч.3,ч.5 п.1 статьи 14 Закона о Счетной палате Российской Федерации

Особенности работы с ГИС объекта аудита

- Необходимость изучения функционала ГИС и возможностей ГИС;
- Риск отсутствия необходимой информации;
- Длительные сроки подключения к ГИС (требования информационной безопасности);
- Риск невозможности выгрузки необходимых данных;
- Риск отказа в доступе к ГИС со стороны объекта аудита.

Особенности работы с данными объекта аудита

- Необходимость определения перечня необходимых данных;
- Необходимость формулирования требований к данным (качество данных);
- Необходимость формулирования требований к аналитическим панелям;
- Необходимость в дата-стюардах, дата-аналитиках.



- Возможность работы с качественными данными;
- Данные доступны «здесь и сейчас»;
- Возможность построения «светофоров»;
- Возможность глубокого анализа с применением математического и статистического анализа.

Спасибо за внимание

Валерий Сохоров
Заместитель директора департамента цифровой
трансформации
Счетная палата Российской Федерации

