

---

**МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ, СВЯЗИ И МАССОВЫХ  
КОММУНИКАЦИЙ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ  
О РЕГЛАМЕНТЕ (СТАНДАРТЕ) УПРАВЛЕНИЯ ДАННЫМИ ГОСУДАРСТВЕННЫХ  
ИНФОРМАЦИОННЫХ СИСТЕМ, СОЗДАВАЕМЫХ, РАЗВИВАЕМЫХ И  
ЭКСПЛУАТИРУЕМЫХ НА ЕДИНОЙ ЦИФРОВОЙ ПЛАТФОРМЕ РОССИЙСКОЙ  
ФЕДЕРАЦИИ «ГОСТЕХ»**

**Версия 1.0**

**2023 г.**

**СОДЕРЖАНИЕ**

<b>I. ОБЩИЕ ПОЛОЖЕНИЯ</b>	<b>5</b>
1. О ДОКУМЕНТЕ	6
<b>II. ФОРМИРОВАНИЕ МОДЕЛИ УПРАВЛЕНИЯ ДАННЫМИ</b>	<b>7</b>
2. ЦЕННОСТЬ УПРАВЛЕНИЯ ДАННЫМИ	7
3. ПОРЯДОК ДЕЙСТВИЙ ПО СОВЕРШЕНСТВОВАНИЮ УПРАВЛЕНИЯ ДАННЫМИ В ОРГАНИЗАЦИИ	7
4. УПРАВЛЕНИЕ ДАННЫМИ НА УРОВНЕ ОРГАНИЗАЦИИ	8
4.1. ПОРЯДОК ПРИМЕНЕНИЯ ПРОЦЕССОВ УПРАВЛЕНИЯ ДАННЫМИ НА УРОВНЕ ОРГАНИЗАЦИИ.....	8
4.2. ОРГАНИЗАЦИОННАЯ СТРУКТУРА УПРАВЛЕНИЯ ДАННЫМИ НА УРОВНЕ ОРГАНИЗАЦИИ.....	10
<b>5. ОСОБЕННОСТИ ОРГАНИЗАЦИИ НАЦИОНАЛЬНОЙ СИСТЕМЫ УПРАВЛЕНИЯ ДАННЫМИ</b>	<b>18</b>
5.1. ПОРЯДОК ПРИМЕНЕНИЯ ПРОЦЕССОВ УПРАВЛЕНИЯ ДАННЫМИ В РАМКАХ НАЦИОНАЛЬНОЙ СИСТЕМЫ УПРАВЛЕНИЯ ДАННЫМИ.....	18
<b>6. ИНСТРУМЕНТЫ УПРАВЛЕНИЯ ДАННЫМИ</b>	<b>19</b>
<b>III. ПРОЦЕССЫ УПРАВЛЕНИЯ ДАННЫМИ</b>	<b>21</b>
<b>7. ПРОЦЕСС «УПРАВЛЕНИЕ ЖИЗНЕННЫМ ЦИКЛОМ ДАННЫХ»</b>	<b>21</b>
7.1. ВВЕДЕНИЕ.....	21
7.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ.....	22
7.3. ЦЕЛЕВОЙ ПРОЦЕСС.....	22
7.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА.....	26
7.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА.....	26
<b>8. ПРОЦЕСС «УПРАВЛЕНИЕ АРХИТЕКТУРОЙ ДАННЫХ»</b>	<b>27</b>
8.1. ВВЕДЕНИЕ.....	27
8.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ.....	30
8.3. ЦЕЛЕВОЙ ПРОЦЕСС.....	30
8.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА.....	35
8.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА.....	36
<b>9. ПРОЦЕСС «УПРАВЛЕНИЕ ИНТЕГРАЦИЕЙ ДАННЫХ»</b>	<b>37</b>
9.1. ВВЕДЕНИЕ.....	37

9.2.	РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ .....	37
9.3.	ЦЕЛЕВОЙ ПРОЦЕСС .....	37
9.4.	ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА.....	40
9.5.	ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА.....	40
<b>10.</b>	<b>ПРОЦЕСС «УПРАВЛЕНИЕ ОСНОВНЫМИ ДАННЫМИ И НСИ»</b>	<b>41</b>
10.1.	ВВЕДЕНИЕ .....	41
10.2.	РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ .....	41
10.3.	ЦЕЛЕВОЙ ПРОЦЕСС .....	42
10.4.	ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА.....	46
10.5.	ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА.....	46
<b>11.</b>	<b>ПРОЦЕСС «УПРАВЛЕНИЕ КАЧЕСТВОМ ДАННЫХ»</b>	<b>47</b>
11.1.	ВВЕДЕНИЕ .....	47
11.2.	РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ .....	48
11.3.	ЦЕЛЕВОЙ ПРОЦЕСС .....	48
11.4.	ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА.....	53
11.5.	ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА.....	53
<b>12.</b>	<b>ПРОЦЕСС «ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ ДАННЫХ»</b>	<b>55</b>
12.1.	ВВЕДЕНИЕ .....	55
12.2.	РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ .....	55
12.3.	ЦЕЛЕВОЙ ПРОЦЕСС .....	55
12.4.	ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА.....	60
12.5.	ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА.....	60
<b>13.</b>	<b>ПРОЦЕСС «АНАЛИЗ И ПРЕДСТАВЛЕНИЕ ДАННЫХ»</b>	<b>61</b>
13.1.	ВВЕДЕНИЕ .....	61
13.2.	РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ .....	62
13.3.	ЦЕЛЕВОЙ ПРОЦЕСС .....	63
13.4.	ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА.....	69
13.5.	ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА.....	69

<b>14. ПРОЦЕСС «УПРАВЛЕНИЕ МЕТАДААННЫМИ И ПРОИСХОЖДЕНИЕМ ДАННЫХ»</b>	<b>71</b>
14.1. ВВЕДЕНИЕ .....	71
14.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ .....	73
14.3. ЦЕЛЕВОЙ ПРОЦЕСС .....	73
14.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА.....	76
14.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА.....	76
<b>IV. ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ ИИ В РАМКАХ ЗАДАЧ ПО РАБОТЕ С ДАННЫМИ</b>	<b>78</b>
<b>ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ</b>	<b>79</b>
<b>ПРИЛОЖЕНИЕ 1. РЕФЕРЕНСНАЯ АРХИТЕКТУРА ХРАНИЛИЩА ДАННЫХ ОРГАНИЗАЦИИ</b>	<b>83</b>
<b>ПРИЛОЖЕНИЕ 2. РЕАЛИЗАЦИЯ ХРАНИЛИЩА ДАННЫХ ОГВ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЧЕСКИХ КОМПОНЕНТОВ «ГОСТЕХ»</b>	<b>85</b>
<b>ПРИЛОЖЕНИЕ 3. РАСПРЕДЕЛЕНИЕ ОТВЕТСТВЕННОСТИ ЗА ВЫПОЛНЕНИЕ МЕРОПРИЯТИЙ ПО УПРАВЛЕНИЮ ДАННЫМИ НА ПЛАТФОРМЕ ГОСТЕХ</b>	<b>86</b>
<b>ПРИЛОЖЕНИЕ 4. УПРАВЛЕНИЕ ДАННЫМИ НА ОСНОВЕ ПРЕДМЕТНЫХ ОБЛАСТЕЙ</b>	<b>91</b>

## **I. ОБЩИЕ ПОЛОЖЕНИЯ**

Методические рекомендации «О регламенте управления данными государственных информационных систем, создаваемых, развиваемых и эксплуатируемых на единой цифровой платформе Российской Федерации «ГосТех» разработаны в соответствии с пунктом 17 Положения о единой цифровой платформе Российской Федерации «ГосТех», утвержденного постановлением Правительства Российской Федерации от 16 декабря 2022 № 2338.

Ответственным за применение настоящих методических рекомендаций являются федеральные органы исполнительной власти, органы исполнительной власти субъектов Российской Федерации, органы местного самоуправления, государственные внебюджетные фонды, а также государственные корпорации, государственные компании, публично-правовые компании и иные организации, наделенные в соответствии с федеральными законами полномочиями по осуществлению от имени Российской Федерации государственного управления в установленной сфере деятельности по обработке информации в информационных системах (далее – организации).

При применении настоящих методических рекомендаций организации определяют объем мероприятий, которые необходимо выполнить операторами информационных систем или пользователями платформы «ГосТех».

Методические рекомендации обязательны для исполнения пользователями платформы «ГосТех». Организациям, не являющимся пользователями платформы «ГосТех», рекомендуется руководствоваться Регламентом управления данными для управления данными на уровне организации в части, не связанной с использованием платформы «ГосТех».

В случае принятия нормативных правовых актов Российской Федерации, требующих пересмотра подходов к управлению государственными данными, а также с целью повышения эффективности управления данными, президиум Правительственной комиссии по цифровому развитию, использованию информационных технологий для улучшения качества жизни и условий ведения предпринимательской деятельности утверждает изменения в настоящие Методические рекомендации.

## 1. О ДОКУМЕНТЕ

Под управлением данными понимается комплекс взаимосвязанных, систематически реализуемых мероприятий, которые обеспечивают передачу, обработку, проверку, защиту и максимально эффективное использование данных для оказания государственных, муниципальных и иных услуг, выполнения государственных функций, принятия и реализации решений, производства научного знания, развития технологий и создания иных экономических и общественных благ.

Все мероприятия сгруппированы в 8 процессов управления данными, выполнение которых позволит организациям повысить эффективность использования данных в своей деятельности:

- Управление жизненным циклом данных;
- Управление архитектурой данных;
- Управление интеграцией данных;
- Управление основными данными и НСИ;
- Управление качеством данных;
- Управление безопасностью данных;
- Анализ и представление данных;
- Управление метаданными и происхождением данных.

Процессы управления данными взаимосвязаны, поэтому выполнение рекомендуемых мероприятий одного процесса, невозможно без выполнения отдельных мероприятий, рекомендуемых для других процессов. В тексте документа приведены соответствующие ссылки.

Мероприятия процессов управления данными приведены в соответствии с рекомендуемой последовательностью их выполнения. При этом допускается пропускать или откладывать выполнение отдельных мероприятий до появления оптимальных условий для их выполнения.

Описание каждого из мероприятий в документе состоит из:

- детального описания мероприятия и задач;
- исходной информации, необходимой для выполнения мероприятия;
- результата выполнения мероприятия;
- используемых инструментов управления данными;
- рекомендаций и примеров выполнения мероприятия (при необходимости);

Реализация мероприятий управления данными выполняется сотрудниками, закрепленными за определенными ролями.

Управление данными реализуется на трех уровнях: уровень организации, уровень домена деятельности, национальный уровень. Для создания управления данными на каждом из уровней в документе приводится перечень необходимых к выполнению процессов и состав организационной структуры для выполнения данных процессов.

Документ ориентирован на руководителей цифровой трансформации организаций, а также иных сотрудников, ответственных за управление данными в организации или отрасли (домене деятельности).

Документ предполагает непрерывное применение процессов управления данными в информационных системах, которые будут созданы или доработаны в рамках развития, а также для уже существующих информационных систем – в него включены критерии оценки уровня зрелости процессов управления данными для каждого процесса управления данными. Такие критерии будут использоваться для формирования рейтинга организаций при оценке их уровня зрелости в области управления данными.

## II. ФОРМИРОВАНИЕ МОДЕЛИ УПРАВЛЕНИЯ ДАННЫМИ

### 2. ЦЕННОСТЬ УПРАВЛЕНИЯ ДАННЫМИ

В условиях создания и накопления большого объема информации организациям необходимо выстроить системную работу с данными для целей их эффективного использования.

При внедрении управления данными организации преследуют следующие цели:

- Создание новых услуг и сервисов в том числе за счет использования данных более высокого качества;
- Получение расширенных сведений об объектах учёта для создания новых услуг и сервисов;
- Повышение качества принятия управленческих решений за счет расширения состава и повышения качества используемых данных, а также применения продвинутых инструментов аналитики;
- Повышение скорости обработки и предоставления данных;
- Снижение количества ошибок данных и затрат на их устранение;
- Сокращение затрат на обработку (сбор, создание, хранение, изменение, предоставление) данных;
- Снижение рисков утечек данных за счет применения проверенных архитектурных решений и методических рекомендаций.

Для этих целей организациям необходимо реализовать мероприятия по управлению данными.

Примеры показателей для оценки ценности от реализации мероприятий по управлению данными приведены в разделе «Оценка эффективности процесса» для каждого процесса управления данными.

### 3. ПОРЯДОК ДЕЙСТВИЙ ПО СОВЕРШЕНСТВОВАНИЮ УПРАВЛЕНИЯ ДАННЫМИ В ОРГАНИЗАЦИИ

Для совершенствования управления данными необходимо сформировать организационную структуру и наладить выполнение процессов в рамках мероприятий управления данными.

При управлении данными на уровне организации необходимо учитывать, что она является частью многоуровневой модели, которая состоит из следующих уровней:

#### - **Управление данными на уровне организации:**

Организациям необходимо обеспечить выполнение всех процессов управления данными для каждой информационной системы в отдельности (включая отдельные компоненты ИС), а также обеспечить выполнение процессов управления данными при взаимодействии нескольких информационных систем (включая отдельные компоненты ИС) внутри организации.

#### - **Управление данными на уровне домена деятельности:**

Организациям необходимо логически отнести данные, обрабатываемые в информационных системах организации, к доменам деятельности, обеспечить согласованность определений данных в рамках домена деятельности, обеспечить доступность и использование данных для целей создания сервисов, ориентированных на потребности потребителей данных домена, обеспечить связанность и консистентность данных.

#### - **Управление данными на национальном уровне:**

Необходимо обеспечить межведомственное взаимодействие информационных систем, согласование данных, связанность и консистентность данных между различными организациями и между доменами деятельности, а также использование общегосударственных данных (например, централизованных справочников).

На рисунке 1 представлена многоуровневая система управления данными:



Рис. 1. Многоуровневая система управления данными.

## 4. УПРАВЛЕНИЕ ДАННЫМИ НА УРОВНЕ ОРГАНИЗАЦИИ

### 4.1. ПОРЯДОК ПРИМЕНЕНИЯ ПРОЦЕССОВ УПРАВЛЕНИЯ ДАННЫМИ НА УРОВНЕ ОРГАНИЗАЦИИ

При организации процессов управления данными необходимо применять принцип дата-центричности<sup>1</sup>, при котором организация управляет данными как единым активом на всех этапах их жизненного цикла. Данный принцип предполагает, что управление данными, включая обеспечение их высокого качества, становится одним из важных аспектов деятельности организации.

Применение процессов управления данными различается в зависимости от статуса информационных систем организации:

- При создании или развитии информационной системы применяются все 8 процессов управления данными в зависимости от уровня зрелости, к которому стремится организация;
- При эксплуатации информационных систем необходимо оценить текущий уровень зрелости процессов управления данными, определить требуемый уровень зрелости и запланировать мероприятия по повышению уровня зрелости до необходимого уровня.

Конкретные мероприятия по внедрению регламента определяются организацией в зависимости от уровня ее зрелости по управлению данными, текущих приоритетов организации, а также доступности соответствующих инструментальных средств и ресурсов.

При создании или развитии информационных систем организациям рекомендуется придерживаться следующего порядка реализации процессов управления данными на уровне организации:

#### 1) Управление жизненным циклом данных:

Внедрение регламента управления данными начинается с реализации процесса жизненного цикла данных, в рамках которого организация контролирует процессы создания, получения и использования данных, включая представление и передачу данных.

Данный процесс является основополагающим, поскольку определяет область дальнейшего применения процессов управления данными в каждой информационной системе.

#### 2) Управление архитектурой данных:

После определения области применения процессов управления данными в организации, необходимо документировать архитектуру данных существующих информационных систем, а для

<sup>1</sup> Дата-центричный подход – подход к управлению информационными системами организации, в котором центральным звеном являются не отдельные информационные системы, а данные, которые хранятся и используются в информационных системах. Дата-центричный подход требует реализации: 1) Единого хранилища данных на уровне организации/Домена, данные в котором объединены единой логической моделью данных с привязкой к физическим моделям хранения данных. Данные единого виртуального хранилища организации/Домена должны быть согласованы с т.з. НПА и НСИ, храниться с учетом исторических изменений данных, учитывать исторические изменения структуры хранения данных. 2) Единых принципов обработки и предоставления данных пользователям. 3) Создания хранилища метаданных на уровне организации/Домена и интеграция с хранилищем метаданных ЕИП НСУД.



создаваемых информационных систем - спроектировать целевую архитектуру данных. Как документирование, так и проектирование архитектуры данных включают в себя определение ключевых сущностей, обрабатываемых в информационной системе, и взаимосвязей между ними.

Также процесс управления архитектурой данных сопровождает физическую реализацию модели данных информационной системы.

### **3) Управление интеграцией данных:**

Процесс управления интеграцией данных реализуется при:

- получении данных от поставщиков данных (других информационных систем и источников данных);
- преобразовании и обработке данных в ИС организации;
- передаче данных потребителям данных (другим информационным системам и источникам данных).

Организацию передачи данных потребителям необходимо выполнять с учетом требований к защите данных, формируемых в рамках процесса управления безопасностью данных.

### **4) Управление основными данными и НСИ:**

Мероприятия по управлению основными данными и НСИ выполняются с учетом реализации **процесса управления жизненным циклом данных** и **процесса архитектуры данных** и начинаются с определения основных и справочных данных, обрабатываемых в информационной системе.

В рамках данного процесса выполняются мероприятия по работе с основными и справочными данными, в том числе при помощи интеграции с централизованными системами инфраструктуры электронного правительства и инструментами ведения основных и справочных данных платформы «ГосТех» и ЕСНСИ.

### **5) Управление качеством данных:**

Управление качеством данных осуществляется от этапа создания данных до их анализа и передачи во внешние информационные системы. Обеспечение высокого качества данных является одной из главных функций управления данными, а реализация мероприятий данного процесса является центральным звеном управления данными.

### **6) Управление безопасностью данных:**

Мероприятия по обеспечению безопасности данных выполняются на протяжении всего жизненного цикла обработки данных с учетом категорий используемых данных, а также с учетом осуществления информационного обмена с внешними организациями.

### **7) Анализ и использование данных:**

Процесс анализа и использования данных выполняется после определения потребителей данных. Мероприятия процесса нацелены на формирование представления данных потребителям, в том числе, для целей проведения последующего анализа.

Процесс может применяться как для информационных систем, поддерживающих выполнение специфичных процессов, со встроенным функционалом формирования отчетности для потребителей, так и для информационных систем, полностью предназначенных для выполнения аналитических функций.

Процесс предполагает предоставление данных потребителям. Требования потребителей к формируемой аналитике должны быть учтены при реализации мероприятий по управлению архитектурой данных и управлению интеграцией данных.

### **8) Управление метаданными:**

Управление метаданными применяется на протяжении всего жизненного цикла данных за исключением этапа архивирования и удаления данных. В рамках управления метаданными осуществляется классификация и описание данных организации на физическом, логическом и концептуальном уровнях; фиксируются сведения, описывающие технические компоненты информационных систем, а также сведения о процессах, выполняемых в информационных системах организации.

Использование концептуального уровня описания данных обеспечивает согласованность терминологии в рамках соответствующей предметной области (домена) и в масштабах всего государства.

Общая схема с областями применения процессов управления данными для информационной системы приведена на рисунке ниже.

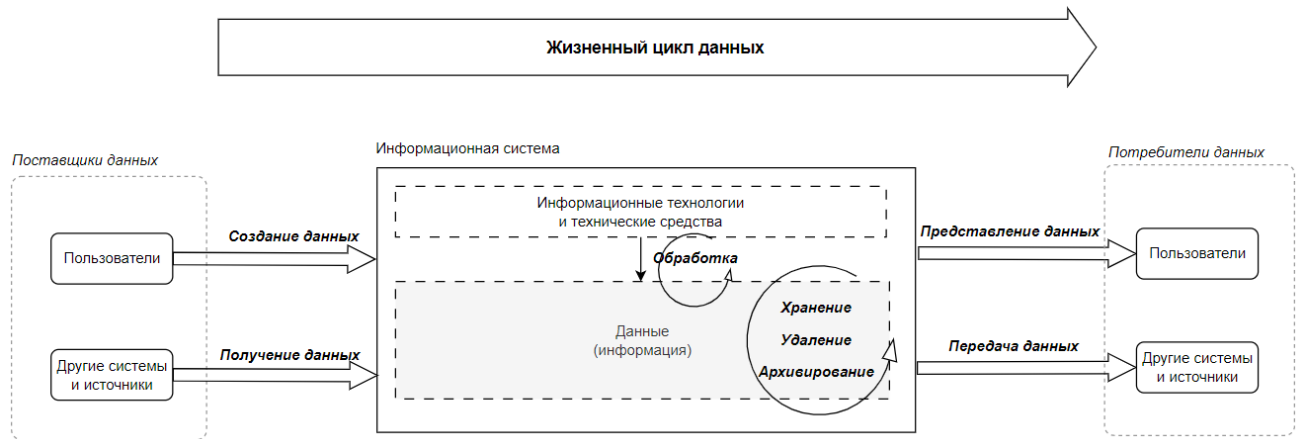


Рис. 2. Область применения процессов управления данными для информационной системы.

В таблице 1 приведен рекомендуемый состав первоочередных процессов, направленных на обеспечение выполнения наиболее распространенных задач организаций, которые связаны с управлением данными.

Таблица 1. Рекомендуемый состав первоочередных процессов и мероприятий.

Процессы	Создание транзакционной ГИС	Создание аналитической ГИС	Создание витрин данных	Текущая эксплуатация ГИС
Управление жизненным циклом данных	✓	✓	✓	✓
Управление архитектурой данных	✓	✓	✓	✓
Управление основными данными и НСИ	✓	✓		✓
Управление качеством данных	✓	✓	✓	✓
Обеспечение безопасности данных	✓	✓	✓	✓
Анализ и представление данных		✓	✓	
Управление метаданными	✓	✓	✓	✓

## 4.2. ОРГАНИЗАЦИОННАЯ СТРУКТУРА УПРАВЛЕНИЯ ДАННЫМИ НА УРОВНЕ ОРГАНИЗАЦИИ

В процессах управления данными участвуют сотрудники, выполняющие роли, показанные на Рис. 3. Команда по управлению данными в организации распределяются по сотрудникам организации в зависимости от размера организации, объемов и сложности обрабатываемых данных. Для небольших организаций, обрабатывающих небольшие объемы данных, допускается совмещение нескольких ролей одним сотрудником.



Рис. 3. Команда по управлению данными в организации.

<p><b>Роль Ответственного за управление данными</b></p>	<p>Ответственный за управление данными информационных систем организации (далее – ответственный за управление данными) обеспечивает исполнение задач, связанных с формированием и функционированием системы управления данными организации:</p> <ul style="list-style-type: none"> <li>- обеспечение исполнения настоящих Методических рекомендаций в организации;</li> <li>- утверждение планов мероприятий по организации управления данными в организации и контроль их исполнения;</li> <li>- ресурсное обеспечение мероприятий по формированию управления данными в организации;</li> <li>- организация мероприятий по обучению и формированию необходимых компетенций по управлению данными в организации;</li> <li>- организация взаимодействия с лицами, выполняющими те или иные функциональные роли в системах управления государственными данными более высокого уровня, в частности – системы управления государственными данными на платформе «ГосТех», а также лицами, представляющими внешних потребителей данных или поставщиков данных, используемых организацией.</li> </ul>
<p><b>Роль Директора по данным</b></p>	<p>Директор по данным отвечает за бесперебойную работу управления данными в организации и выполняет следующие функции:</p> <ul style="list-style-type: none"> <li>- планирование и организацию мероприятий по управлению данными, включая запуск и поддержание процессов управления данными;</li> <li>- контроль за реализацией мероприятий по управлению данными;</li> <li>- руководство другими лицами, которым поручено выполнение иных функциональных задач по управлению данными;</li> <li>- оценку компетенций по управлению данными лиц, участвующих в управлении данными, планирование и проведение мероприятий по обучению / развитию компетенций указанных лиц;</li> <li>- координацию с кураторами ГИС, экспертами по информационной безопасности и потребителями данных в части развития управления данными;</li> <li>- взаимодействие с лицами, выполняющими те или иные функциональные роли в системах управления государственными данными более высокого уровня, в частности – системы управления</li> </ul>

	<p>государственными данными на платформе «ГосТех», а также лицами, представляющими внешних потребителей данных или поставщиков данных.</p> <p>В случаях, когда количество видов данных, обрабатываемых организацией, и/или число и разнообразие процессов, в рамках которых ведется обработка данных, не велики, директор по данным может совмещать выполнение своих функций с выполнением иных функциональных ролей в рамках организационной системы управления данными кроме ответственного за управление данными.</p>
<p><b>Роль Куратора данных</b></p>	<p>Кураторы данных – работники организации, которые осуществляют выполнение внутренних процессов организации, в рамках которых выполняется создание, обработка и предоставление данных.</p> <p>Кураторы данных выполняют следующие задачи, связанные с управлением данными организации:</p> <ul style="list-style-type: none"> <li>- формирование состава терминов и показателей в ИС;</li> <li>- организация работы по описанию терминов и показателей в ИС;</li> <li>- решение разногласий по данным в ИС;</li> <li>- консолидация требований по развитию предметной области и формируемых в рамках нее данных;</li> <li>- согласование локальных нормативных актов по ведению справочных данных и метаданных;</li> <li>- согласование показателей и критериев качества данных;</li> <li>- определение экспертов по данным;</li> </ul> <p>Число кураторов данных зависит от количества ведущихся в организации процессов обработки данных, предоставления данных, а также процессов, в результате которых данные создаются или предоставляются.</p>
<p><b>Роль Эксперта по данным</b></p>	<p>В управлении данными должны участвовать работники организации, которые являются экспертами с отраслевой экспертизой и знанием НПА.</p> <p>Эксперты по данным могут назначаться организацией в соответствующем порядке из числа работников как занимающих должности государственных гражданских служащих, так и находящихся в подведомственных организациях.</p> <p>Эксперты по данным выполняют следующие задачи, связанные с управлением данными организации:</p> <ul style="list-style-type: none"> <li>- разработка документации бизнес-терминов;</li> <li>- организация необходимых процессов согласования и принятия решений о методике расчета показателей;</li> <li>- проверка корректности использования терминов и показателей (при необходимости).</li> </ul> <p>Число экспертов по данным зависит от количества видов данных, обрабатываемых в организации.</p>
<p><b>Роль Эксперта по качеству данных</b></p>	<p>В управлении данными участвуют работники организации, которые являются ответственными за качество данных.</p> <p>Эксперты по качеству данных могут назначаться организацией в соответствующем порядке как из числа работников организации, так из числа работников подведомственных организаций.</p> <p>Эксперт по качеству данных выполняет следующие задачи, связанные с управлением данными организации:</p> <ul style="list-style-type: none"> <li>- измерение метрик качества данных;</li> <li>- разработка показателей и требований для мониторинга качества</li> </ul>

	<p>данных;</p> <ul style="list-style-type: none"> <li>- исследование и мониторинг качества данных;</li> <li>- разработка требований к системам и процессам с целью повышения качества данных;</li> <li>- профилирование данных на предмет соответствия требованиям к качеству данных и выработка рекомендаций по устранению выявленных проблем с качеством данных;</li> <li>- организация работы с инцидентами и проблемами качества данных, взаимодействие с кураторами данных;</li> <li>- детализация требований к данным, необходимых для обеспечения качества данных;</li> <li>- подготовка данных для решения разовых задач;</li> <li>- участие в разработке инструментов для автоматизированного мониторинга качества данных;</li> <li>- организацию работ по повышению качества данных.</li> </ul>
<p><b>Роль Архитектора данных</b></p>	<p>Архитектор данных выполняет следующие задачи, связанные с управлением данными организации:</p> <ul style="list-style-type: none"> <li>- развитие архитектуры данных, включая оптимизацию потоков данных;</li> <li>- разработка и контроль соблюдения правил и стандартов построения архитектуры данных;</li> <li>- проверка соблюдения правил и стандартов построения архитектуры данных;</li> <li>- контроль по исправлению выявленных отклонений от стандарта построения архитектуры данных и блокирование разработки решений, не соответствующих целевой архитектуре данных;</li> <li>- синхронизацию архитектуры данных с ИТ-архитектурой;</li> <li>- организацию работ по формированию и ведению модели данных организации.</li> </ul>
<p><b>Роль Эксперта по информационной безопасности</b></p>	<p>Эксперты по информационной безопасности в организации участвуют в решении следующих задач по управлению данными:</p> <ul style="list-style-type: none"> <li>- выявление отклонений от стандартов предоставления доступа к данным;</li> <li>- категоризация наборов данных в соответствии с требованиями информационной безопасности;</li> <li>- разработка правил информационной безопасности при обработке данных в организации;</li> <li>- определение требований по информационной безопасности к архитектуре данных;</li> <li>- организация разработки правил безопасности при передаче данных третьим лицам и/или предоставления API для запроса данных извне;</li> <li>- анализ данных, критичных к передаче третьим лицам и/или к предоставлению API для запроса данных извне;</li> <li>- согласование запросов на выгрузку данных из ГИС;</li> <li>- согласование требований к разграничению доступа к данным по ролям;</li> <li>- регулярный пересмотр предоставленных доступов;</li> <li>- формирования требований к защите информации, содержащейся в информационной системе;</li> <li>- разработка системы защиты информации информационной системы;</li> <li>- внедрение системы защиты информации информационной</li> </ul>

	<p>системы;</p> <ul style="list-style-type: none"> <li>- аттестация информационной системы по требованиям защиты информации (далее - аттестация информационной системы) и ввод ее в действие;</li> <li>- обеспечение защиты информации в ходе эксплуатации аттестованной информационной системы;</li> <li>- обеспечение защиты информации при выводе из эксплуатации аттестованной информационной системы или после принятия решения об окончании обработки информации</li> <li>- обнаружение и идентификации инцидентов, в том числе отказов в обслуживании, сбоев (перезагрузок) в работе технических средств, программного обеспечения и средств защиты информации, нарушений правил разграничения доступа, неправомерных действий по сбору информации, внедрений вредоносных компьютерных программ (вирусов) и иных событий, приводящих к возникновению инцидентов проведение проверок соблюдения правил безопасности по работе с данными;</li> <li>- проведение контроля (анализа) защищенности информации с учетом особенностей функционирования информационной системы;</li> <li>- проведение анализа и оценки функционирования информационной системы и ее системы защиты информации, включая анализ и устранение уязвимостей и иных недостатков в функционировании системы защиты информации информационной системы;</li> <li>- документирование процедур и результатов контроля по обеспечению уровня защищенности информации, содержащейся в информационной системе;</li> <li>- принятие решения по результатам контроля за обеспечением уровня защищенности информации, содержащейся в информационной системе,</li> </ul> <p>о необходимости доработки (модернизации) модуля системы защиты информации.</p>
<p><b>Роль Инженер данных</b></p>	<p>Инженеры данных в организации участвуют в решении следующих задач по управлению данными:</p> <ul style="list-style-type: none"> <li>- разработка механизмов передачи данных между информационными системами организации;</li> <li>- разработка физической модели данных информационных систем и хранилища данных организации;</li> <li>- реализация витрин данных организации;</li> <li>- техническая реализация правил контроля качества данных на уровне информационных систем и хранилища данных организации;</li> <li>- документирование разработки информационной системы и хранилища данных организации.</li> </ul>
<p><b>Роль Аналитика данных</b></p>	<p>Аналитики данных в организации участвуют в решении следующих задач по управлению данными:</p> <ul style="list-style-type: none"> <li>- реализация расчетов аналитических метрик в хранилище данных или аналитической СУБД;</li> <li>- настройка аналитических инструментов;</li> <li>- формирование аналитики и ее представление конечным пользователям;</li> <li>- эксплуатация и поддержка работоспособности аналитических расчетов и отображения данных;</li> <li>- (опционально) разработка аналитических информационных панелей в инструментах визуализации данных (BI инструментах).</li> </ul>

<b>Роль Куратора ГИС</b>	<p>Лицо, назначенное на роль Куратора ГИС органа или организации государственного сектора, участвует в выполнении следующих задач по управлению данными:</p> <ul style="list-style-type: none"> <li>- в развитии ГИС как источника данных;</li> <li>- в обеспечении выполнения требований к качеству данных;</li> <li>- в обеспечении устранения проблем и инцидентов качества данных в ГИС;</li> <li>- в формировании требований к разграничению прав пользователей ГИС и согласовании доступа пользователей к ГИС;</li> <li>- в формировании требований к архивации и удалению данных ГИС.</li> </ul> <p>Роль Куратора ГИС выполняет Руководитель Цифровой Трансформации органов власти.</p>
<b>Роль Потребителя данных</b>	<p>Потребителями данных являются лица, использующие данные для выполнения стоящих перед ними задач и/или принятия управленческих решений по итогам анализа данных.</p>

В Таблице 2 представлены основные компетенции, которые должны быть у ролей по управлению данными.

*Таблица 2. Компетенции ролей по управлению данными организации*

<b>Компетенции Куратора данных</b>	<ol style="list-style-type: none"> <li>1. Понимание процессов и потребностей потребителей данных;</li> <li>2. Понимание предметной области данных;</li> <li>3. Навык планирования, координирования и контроля процессов сбора, хранения, использования и обновления данных;</li> <li>4. Знание законодательства и требований в области хранения и обеспечения безопасности данных;</li> <li>5. Компетенции в области управления качеством данных в части требований, контролей и мониторинга;</li> <li>6. Знание методов сбора, хранения и обработки данных;</li> <li>7. Понимание процессов управления метаданными, основными данными и НСИ;</li> <li>8. Навык разработки регламентов ведения справочников основных данных.</li> </ol>
<b>Компетенции Эксперта по данным</b>	<ol style="list-style-type: none"> <li>1. Понимание процессов и потребностей потребителей данных;</li> <li>2. Понимание предметной области данных;</li> <li>3. Знание утвержденных шаблонов и правил наименования терминов;</li> <li>4. Компетенции в разработке и согласовании методик расчета показателей;</li> <li>5. Понимание процессов разработки и ведения глоссариев данных;</li> <li>6. Навык определения и описания критических данных;</li> <li>7. Навык работы с базами данных и понимание структуры данных и моделей данных;</li> <li>8. Понимание моделей архитектуры данных и потоков данных.</li> </ol>
<b>Компетенции Эксперта по качеству данных</b>	<ol style="list-style-type: none"> <li>1. Понимание методологий управления качеством данных, включая профилирование, очистку и валидацию.</li> <li>2. Знание инструментов и технологий управления качеством данных;</li> <li>3. Навык разработки метрик качества данных и проведения оценки качества данных;</li> <li>4. Навык работы с метаданными и основными данными;</li> <li>5. Навык анализа данных и выявления проблем качества данных;</li> <li>6. Навык работы с базами данных, понимание структуры данных и моделей данных.</li> </ol>

<p><i>Компетенции Архитектора данных</i></p>	<ol style="list-style-type: none"> <li>1. Компетенции моделирования и ведения государственной модели данных (концептуальной, логической, физической);</li> <li>2. Экспертиза в области выстраивания архитектуры данных, в частности оптимизация потоков данных;</li> <li>3. Навык интеграции данных из различных источников;</li> <li>4. Навык проектирования баз данных.</li> </ol>
<p><i>Компетенции Эксперта по информационной безопасности</i></p>	<ol style="list-style-type: none"> <li>1. Понимание законодательных и нормативных требований в области защиты данных;</li> <li>2. Знание методологий и стандартов информационной безопасности (например, ISO 27001);</li> <li>3. Знание правил предоставления доступов к данным;</li> <li>4. Понимание принципов категоризации, классификации и шифрования данных;</li> <li>5. Понимание устройства архитектуры и потоков данных;</li> <li>6. Понимание принципов обмена данными с точки зрения информационной безопасности;</li> <li>7. Навык построения модели угроз информационной безопасности, методов их обнаружения и предотвращения;</li> <li>8. Понимание подходов к расследованию инцидентов информационной безопасности;</li> <li>9. Навыки работы с инструментами обеспечения безопасности данных;</li> <li>10. Умение разрабатывать и внедрять политики безопасности данных;</li> <li>11. Навыки аудита информационной безопасности и умение разрабатывать и проводить проверки безопасности данных;</li> <li>12. Навыки определения актуальных угроз безопасности информации, разработки и актуализация модели угроз безопасности информации и модели нарушителя;</li> <li>13. Навыки разработки и актуализации и на основе типовой модели угроз безопасности информации и модели нарушителя частных моделей угроз безопасности информации и моделей нарушителя ГИС, сопровождение согласования частных моделей угроз безопасности информации и моделей нарушителя ГИС со ФСТЭК России и ФСБ России.</li> <li>14. Наличие высшего образования по направлению подготовки «Информационная безопасность» или иного высшего профессионального образования и обучения по программам повышения квалификации по направлению «Информационная безопасность»: программы повышения квалификации специалистов, работающих в области технической защиты конфиденциальной информации и/или в области обеспечения безопасности критической информационной инфраструктуры Российской Федерации.</li> </ol>
<p><i>Компетенции Инженера данных</i></p>	<ol style="list-style-type: none"> <li>1. Навыки разработки и актуализации механизмов передачи и обновления данных (ETL/ELT);</li> <li>2. Знания языка структурированных запросов к данным (SQL);</li> <li>3. Навык проектирования баз данных и разработки витрин данных.</li> </ol>
<p><i>Компетенции Аналитика данных</i></p>	<ol style="list-style-type: none"> <li>1. Навыки анализа данных;</li> <li>2. Навыки проектирования аналитических отчетов;</li> <li>3. Навыки формулирования результатов и выводов по итогам анализа данных;</li> <li>4. Знания высокоуровневых языков программирования для анализа данных (Python и пр.);</li> </ol>



	5. Навыки работы в одном или более аналитических BI-инструментов для визуализации данных.
<i>Компетенции Куратора ГИС</i>	<ol style="list-style-type: none"> <li>1. Знание состава источников данных и потребителей данных ГИС;</li> <li>2. Знание основных принципов архитектуры ГИС и умение принимать решения по ее улучшению и поддержке;</li> <li>3. Навыки работы с поставщиками ИТ-услуг и умение управлять отношениями с ними;</li> <li>4. Умение управлять бюджетом на развитие и поддержку ГИС;</li> <li>5. Знание основных стандартов и методологий управления ИТ-сервисами;</li> <li>6. Навыки работы с данными и базами данных;</li> <li>7. Навыки коммуникации и работы с другими кураторами ГИС, архитекторами данных и специалистами по ИТ.</li> </ol>
<i>Компетенции Потребителя данных</i>	<ol style="list-style-type: none"> <li>1. Навык формулирования требований к данным и отчетам, а также запросов на поставку данных;</li> <li>2. Знание методов анализа и интерпретации данных;</li> <li>3. Знание основных инструментов анализа данных и умение работать с ними;</li> <li>4. Умение определять критерии качества данных;</li> <li>5. Умение принимать решения на основе данных и делать выводы на основе анализа данных.</li> </ol>

## **5. ОСОБЕННОСТИ ОРГАНИЗАЦИИ НАЦИОНАЛЬНОЙ СИСТЕМЫ УПРАВЛЕНИЯ ДАННЫМИ**

### **5.1. ПОРЯДОК ПРИМЕНЕНИЯ ПРОЦЕССОВ УПРАВЛЕНИЯ ДАННЫМИ В РАМКАХ НАЦИОНАЛЬНОЙ СИСТЕМЫ УПРАВЛЕНИЯ ДАННЫМИ**

При создании или развитии информационных систем организациям рекомендуется организовать выполнение следующих процессов управления данными в рамках национальной системы управления данными:

#### **1) Управление архитектурой данных:**

В случае, если при формировании концептуальной модели данных и создании сущностей, которые будет обрабатывать информационная система, необходимо использовать единые общепринятые сущности на государственном уровне, определенные для всех доменов деятельности одинаково - необходимо использовать сущности государственной модели данных из ФГИС ЕИП НСУД.

Государственная концептуальная модель данных должна объединять в себе концептуальные модели данных разных доменов деятельности.

Решение об отнесении сущности к типу общедоступной сущности на государственном уровне и последующем использовании ее во всех доменах деятельности принимают сотрудники Коллегии по управлению данными совместно с сотрудниками центра компетенции НСУД, обеспечивающими эксплуатацию ФГИС ЕИП НСУД.

#### **2) Управление потоками и интеграцией данных:**

При настройке передачи данных между информационными системами разных организаций рекомендуется использовать систему межведомственного электронного взаимодействия (СМЭВ).

#### **3) Управление основными данными и НСИ:**

При реализации мероприятий по управлению справочными данными, в случае выявления потребности в использовании общедоступных, общероссийских справочников - необходимо использовать справочники, содержащиеся в ФГИС ЕСНСИ.

При необходимости регулярного обновления данных справочников в информационной системе организации – необходимо обеспечить интеграцию информационной системы с ФГИС ЕСНСИ посредством СМЭВ.

#### **4) Управление метаданными:**

При организации межведомственного взаимодействия между информационными системами различных организаций необходимо обеспечить описание витрин данных НСУД в ФГИС ЕИП НСУД с последующей регистрацией их в СМЭВ.

При описании витрин данных НСУД в ФГИС ЕИП НСУД необходимо обеспечить привязку их атрибутов к сущностям, созданным в рамках процесса управления архитектурой данных.

При выполнении мероприятий по описанию общепринятых сущностей на уровне национальной системы управления данными необходимо обеспечить создание и использование единой согласованной терминологии для всех доменов деятельности.

При использовании организациями сервиса ГосТех «Сервис каталога данных» необходимо обеспечить согласованность показателей в части перечня и описания показателей с репозиторием показателей ФГИС ЕИП НСУД.

#### **5) Управление качеством данных:**

При использовании межведомственного взаимодействия между информационными системами различных организаций посредством обращения к витринам данных НСУД - необходимо настроить проверки качества данных на витринах данных НСУД в ФГИС ЕИП НСУД в соответствии с требованиями потребителей данных.

## 6. ИНСТРУМЕНТЫ УПРАВЛЕНИЯ ДАННЫМИ

Для реализации процессов управления данными на платформе «ГосТех» необходимо использовать компоненты базовых сервисов платформы «ГосТех». На Рис. 4 представлено общее описание доступных базовых сервисов платформы «ГосТех» и смежных решений, необходимых для реализации процессов управления данными.

Для реализации процессов управления данными на платформе «ГосТех» предполагается использование 8 сервисов управления данными<sup>2</sup>:

- Сервисы базового набора «ГосТех» (сервисы работы с данными, интеграционные сервисы, сервисы управления, служебные технологические сервисы, сервисы интеграции с инфраструктурой электронного правительства)<sup>3</sup>;

- Сервис ГосТех «Сервис управления НСИ»;

- Сервис ГосТех «Сервис контроля качества данных»;

- Сервис ГосТех «Сервис каталога данных»;

- Сервис ГосТех «Сервис управления потоками данных и извлечения, преобразования и загрузки данных»;

- Сервис ГосТех «Сервис визуализации данных»;

- Сервис ГосТех «Сервис управление архитектурой данных»;

Сервисы базового набора «ГосТех» доступны для реализации процессов по управлению данными (например, сервис работы с данными – «Сервис СУБД аналитического хранилища данных»). Прочие сервисы, перечисленные выше, находятся на стадии формулирования функционально-технических требований / разработки и будут доступные на платформе «ГосТех» не ранее 2024–2025 года.

Для реализации процессов управления метаданными и качеством данных при межведомственном взаимодействии посредством СМЭВ необходимо использовать решение ФГИС ЕИП НСУД, в части публикация метаданных витрины данных НСУД в общем хранилище метаданных ФГИС ЕИП НСУД, и в части контроля качества данных на витринах данных НСУД.

Компоненты ЕЦП «ГосТех» используются только при создании и развитии ГИС в периметре ОГВ.

Для реализации процессов управления справочными данными необходимо использовать ФГИС ЕСНСИ в части использования общедоступных, общероссийских справочников в ГИС ОГВ.

---

<sup>2</sup> Целевое состояние в части управления данными на платформе «ГосТех».

<sup>3</sup> Источник: [https://platform.gov.ru/wp-content/uploads/2022/12/МЕТОДРЕКОМЕНДАЦИИ\\_БАЗОВЫЕ\\_СЕРВИСЫ\\_ЕЦП\\_ГОСТЕХ-1.pdf](https://platform.gov.ru/wp-content/uploads/2022/12/МЕТОДРЕКОМЕНДАЦИИ_БАЗОВЫЕ_СЕРВИСЫ_ЕЦП_ГОСТЕХ-1.pdf)

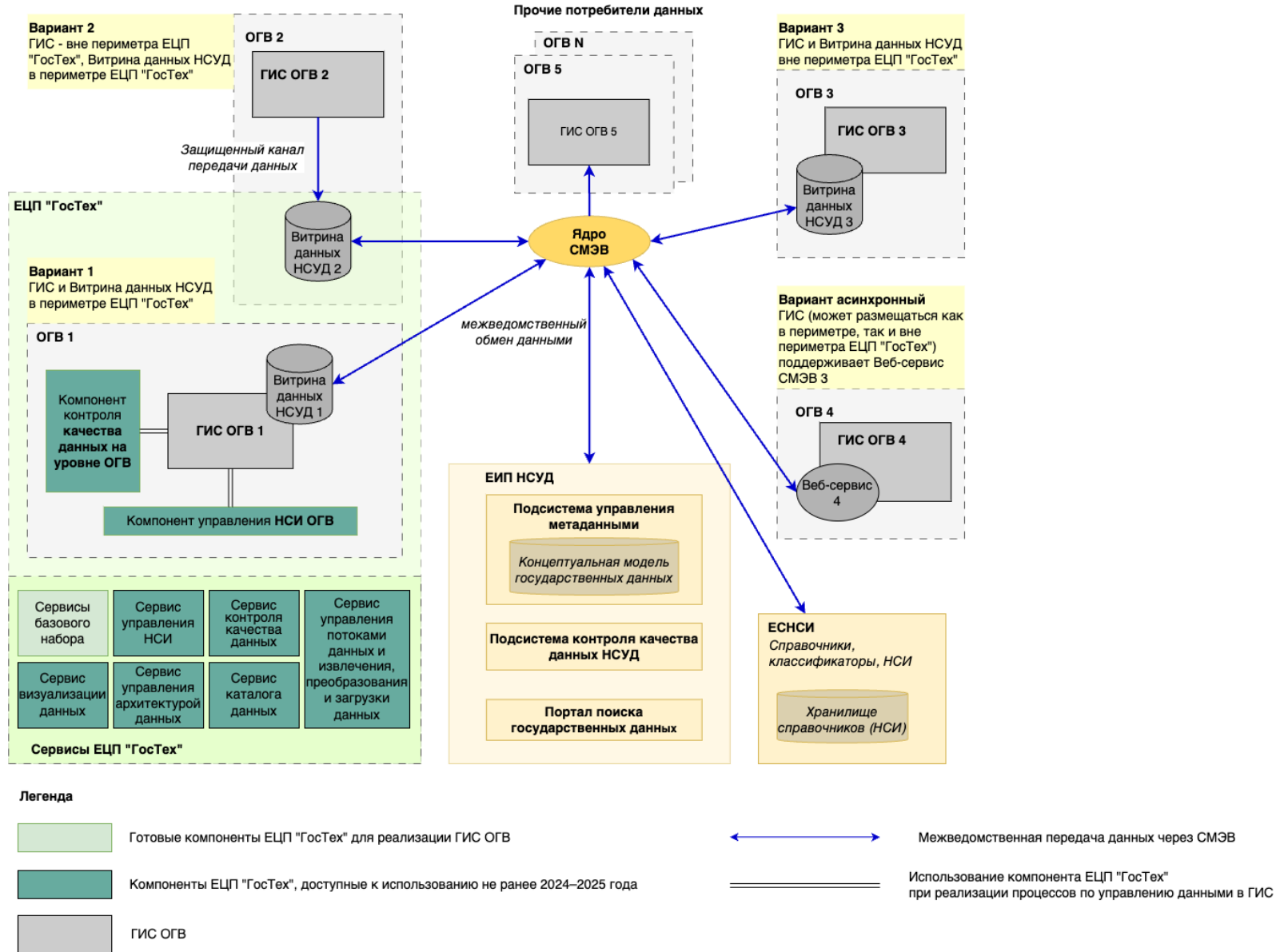


Рис. 4. Общая типовая схема взаимодействия информационных систем и инструментов управления данными.

### III. ПРОЦЕССЫ УПРАВЛЕНИЯ ДАННЫМИ

#### 7. ПРОЦЕСС «УПРАВЛЕНИЕ ЖИЗНЕННЫМ ЦИКЛОМ ДАННЫХ»

##### 7.1. ВВЕДЕНИЕ

###### Назначение и цели процесса

Данные являются одним из ключевых активов организации, для эффективного управления которыми необходимо понимание и планирование их жизненного цикла – от этапа создания данных до использования данных и их последующего удаления.

Целью управления жизненным циклом данных является обеспечение необходимого уровня доступности, качества и безопасности данных, достигаемых за счет:

- стандартизованного описания состава обрабатываемых данных;
- учета источников и потребителей данных;
- осуществления контроля качества данных на этапах жизненного цикла, связанных с созданием (сбором), использованием и предоставлением данных;
- регламентирования хранения / архивирования данных;
- отслеживания актуальности нормативных оснований для обработки данных.

###### Основные понятия

**Жизненный цикл данных** – совокупность мероприятий или этапов, от создания и получения данных до их архивирования и удаления из информационных систем.

Таблица 3. Основные этапы жизненного цикла данных.

НАИМЕНОВАНИЕ ЭТАПА	ОПИСАНИЕ ЭТАПА ИЛИ МЕРОПРИЯТИЯ
Создание и получение данных	Данные поступают / формируются в информационной системе. На данном этапе важно снизить риски получения и создания некачественных данных в рамках процесса <b>обеспечения качества данных</b> .
Трансформация данных	Подготовка данных для дальнейшего хранения / передачи / анализа данных. На этом этапе выполняются необходимые преобразования данных, включая стандартизацию и иные изменения, связанные с повышением качества данных.
Хранение данных	Хранение данных в ИС с учетом нормативных требований к исторической глубине данных и объему хранимых данных.
Анализ и представление данных	Применение подготовленных данных для принятия управленческих решений, выполнения функций или оказания государственных услуг.
Передача данных	Предоставление подготовленных данных для внешних потребителей.
Архивирование и удаление данных	Оптимизация объемов хранимых данных. Хранение слишком большого количества данных увеличивает стоимость управления и владения данными, поэтому неактуальные, невостребованные и некорректные данные с учетом нормативных требований архивируются, удаляются или уничтожаются.

## 7.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ

В рамках текущего процесса задействованы следующие роли:

- **Архитектор данных:** Проектирование архитектуры данных и потоков данных.
- **Инженер данных:** Реализация задач по обработке данных.
- **Аналитик данных:** Определение требований к источникам и составу данных, алгоритмам трансформации данных.
- **Эксперт по качеству данных:** Формирование требований к качеству данных на этапах их создания / сбора, трансформации и передачи. Мониторинг и разрешение инцидентов качества данных.
- **Эксперт по информационной безопасности:** Контроль предоставления доступов к данным, контроль актуальности нормативных оснований для обработки данных.

## 7.3. ЦЕЛЕВОЙ ПРОЦЕСС

Целевой процесс управления жизненным циклом данных выполняется на уровне организации и состоит из следующих мероприятий:

- 1) определение мероприятий, необходимых для реализации управления жизненным циклом данных;
- 2) определение обрабатываемых данных и требований к инфраструктуре для их обработки;
- 3) управление созданием / получением данных;
- 4) определение потребителей данных;
- 5) управление трансформацией, хранением, использованием данных;
- 6) управление передачей данных;
- 7) управление архивацией и удалением данных.

### 1. Определение мероприятий, необходимых для реализации управления жизненным циклом данных

*Описание мероприятия*

В рамках мероприятия необходимо:

- 1) Для ИС на стадии разработки, а также для ИС на стадии развития, предполагающего существенные изменения в составе и объеме обрабатываемых данных, выполнить мероприятия **2 – 7** текущего процесса.
- 2) При появлении новых потребителей данных выполнить мероприятие **6** текущего процесса.
- 3) При возникновении необходимости изменить алгоритмы трансформации данных или условия / механизмы хранения данных выполнить мероприятие **5** текущего процесса.

*Исходная информация*

- НПА, регламентирующие деятельность организации;
- Результаты мониторинга функционирования ИС организации.

*Результат выполнения мероприятия*

Улучшены значения показателей эффективности процесса (раздел.7.4).

*Рекомендуемые инструменты для реализации*

Приведены в описаниях соответствующих мероприятий.

### 2. Определение обрабатываемых данных и требований к инфраструктуре для их обработки

*Описание мероприятия:*

В рамках мероприятия необходимо выполнить следующие задачи:

- определить периметр моделирования и проектирования данных – ИС организации и/или хранилище данных организации;
- определить данные, которые будут использоваться в организации;

- определить категории создаваемых данных (основные данные, справочные данные, транзакционные данные, служебные данные);
- определить формат и структуру создаваемых данных (структурированные, полуструктурированные и неструктурированные данные);
- описать создаваемые данные в соответствии с **процессом управления метаданными и процессом управления основными данными и НСИ**;
- определить потенциальный объем хранимых данных за период эксплуатации ИС и/или хранилища данных организации;
- сформировать требования к инфраструктуре обработки данных в соответствии с процессом управления архитектурой данных.

*Исходная информация*

- НПА, регламентирующие деятельность организации;
- Требования потребителей данных;
- Проектная / эксплуатационная документация на ИС организации.

*Результат выполнения мероприятия*

- Информация о данных, обрабатываемых в организации, загружена в **хранилище метаданных**;
- Подготовлены требования к инфраструктуре обработки данных.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных».

### 3. Управление созданием и сбором данных

*Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи:

- определить организации и ИС, которые являются источниками данных;
- при отсутствии интеграции с источником данных реализовать интеграции в соответствии с **процессом управления интеграцией данных**;
- при отсутствии доступа к данным, предоставляемым внешней организацией, запросить доступ в соответствии с **процессом обеспечения безопасности данных**;
- описать процесс, в рамках которого выполняется создание / сбор данных в соответствии с **процессом управления архитектурой данных**;
- добавить в описание данных сведения об источниках данных в соответствии с **процессом управления метаданными**;
- определить владельца данных для создаваемых данных;
- обеспечить контроль качества данных на этапе создания / сбора первичных данных в соответствии с **процессом управления качеством данных**.

*Исходная информация*

- НПА, регламентирующие деятельность организации;
- эксплуатационная документация на ИС организации;
- соглашения об обмене данными с внешними организациями.

*Результат выполнения мероприятия*

- определены источники данных;
- реализована интеграция с источниками данных, получен доступ к данным, предоставляемым внешней организацией;
- описан процесс создания / сбора данных;
- определен Куратор данных, он наделен соответствующими правами в ЕИП НСУД;
- обеспечен контроль качества данных на этапе создания / сбора первичных данных.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»;

- ЕИП НСУД.

#### 4. Определение потребителей данных

##### *Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи:

- определить внутренних и внешних потребителей (категории потребителей) данных и цели использования данных, внести информацию о потребителях данных в хранилище метаданных;
- определить и зафиксировать SLA на использование данных и доступность данных с внутренними потребителями данных;
- сформировать ролевую модель доступа к данным для пользователей ИС организации и критерии доступа к данным для внешних потребителей в соответствии с **процессом обеспечения безопасности данных**.

##### *Исходная информация*

- НПА, регламентирующие деятельность организации;
- Реестр сервисов / услуг, функционирующих на основе предоставляемых данных;
- Требования к доступности и качеству данных от потребителей данных.

##### *Результат выполнения мероприятия*

- Информация о потребителях данных внесена в хранилище метаданных;
- Зафиксированы SLA на доступность и качество данных с внутренними потребителями данных;
- Сформированы: ролевая модель доступа к данным для пользователей ИС организации и критерии доступа к данным для внешних потребителей.

##### *Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных».

#### 5. Управление трансформацией, хранением, использованием данных

##### *Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи:

- разработать алгоритмы трансформации данных, необходимые для их предоставления потребителям, в соответствии с **процессом управления архитектурой данных**;
- реализовать меры, обеспечивающие соблюдение регуляторных требований и отраслевых стандартов, регламентирующих хранение данных;
- определить срок доступности данных и способ хранения данных;
- определить требования к резервному копированию данных;
- реализовать алгоритмы трансформации данных и требования по хранению данных в соответствии с **процессом управления архитектурой данных**;
- реализовать контроль качества данных по результатам трансформации данных в соответствии с **процессом управления качеством данных**;
- предоставить пользователям ИС организации доступ к данным согласно ролевой модели в соответствии с процессом обеспечения безопасности данных.

##### *Исходная информация*

- результаты выполнения предыдущего мероприятия.
- НПА, регламентирующие деятельность организации;
- регуляторные требования и отраслевые стандарты, регламентирующие хранение данных.

##### *Результат выполнения мероприятия*

- осуществляется трансформация данных в соответствии с реализованными алгоритмами;



- осуществляется хранение данных в соответствии с регуляторными требованиями и отраслевыми стандартами, регламентирующими хранение данных, а также с требованиями к резервному копированию данных;
- осуществляется контроль качества данных по результатам их трансформации в соответствии с SLA по качеству и доступности данных для потребителей данных;
- пользователям ИС организации предоставлен доступ к данным в соответствии с ролевой моделью.

*Рекомендуемые инструменты для реализации*

- ГосТех. «Сервис управления потоками данных и извлечения, преобразования и загрузки данных».

## **6. Управление передачей данных**

*Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи:

- при отсутствии интеграции с ИС внешних потребителей данных реализовать интеграцию в соответствии с **процессом управления интеграцией данных**;
- обеспечить требуемое качество данных, предоставляемых внешним потребителям в соответствии с **процессом управления качеством данных**.

*Исходная информация*

Результаты выполнения мероприятий **4** и **5** текущего процесса.

*Результат выполнения мероприятия*

Осуществляется предоставление данных внешним потребителям в соответствии с критериями доступа и с требуемым качеством данных.

*Рекомендуемые инструменты для реализации*

- Личный кабинет участника электронного взаимодействия (ЛК ЭВ).

## **7. Управление архивацией и удалением данных**

*Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи:

- оценить затраты на хранение / архивацию данных;
- определить требования к архивации (разархивации) и/или удалению данных с учетом нормативных и регуляторных требований, а также затрат на хранение / архивацию данных;
- разработать регламенты архивирования и/или удаления данных, включая ведение учета архивов;
- реализовать требования к архивации (разархивации) и/или удалению данных в соответствии с процессом управления жизненным циклом данных.

*Исходная информация*

- НПА, регламентирующие деятельность организации;
- регуляторные требования и отраслевые стандарты, регламентирующие хранение данных.

*Результат выполнения мероприятия*

Осуществляется архивация / удаление данных в соответствии с установленными требованиями.

*Рекомендуемые инструменты для реализации*

- отсутствуют

#### 7.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА

Для оценки эффективности процесса рекомендуется применять следующие показатели:

- доля описаний показателей, сведения о которых внесены в **хранилище метаданных организации** (целевое значение: 100 %);
- доля подлежащих контролю качества данных, для которых выполняется контроль качества (целевое значение: 100 %).

#### 7.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА

Для оценки уровня зрелости процесса рекомендуется применять шкалу оценок, указанную в таблице 4.

Таблица 4. Описание уровня зрелости процесса

Критерии	Уровни зрелости процесса			
	1	2	3	4
Описание ключевых данных загружено в хранилище метаданных организации.		✓	✓	✓
Осуществляется регулярный контроль качества данных на этапе передачи данных.		✓	✓	✓
Описание всех данных загружено в хранилище метаданных организации.			✓	✓
Разработаны и используются регламенты архивирования и/или удаления данных.			✓	✓
Осуществляется регулярный контроль качества данных на этапе создания / сбора данных.			✓	✓
Осуществляется регулярный контроль качества данных на этапе обработки и миграции данных.				✓
Выполняется постоянный контроль действия НПА, регламентирующих доступ к данным.				✓

## 8. ПРОЦЕСС «УПРАВЛЕНИЕ АРХИТЕКТУРОЙ ДАННЫХ»

### 8.1. ВВЕДЕНИЕ

#### Назначение и цели процесса:

**Управление архитектурой данных** – совокупность мероприятий по определению потребностей в данных, проектированию и организации способов обработки и хранения данных для удовлетворения требований потребителей данных.

Ключевая цель процесса – определение потребностей организации в данных, формирование требований к хранению и обработке данных, и создание такой архитектуры обработки и хранения данных, которая позволит:

- обеспечить выполнение процессов, заложенных в информационную систему, включая выполнение аналитических задач организации;
- приносить организации ценность за счет расширения возможностей по использованию данных;
- обеспечить возможность внесения изменений в структуру данных и поддерживать рост объема обрабатываемых данных при минимальных затратах;
- организовать эффективное выполнение операций при обработке данных;
- обеспечить доступность данных на протяжении всего их жизненного цикла;
- обеспечить возможность управлять качеством данных на требуемых этапах обработки данных;
- приносить организации ценность за счет оптимизации требуемых ресурсов на реализацию архитектуры хранения и обработки данных.

#### Основные понятия

Архитектура данных информационной системы (ИС) зависит от **архитектуры ИС**. В целях настоящего документа ИС рассматриваются с точки зрения<sup>4</sup> их **назначения**:

- **Учетная ИС** обеспечивает сбор, регистрацию, обработку и представление учитываемых данных. Как правило, учетные системы реализуются для управления основными данными.
- **Транзакционная ИС** обеспечивает в автоматическом режиме учет, хранение и оперативную обработку непрерывно поступающих данных.
- **Аналитическая ИС** обеспечивает сбор, хранение, структурирование и обобщение (агрегацию) информации, а также её представление пользователю в виде, удобном для принятия решений.
- **Исследовательская ИС** предназначена для получения уникальных результатов обработки больших объемов слабо структурированной информации. Как правило, исследовательские системы применяются в проектах машинного обучения, ИИ, экспертных систем и т. п.
- **Служебная ИС** не реализует собственных бизнес-функций, но обеспечивает функционирование иных ИС организации.

ИС уровня организации, как правило, являются **комплексными**, так как содержат в составе отдельные **подсистемы**, каждая из которых имеет собственное назначение из числа приведенных выше. ИС, не имеющие в составе подсистем, называются **монолитными**.

На технологическую архитектуру ИС и архитектуру данных существенно влияют также требования по безопасности и защищенности данных. Например, для их удовлетворения может потребоваться физическое разграничение данных в СУБД, содержащие данные разных классов защищенности.

Данные, обрабатываемые ИС, хранятся в структурах, называемых **Базами данных (БД)**. Доступ к БД осуществляется помощью **системы управления базами данных (СУБД)**.

<sup>4</sup> См. ГОСТ Р 57100—2016/ISO/IEC/IEEE 42010:2011 «Описание архитектуры»

В зависимости от требований архитектуры, БД и соответствующие СУБД могут существенно различаться. В ГИС часто применяются следующие типы архитектур БД:

- **Реляционная БД** – наиболее универсальная и распространенная архитектура, ориентированная на организацию данных в виде связанных записей и атрибутов в двумерных таблицах. Используется для любых видов ИС.
- **Колоночная БД** – архитектура, отличающаяся от реляционных БД хранением записей не построчно, а по колоночно, что усложняет вставку записей в такую БД, но существенно ускоряет аналитические выборки. Преимущественно используется для Аналитических и Исследовательских ИС.
- **БД ключ-значение** – архитектура, ориентированная на возможно более быстрое получение данных по заданному ключу. Преимущественно используется для Транзакционных и Учетных ИС.
- **Документно-ориентированная БД** – архитектура, предназначенная для хранения больших объемов слабо структурированной информации. Преимущественно используется для Учетных ИС.
- **БД временных рядов** – архитектура предназначена для хранения упорядоченных по времени данных, имеющих метки времени. Преимущественно используется для Транзакционных ИС.
- **Графовая БД** - архитектура, ориентированная на хранение связей между данными, что упрощает постановку и решение аналитических и экспертные задач. Преимущественно используется для Аналитических ИС.
- **Многомодельная БД** – объединяет в одной СУБД несколько параллельных архитектур хранения данных. Может быть использована для всех видов ИС.

## Типы моделей данных

**Модель данных организации состоит из следующих компонентов:**

- **Концептуальная модель данных организации (КМД)** – целостное, не зависящее от технических средств реализации, высокоуровневое представление данных, обрабатываемых в организации. КМД организации включает данные о ключевых сущностях и связях между ними.
- **Логические модели данных (ЛМД)** отдельных предметных областей, реализуемых ИС организации – представление сущностей, их атрибутов и связей.
- **Физические модели данных ИС организации** – детальное описание организации данных в конкретных архитектурах БД ИС.

Примеры уровня детализации описания концептуальной, логической и физической моделей данных для реляционных БД организации приведены на Рис. 5.

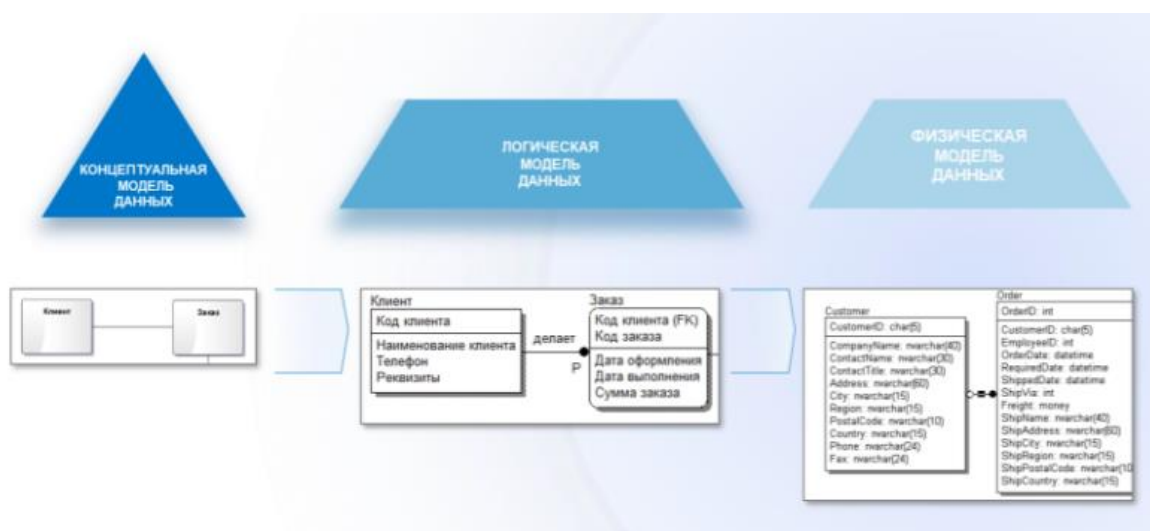
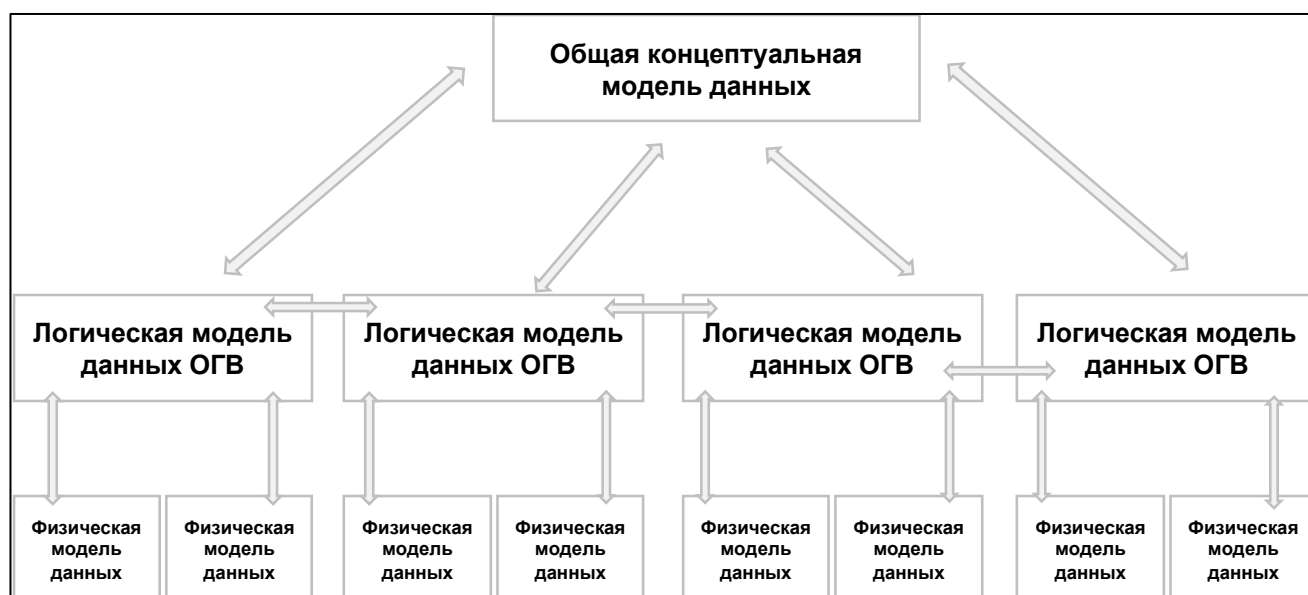


Рис. 5. Пример уровня детализации описания модели на концептуальном логическом и физическом уровне

Указанные модели должны соответствовать **государственной модели данных** как показано на Рис. 6.



*Рис. 6. Единая государственная модель данных (концептуальная, логическая, физическая)*

На самом верхнем уровне представлена **концептуальная модель государственных данных**, ведение которой осуществляется в ЕИП НСУД, на втором уровне – концептуальные модели данных организации, входящих в Домен, на третьем уровне – физические модели данных ГИС.

**Поток данных** – передача информации между ИС в автоматическом режиме в машиночитаемом формате.

## 8.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ

В рамках текущего процесса задействованы следующие роли:

- **Директор по данным:** утверждает приоритеты документирования архитектуры данных, описания текущей и целевой архитектуры данных, результаты оценки архитектуры данных, дорожные карты проектов по изменению архитектуры данных, целевые значения показателей.
- **Архитектор данных:** обеспечивает документирование и сопровождение архитектуры данных, подготовку целевой архитектуры данных.
- **Эксперты по данным:** участвуют в документировании архитектуры данных, подготовке целевой архитектуры данных, оценке текущей архитектуры данных, приемке результатов реализации целевой архитектуры данных.
- **Куратор данных:** участвуют в документировании архитектуры данных, подготовке целевой архитектуры данных, оценке текущей архитектуры данных, приемке результатов реализации целевой архитектуры данных.

## 8.3. ЦЕЛЕВОЙ ПРОЦЕСС

Целевой процесс управления архитектурой данных государственных информационных систем состоит из 6 мероприятий, приведенных ниже:

- 1) Анализ существующей ИТ-архитектуры и архитектуры данных организации;
- 2) Документирование существующей архитектуры данных, создание модели данных организации;
- 3) Оценка текущей архитектуры;
- 4) Разработка и документирование целевой архитектуры данных, планирование перехода к целевой архитектуре данных;
- 5) Реализация целевой архитектуры данных;
- 6) Сопровождение архитектуры данных.

Мероприятия 1 и 2 выполняются однократно, мероприятия 3-6 на регулярной основе.

### 1. Анализ существующей ИТ-архитектуры и архитектуры данных организации

#### *Описание мероприятия*

В рамках мероприятия необходимо:

1. Сформировать реестр информационных систем организации, указав для каждой ИС:
  - Наименование;
  - Контакты лица, ответственного за создание и поддержку сведений об ИС;
  - Код ИС в ФГИС «Координация информатизации»;
  - НПА, регламентирующие создание, ввод в эксплуатацию и эксплуатацию ИС;
  - Назначение ИС (транзакционная, учетная, аналитическая, исследовательская, служебная<sup>5</sup>);
  - Физические модели данных, используемые в ИС, и наименования соответствующих СУБД;
  - Для комплексных ИС – названия и назначения подсистем, входящих в состав таких ИС;
  - Основные бизнес-объекты, сведения о которых обрабатываются в ИС, включая виды основных данных, ведение которых обеспечивает ИС.
2. Сформировать реестр нормативных требований к данным организации, включая:
  - Требования НПА к составу данных, обрабатываемых в ГИС организации;

<sup>5</sup> Поскольку служебные ИС (а также служебные подсистемы, входящие в состав ИС), как правило, не реализуют бизнес-задачи организации, создание моделей данных для таких систем (подсистем) не является обязательным.

- Перечень поставщиков данных в ГИС организации в соответствии с НПА и межведомственными соглашениями;
  - Требования НПА и соглашений к составу данных, поступающих в ГИС от поставщиков данных;
  - Перечень потребителей данных из ГИС организации в соответствии с НПА и межведомственными соглашениями;
  - Требования НПА и соглашений к составу данных, поставляемых из ГИС организации потребителям данных;
  - Требования НПА и соглашений о взаимодействии к срокам и графикам получения, обработки и поставки данных;
  - Требования НПА по ограничению доступа к данным, обрабатываемых в ГИС организации.
3. Сформировать реестр средств информационного взаимодействия, включая:
- Перечень внешних источников данных, включая сведения о получаемых данных и, при наличии, сведения об основаниях получения организацией внешних данных;
  - Перечень внешних потребителей данных, включая сведения о передаваемых данных и, при наличии, сведения об основаниях получения потребителями данных организации;
  - Перечень средств информационного взаимодействия (виды сведений СМЭВ-3, витрины НСУД) в разрезе получаемых и передаваемых данных.
4. Проверить наличие в документации описаний концептуальных, логических и физических моделей данных и диаграмм потоков данных для всех ИС организации.
5. Проанализировать соглашения об информационном взаимодействии, внутренние нормативные акты, документы, содержащие описание ИТ-архитектуры организации и архитектур ИС организации на предмет полноты и актуальности. Зафиксировать все обнаруженные несоответствия.

#### *Исходная информация*

- Результаты выполнения процесса «Управление жизненным циклом данных»;
- НПА, касающиеся данных, обрабатываемых в ГИС организации;
- Межведомственные соглашения об информационном взаимодействии;
- Описание текущей архитектуры хранилища данных организации (при наличии);
- Внутренние нормативные акты организации, касающиеся порядка получения, обработки, передачи и/или публикации информации;
- Описание ИТ-архитектуры организации и архитектур ИС организации;
- Техническая документация существующих ИС организации.

#### *Результат выполнения мероприятия*

- Реестр информационных систем организации (п.1);
- Реестр нормативных требований к данным организации (п.2);
- Реестр средств информационного взаимодействия (п.3);
- Сведения о документах, содержащих описания концептуальных, логических и физических моделей данных и диаграмм потоков данных для каждой ИС организации (п.4);
- Сведения о несоответствиях фактической ИТ-архитектуры организации нормативным актам, имеющейся архитектурной и технологической документации (при наличии).

#### *Рекомендуемые инструменты для реализации*

Результаты работ размещаются в:

- ГосТех «Сервис каталога данных»;
- ЕИП НСУД в части реестра средств информационного взаимодействия.

## **2. Документирование существующей архитектуры данных**

### *Описание мероприятия*

В рамках мероприятия необходимо:

1. Определить методы моделирования, которые будут применяться для первичного описания архитектуры организации в зависимости приоритетов организации и типов ИС, принимая во внимание что:

- для всех ИС необходимо создать концептуальную и логическую модели данных. Данные всех ИС организации/Домена должны быть объединены единой логической моделью данных организации/Домена с привязкой физических моделей данных к единой логической модели данных организации/домена;
- для всех ИС, обладающих модульной архитектурой, необходимо создать схему подсистем и/или модулей и диаграмму потоков данных между подсистемами/модулями;
- для всех ИС, являющихся поставщиками данных для внешних потребителей необходимо сформулировать описания интерфейсов доставки данных: описания витрин данных НСУД и регламентированных запросов СМЭВ-4, видов сведений СМЭВ-3, способов и форматов публикации открытых данных;
- для транзакционных и учетных ИС рекомендуется создать схемы потоков данных внутри основных модулей;
- для аналитических ИС необходимо указать основные аналитические представления (отчеты, дашборды), формируемые указанными ИС и данные;
- физические модели данных рекомендуется создать для всех ИС, в технической документации которых отсутствуют актуальные описания физических моделей данных.

2. Сформировать план-график документирования данных ИС организации в зависимости от важности данных, обрабатываемых в ИС, для выполнения бизнес-процессов организации и достижения стратегических целей.

3. При отсутствии – сформировать, при наличии – привести в соответствие концептуальную модель данных организации концептуальной модели государственных данных и моделями данных доменов, относящихся к организации.

4. Сформировать концептуальную модель данных, соответствующую глоссарию государственных данных и глоссарию данных доменов деятельности. При необходимости сформировать предложения по расширению государственного глоссария данных и глоссария данных домена деятельности.

5. При отсутствии – сформировать, при наличии – привести в соответствие концептуальной модели данных и глоссарию данных логические модели данных ИС организации.

6. При отсутствии – сформировать диаграммы потоков данных между подсистемами каждой ИС организации, а также между ИС организаций.

7. При отсутствии – сформировать диаграммы потоков данных внутри ИС (подсистем ИС) организации.

8. Проанализировать описания средств информационного взаимодействия (видов сведений, витрин данных) на предмет наличия, полноты, корректности, актуальности и соответствия логическим моделям данных. Связать элементы описания средств информационного взаимодействия с логическими моделями данных организации в сервисе ГосТех «Каталог данных». Создать недостающие и доработать несоответствующие описания средств информационного взаимодействия.

9. Проанализировать техническую документацию к ИС на предмет наличия, полноты, корректности и актуальности описания физических моделей данных. Связать элементы физических моделей данных с логическими моделями данных. Создать недостающие физические модели данных, доработать некорректные и неактуальные модели.

*Исходная информация*

- Результаты мероприятия 1;
- Модель государственных данных;



- Концептуальные модели данных доменов деятельности организации;
- Техническая документация существующих ИС организации.

*Результат выполнения мероприятия*

- Концептуальная модель данных организации, синхронизированная с ЕИП НСУД и КМД домена деятельности;
- Логические модели данных ИС организации;
- Диаграммы потоков данных между ИС организаций, а также внутренних потоков данных внутри ИС организации;
- Физические модели данных ИС организации;
- Логическая модель данных организации/ Домена сопоставленная с физической моделью данных организации;
- «Технический долг» по описанию архитектуры данных организации – перечень работ по документированию архитектуры данных организации, в соответствии с приоритетами запланированных на сроки в будущем.

*Рекомендуемые инструменты для реализации*

Работы выполняются в ПО, предназначенном для моделирования данных:

- ГосТех «Сервис каталога данных» и «Сервис для управления архитектурой данных».
- В ФГИС ЕИП НСУД в части моделей данных информационных систем и описания средств информационного взаимодействия, применяемых для межведомственного информационного взаимодействия.

### **3. Оценка текущей архитектуры**

*Описание мероприятия*

В рамках мероприятия не реже одного раза в год проводится анализ жалоб пользователей и потребителей, бизнес-планов организации, прогнозов изменений объемов данных и состава потребителей.

По результатам анализа выявляются проблемы архитектуры данных, новые возможности по её развитию, а также принимаются решения о целесообразности и приоритетах изменения моделей данных, создании новых либо деактивации неактуальных баз данных, выводе из эксплуатации сервисов, витрин данных, аналитических представлений данных, не имеющих потребителей.

Принятые решения не ограничиваются изменением документов, описывающих архитектуру данных и требуют соответствующих изменений в ИС организации. Например, при проблемах с оперативностью и качеством подготовки аналитических отчетов в учетной системе может потребоваться создание отдельной аналитической подсистемы; при проблемах со скоростью обработки информации может потребоваться разделение потоков данных на синхронную и асинхронную составляющую.

*Исходная информация*

- Результаты мероприятия 2;
- Обращения потребителей данных о нарушении требований по полноте, качеству и доступности данных;
- Журналы загрузки и мониторинга нагрузки на базы данных;
- Планы по развитию ИС организации, расширению исполняемых функций, обрабатываемых данных;
- Нормативные и методические документы, касающиеся обработки данных, опубликованные с момента предыдущего мероприятия по оценке архитектуры данных.

*Результат выполнения мероприятия*

- Описание выявленных проблем и несоответствий архитектуры данных организации;
- Перечень принятых решений по изменению архитектуры данных организации, включая:

- Приоритет изменения;
- Измеримые показатели, на достижение которых направлено изменение;
- Перечень доработок ИС, которые потребуются для реализации изменения (при наличии).

*Рекомендуемые инструменты для реализации*

- отсутствуют

#### **4. Разработка и документирование целевой архитектуры данных, планирование перехода к целевой архитектуре данных**

*Описание мероприятия*

Необходимо разработать целевую архитектуру данных организации на основе существующей архитектуры и принятых решений по её изменению, в том числе:

- подготовить и согласовать дорожные карты проектов по изменению архитектуры данных организации, включая:

- состав работ;
- ответственных за выполнение работ;
- согласованные с владельцем ИС организации планы по изменению организации, в т.ч. в результате проектирования доменов деятельности (согласно Методическим рекомендациям по проектированию и утверждению целевой архитектуры домена с использованием единой цифровой платформы «ГосТех»);
- целевые сроки проверки результативности изменений: достижения требуемых показателей.

Для пользователей платформы «ГосТех» необходимо разработать целевую архитектуру данных и согласовать ее с оператором платформы «ГосТех». Целевая архитектура данных формируется с учетом практик построения Хранилища данных, использования сервисов «ГосТех» и матрицы распределения ответственности (Приложение 1, 2 и 3).

*Исходная информация*

- Результаты мероприятий 2 и 3;
- Обращения потребителей данных о нарушении требований по полноте, качеству и доступности данных;
- Журналы загрузки и мониторинга нагрузки на базы данных;
- Планы по развитию ИС организации, расширению исполняемых функций, обрабатываемых данных;
- Нормативные и методические документы, касающиеся обработки данных, опубликованные с момента предыдущего мероприятия по оценке архитектуры данных.

*Результат выполнения мероприятия*

- Описание целевой архитектуры данных организации, направленной на устранение выявленных проблем и несоответствий;
- Дорожные карты проектов по изменению архитектуры данных организации, согласованные с руководством организации.

*Рекомендуемые инструменты для реализации*

- Работы по разработке целевой архитектуры данных выполняются в ПО, предназначенном для моделирования данных;
- ГосТех «Сервис для управления архитектурой данных»;
- В ФГИС ЕИП НСУД в части моделей данных информационных систем и описания средств информационного взаимодействия, применяемых для межведомственного информационного взаимодействия.

## 5. Реализация целевой архитектуры данных

### *Описание мероприятия*

В рамках мероприятия необходимо:

- выполнить проекты по изменению архитектуры данных организации в соответствии с согласованными дорожными картами;
- выполнить проверки результативности изменений: достижения требуемых показателей;
- внести изменения в документы, описывающие текущую архитектуру данных организации в соответствии с реализованной целевой архитектурой.

### *Исходная информация*

- Результаты мероприятий 2 и 4;
- Подтвержденные значения показателей, полученные по результатам проведенных изменений архитектуры данных организации;
- Актуальные документы по архитектуре данных и функциональным компонентам ГИС.

### *Рекомендуемые инструменты для реализации*

- Работы по разработке целевой архитектуры данных выполняются в ПО, предназначенном для моделирования данных;
- ГосТех «Сервис для управления архитектурой данных»;
- В ФГИС ЕИП НСУД в части моделей данных информационных систем и описания средств информационного взаимодействия, применяемых для межведомственного информационного взаимодействия.

## 6. Сопровождение архитектуры данных

### *Описание мероприятия*

В рамках мероприятия необходимо:

- Поддерживать описания архитектуры данных в актуальном состоянии. Документация по архитектуре данных организации приведена в соответствие с изменениями, проведенными в ИС организации в срок, не превышающий одну рабочую неделю после выполнения таких изменений.

### *Исходная информация*

- Результаты мероприятия 2;
- Описания работ по изменениям ИС организации.

### *Результат выполнения мероприятия*

- Описания архитектуры данных организации своевременно приведены в соответствие с изменениями в ИС организации.

### *Рекомендуемые инструменты для реализации*

- Работы по разработке целевой архитектуры данных выполняются в ПО, предназначенном для моделирования данных;
- ГосТех «Сервис для управления архитектурой данных»;
- В ФГИС ЕИП НСУД в части моделей данных информационных систем и описания средств информационного взаимодействия, применяемых для межведомственного информационного взаимодействия.

## 8.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА

Для оценки эффективности процесса рекомендуется применять следующие показатели:

1. Полнота документирования архитектуры данных в части описания ИС, целевое значение: 100%;
2. Полнота документирования архитектуры данных в части концептуальной модели данных, целевое значение: 100% для всех ИС организации за исключением служебных, 100% для всех объектов, указанных в НПА и соглашениях;
3. Полнота документирования архитектуры данных в части логических моделей данных, целевое значение: 100% для всех ИС организации за исключением служебных, 100% для всех атрибутов всех объектов, указанных в НПА и соглашениях;
4. Полнота документирования архитектуры данных в части потоков данных, целевое значение: 100% для всех ИС организации (включая подсистемы) за исключением служебных;
5. Своевременность поддержки архитектуры данных в актуальном состоянии; целевое значение: 100% изменений всех ИС организации за исключением служебных отражаются в документации архитектуры данных в срок, не превышающий одной рабочей недели;
6. Результативность перехода к целевой архитектуре данных, целевое значение: в 100% достигнуто улучшение значений целевых показателей на не менее чем 75% от запланированных значений.

### 8.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА

Таблица 5. Описание уровня зрелости процесса

Критерии	Уровни зрелости процесса			
	1	2	3	4
Проанализирована и частично (основные критичные ГИС) документирована текущая архитектура данных организации.		✓	✓	✓
Архитектура данных для большинства ИС организации документирована.		✓	✓	✓
Вся существующая архитектура данных документирована (все ГИС).			✓	✓
Архитектура данных организации поддерживается в актуальном состоянии (отсутствуют неактуальные внутренние документы в части архитектуры данных).			✓	✓
Проводится регулярная (не реже раза в год) оценка текущей архитектуры данных, по результатам оценки планируются и реализуются изменения архитектуры данных.				✓
Своевременно погашается «технический долг» по описанию архитектуры данных.				✓
Результаты изменений архитектуры данных по приведению её к целевой архитектуре измеряются и близки к планируемым				✓

## 9. ПРОЦЕСС «УПРАВЛЕНИЕ ИНТЕГРАЦИЕЙ ДАННЫХ»

### 9.1. ВВЕДЕНИЕ

#### Назначение и цели процесса

Интеграция данных — это процесс объединения данных из нескольких разрозненных источников данных для предоставления потребителям данных возможности сквозного анализа данных.

Основными целями управления интеграцией данных являются:

- обеспечение необходимого уровня доступности данных из внутренних и внешних источников;
- минимизация повторного ввода и хранения данных;
- обеспечение оперативности передачи данных между различными системами;
- оптимизация затрат на хранение, обработку и передачу данных;
- упрощение взаимодействия между государственными органами в части обмена данными;
- ускорение процессов принятия решений и повышение эффективности государственной деятельности;
- упрощение взаимодействия между организациями в части обмена данными.

#### Основные понятия

В рамках процесса интеграции данных регулируются потоки данных внутри организации и между организациями. Определение понятия «поток данных» приведено в описании **процесса управления архитектурой**.

### 9.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ

В рамках текущего процесса задействованы следующие роли:

- **Архитектор данных:** Проектирование модели потоков данных.
- **Инженер данных:** Реализация передачи данных внутри каждой ИС организации, между различными ИС организации, а также между ИС организации и внешними ИС. Разработка потоков данных и настройка модели загрузки данных.

### 9.3. ЦЕЛЕВОЙ ПРОЦЕСС

Целевой процесс управления интеграцией данных выполняется на уровне организации и состоит из следующих мероприятий:

- 1) определение мероприятий, необходимых для интеграции данных;
- 2) учет потоков данных;
- 3) проектирование потоков данных;
- 4) реализация потоков передачи данных.

#### 1. Определение мероприятий, необходимых для интеграции данных

##### *Описание мероприятия*

В рамках мероприятия необходимо:

- проверить, внесены ли описания потоков данных внутри каждой ИС организации, между различными ИС организации, а также между ИС организации и внешними ИС в **хранилище метаданных** организации. Если нет, то выполнить мероприятие **2**.
- в случае необходимости добавления / изменения потоков данных внутри ИС организации, между различными ИС организации, а также между ИС организации и внешними ИС выполнить мероприятия **3** и **4**.

##### *Исходная информация*

- текущая модель потоков данных организации, загруженная в **хранилище метаданных**;
- НПА, регламентирующие деятельность организации;

- соглашения об обмене данными с внешними организациями;
- техническая документация на ИС организации;
- сведения о внешних ИС – потребителях и источниках данных;
- описания данных организации на физическом и логическом уровнях;
- требования к качеству и доступности данных источников и потребителей данных организации.

*Результат выполнения мероприятия*

Инициатива(ы) по разработке/доработке/оптимизации потоков данных.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»;
- ГосТех «Сервис управления потоками данных и извлечения, преобразования и загрузки данных».
- ГосТех. Базовые сервисы работы с данными (СУБД);

## 2. Учет потоков данных

*Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи.

- провести анализ текущей модели потоков данных внутри каждой ИС организации, между различными ИС организации, а также между ИС организации и внешними ИС и сформировать описание потоков данных в соответствии с **процессом управления архитектурой**;
- внести описания потоков данных в **хранилище метаданных** организации в соответствии с **процессом управления метаданными**;

*Исходная информация*

- НПА, регламентирующие деятельность организации;
- соглашения об обмене данными с внешними организациями;
- техническая документация на ИС организации;
- сведения о внешних ИС – потребителях и источниках данных;
- описания данных организации на физическом и логическом уровнях;
- требования к качеству данных источников и потребителей данных организации.

*Результат выполнения мероприятия*

- текущая модель потоков данных организации, загруженная в **хранилище метаданных**.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»;
- ГосТех «Сервис управления потоками данных и извлечения, преобразования и загрузки данных»;
- ГосТех. Базовые сервисы работы с данными (СУБД).

## 3. Проектирование потоков данных

*Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи.

- разработать проект целевой модели потоков данных, включающий:
  - источник данных, перечень промежуточных и конечных приемников данных;
  - описание каналов передачи данных (к примеру API, СМЭВ);
  - логическая и физическая модель данных источника;
  - логические и физические модели данных конечных и промежуточных приемников (витрина данных, аналитический отчет);
  - выбранный подход к передаче данных (инкрементальный/полный);
  - расписание передачи данных;

- порядок трансформации данных (ETL/ELT);
- средства и точки мониторинга потоков данных;
- операционную модель и план реагирования на внештатные ситуации ошибки и сбои в ходе промышленной эксплуатации с указанием должностных лиц и потребителей данных.
- сформировать план перехода от текущей модели потоков данных к целевой модели потоков данных в соответствии с утвержденным проектом;
- в случае отсутствия доступа к данным внешнего источника запросить доступ в соответствии с **процессом обеспечения безопасности данных**;
- в случае если конечным приемником потока данных является внешняя ИС, предоставить доступ к данным внешним потребителям согласно критериям доступа к данным в соответствии с **процессом обеспечения безопасности данных**.

Для пользователей платформы «ГосТех» необходимо согласовать целевую модель потоков данных и план перехода к целевой модели с Оператором платформы «ГосТех».

#### *Исходная информация*

- текущая модель потоков данных организации, загруженная в **хранилище метаданных**;
- НПА, регламентирующие деятельность организации;
- соглашения об обмене данными с внешними организациями;
- техническая документация на ИС организации;
- сведения о внешних ИС – потребителях и источниках данных;
- описания данных организации на физическом и логическом уровнях;
- требования к качеству данных источников и потребителей данных организации.

#### *Результат выполнения мероприятия*

- Проект согласованной целевой модели потоков данных;
- План перехода от текущей модели потоков данных к целевой.

#### *Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»;
- ГосТех «Сервис управления потоками данных и извлечения, преобразования и загрузки данных»;
- ГосТех. Базовые сервисы работы с данными (СУБД).

## **4. Реализация потоков передачи данных**

#### *Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи:

- осуществить переход к целевой модели потоков данных в соответствии с утвержденным планом перехода;
- обеспечить тестирование разработанных потоков данных;
- внести описания новых потоков данных в **хранилище метаданных** организации в соответствии с **процессом управления метаданными**;
- реализовать систему мониторинга потоков данных;
- обеспечить регистрацию всех нештатных ситуаций, связанных с потоками данных;
- реализовать оповещение должностных лиц организации и Потребителей данных о возникновении ошибок и нештатных ситуаций.

#### *Исходная информация*

Утвержденный план перехода от текущей модели потоков данных к целевой

#### *Результат выполнения мероприятия*

- осуществлен переход к целевой модели потоков данных в соответствии с утвержденным планом перехода;
- проведено тестирование разработанных потоков данных;

- выполняется мониторинг потоков данных, производится оповещение должностных лиц организации и Потребителей данных о возникновении ошибок и нештатных ситуаций.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»;
- ГосТех «Сервис управления потоками данных и извлечения, преобразования и загрузки данных»;
- ГосТех Базовые сервисы работы с данными (СУБД);
- Сервис ГосТех «Мониторинг функционирования сервисов и управление инцидентами».

#### 9.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА

Для оценки эффективности выполнения процесса рекомендуется применять следующие показатели:

- количество сбоев, ошибок и нештатных ситуаций в работе потоков данных за период всего по ИС / всего по организации;
- доля ошибок, сбоев и нештатных ситуаций в работе потоков данных, устраненных в установленные сроки.

#### 9.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА

Для оценки уровня зрелости процесса рекомендуется применять шкалу оценок, указанную в таблице 6.

Таблица 2. Описание уровня зрелости процесса

Критерии	Уровни зрелости процесса			
	1	2	3	4
Модель основных потоков данных организации загружена в <b>хранилище метаданных</b>		✓	✓	✓
Осуществляется мониторинг основных потоков данных, производится оповещение должностных лиц организации и Потребителей данных о возникновении ошибок и нештатных ситуаций		✓	✓	✓
Модель всех потоков данных организации загружена в <b>хранилище метаданных</b>			✓	✓
Осуществляется мониторинг всех потоков данных, производится оповещение должностных лиц организации и Потребителей данных о возникновении ошибок и нештатных ситуаций			✓	✓
Создание новых потоков данных осуществляется на основе утвержденного проекта и плана				✓



## 10. ПРОЦЕСС «УПРАВЛЕНИЕ ОСНОВНЫМИ ДАННЫМИ И НСИ»

### 10.1. ВВЕДЕНИЕ

#### Назначение и цели процесса

**Управление основными данными и нормативно-справочной информацией (НСИ)** - это совокупность мероприятий, направленных на систематизацию операций с основными и справочными данными, минимизацию рисков, обусловленных использованием в основных данных некачественных справочных данных.

Целью процесса управления основными данными и НСИ является обеспечение необходимого уровня качества и доступности основных и справочных данных, достигаемых в результате осуществления:

- классификации и нормализации основных и справочных данных;
- контроля качества основных и справочных данных;
- своевременной актуализации справочных данных в локальных справочниках за счет реализации интеграции с ЕСНСИ.

#### Основные понятия

**Основные данные** — это находящиеся в распоряжении организации сведения о субъектах и объектах права<sup>6</sup>.

**Справочные данные** — это данные, используемые для структуризации, классификации и интерпретации основных данных.

**Транзакционные данные** — это данные, которые порождаются в результате выполнения операций (например, запросов) с использованием основных данных, а также данные с высокой степенью изменчивости (сотни изменений в сутки).

**Служебные данные** — это данные, связанные с эксплуатацией технических и программных средств, иного оборудования, сведения о функционировании различных сервисов и служб.

**Локальные справочники** – справочники организации, которые используются при выполнении процессов организации, при этом данные справочника не используются другими организациями.

Для снижения риска появления некачественных справочных и основных данных необходимо обеспечить централизованное ведение справочников, востребованных в нескольких организациях, с использованием **ЕСНСИ**. На Рис.7 приведена общая схема использования основных данных ГИС.

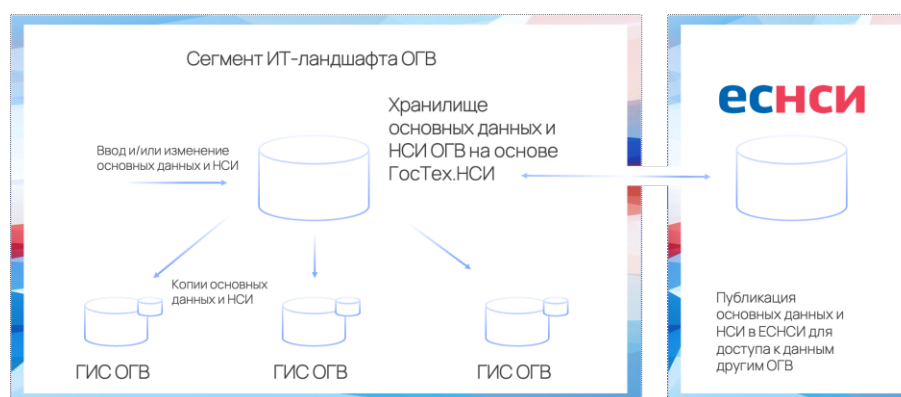


Рис. 7. Централизованная архитектура ведения НСИ с использованием ЕИС НСИ и базового сервиса ГосТех «Сервис управления НСИ»

### 10.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ

<sup>6</sup> См. Глоссарий для определения понятий «субъект права» и «объект права».

В рамках текущего процесса задействованы следующие роли:

- **Куратор данных:** Формирование требований к ведению основных и справочных данных. Контроль и согласование записей в реестрах и справочниках.
- **Архитектор данных:** Проектирование логической и физической модели данных хранилища основных данных и НСИ.
- **Аналитик данных:** Детализация требований к ведению основных и справочных данных;
- **Эксперт по качеству данных:** Формирование требований к качеству основных и справочных данных. Мониторинг и разрешение инцидентов по качеству данных.
- **Инженер данных:** Реализация обмена справочными данными с ЕСНСИ.

### 10.3. ЦЕЛЕВОЙ ПРОЦЕСС

Целевой процесс управления основными данными и НСИ выполняется на уровне организации и состоит из следующих мероприятий:

- 1) Оценка доступности и качества основных данных и НСИ;
- 1) Инвентаризация основных и справочных данных.
- 2) Интеграция справочных данных с ЕСНСИ;
- 3) Ведение основных и справочных данных;
- 4) Нормализация основных и справочных данных;
- 5) Контроль качества основных и справочных данных.

#### 1. Оценка доступности и качества основных данных и НСИ

##### *Описание мероприятия*

В рамках мероприятия необходимо:

- 1) Проверить выполнены ли следующие задачи:
  - все справочники организации загружены в инструмент ГосТех «Сервис управления НСИ»;
  - описания всех справочников (метаданные) внесены в хранилище метаданных;
  - осуществлена проверка каждого справочника на наличие аналогичного справочника в ЕСНСИ;
  - для справочников реализованы и применены правила контроля качества данных.

В случае невыполнения хотя бы одной из перечисленных задач, необходимо выполнить мероприятие **2** текущего процесса.

2) Проверить, производится ли для справочников, дублирующих справочники из ЕСНСИ, а также для справочников, подлежащих публикации в ЕСНСИ, обмен данными с ЕСНСИ. В случае отсутствия такого обмена необходимо выполнить мероприятие **3** текущего процесса.

3) Проверить выполнена ли нормализация справочных и основных данных. В случае необходимости осуществления нормализации справочных и основных данных следует выполнить мероприятие **5** текущего процесса.

4) Проверить, осуществляется ли контроль качества основных данных.

Если такой контроль не осуществляется, необходимо выполнить мероприятие **6** текущего процесса.

##### *Исходная информация*

- НПА, регламентирующие деятельность организации;
- Локальные справочники ИС организации, загруженные в инструмент ГосТех «Сервис управления НСИ»;
- Информация о реестрах и справочниках в инструменте ГосТех «Сервис каталога данных».

##### *Результат выполнения мероприятия*

- Улучшены значения показателей эффективности процесса (п 4.4.).

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис управления НСИ»;
- ЕСНСИ;
- ГосТех «Сервис каталога данных».

**2. Инвентаризация основных и справочных данных***Описание мероприятия*

В рамках мероприятия необходимо:

- определить перечень справочников, которые хранятся в ИС организации (локальные справочники), а также перечень справочников, импортируемых из внешних информационных ресурсов;
- загрузить справочники в инструмент **ГосТех. Сервис управления НСИ**;
- определить, какие данные, обрабатываемые в организации, являются основными, а какие справочными. При наличии справочных данных, не включенных в справочники, создать соответствующие справочники в инструменте **ГосТех. Сервис управления НСИ**;
- для каждого справочника выполнить создание и/или актуализацию модели данных справочника, метаданные справочника внести в **хранилище метаданных**, скорректировать метаданные, описывающие данные, использующие новый справочник в соответствии с **процессом управления метаданными**;
- для каждого справочника выполнить проверку на наличие аналогичного справочника в **ЕСНСИ**;
- для каждого созданного справочника (а при необходимости и для имеющихся справочников) реализовать и применить правила контроля качества данных в справочнике (см. **процесс управления качеством данных**);
- для локальных справочников, востребованных другими организациями, инициировать процедуру их включения в **ЕСНСИ** в качестве общедоступных.

*Исходная информация*

- НПА, регламентирующие деятельность организации;
- техническая документация на ИС организации

*Результат выполнения мероприятия*

- Определены основные и справочные данные, обрабатываемые в организации;
- Определен перечень используемых справочников;
- Для каждого используемого справочника определена дополнительная информация (описание, порядок обновления);
- Определены владельцы справочников;
- (при наличии) Выявлены дубли справочников и справочники, получаемые не из ЕСНСИ.
- Локальные справочники ИС организации загружены в инструмент ГосТех «Сервис управления НСИ»;
- Информация о справочниках внесена в **хранилище метаданных**;
- Реализованы правила контроля качества данных в справочниках;
- Получена оценка качества данных справочников ИС организации;
- Подготовлены предложения по публикации локальных справочников (в случае необходимости).

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис управления НСИ»;
- ЕИП НСУД;
- ЕСНСИ;
- ГосТех «Сервис каталога данных».

**3. Интеграция справочных данных с ЕСНСИ**

*Описание мероприятия*

Если по результатам выполнения мероприятия 2 текущего процесса выявлены справочники, получаемые не из ЕЧНСИ, а также локальные справочники, подлежащие публикации, необходимо реализовать интеграцию инструмента ГосТех «Сервис управления НСИ» с ЕЧНСИ для осуществления:

- импорта из ЕЧНСИ актуальных записей общедоступных справочников;
- экспорта в ЕЧНСИ актуальных записей публикуемых локальных справочников.

*Исходная информация*

Результаты выполнения мероприятия 2 текущего процесса.

*Результат выполнения мероприятия*

Настроена интеграция инструмента ГосТех. Сервис управления НСИ с ЕЧНСИ для импорта из ЕЧНСИ актуальных записей общедоступных справочников и / или экспорта в НСИ актуальных записей публикуемых локальных справочников.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис управления НСИ»;
- ЕЧНСИ.

**4. Ведение основных и справочных данных***Описание мероприятия*

В рамках мероприятия осуществляется внесение изменений в основные и справочные данные. При создании / изменении / удалении / блокировке записей в реестре / справочнике должны быть выполнены следующие задачи:

- подача заявки на изменение;
- оценка заявки на изменение;
- выявление заинтересованных третьих сторон и консультация с ними;
- определение ответственных за внесение изменений в ИС организации;
- внесение новых значений справочных данных в ИС организации;
- осуществление проверки качества данных измененного реестра / справочника;
- уведомление заинтересованных третьих сторон о внесенных изменениях;
- фиксация внесенных изменений в инструменте ГосТех «Сервис управления НСИ».

*Исходная информация*

- Локальные справочники ИС организации, загруженные в инструмент ГосТех «Сервис управления НСИ».
- Информация о реестрах и справочниках, актуализированная в хранилище метаданных.

*Результат выполнения мероприятия*

- Согласованные версии реестра / справочника с внесенными изменениями;

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис управления НСИ»;
- ГосТех «Сервис каталога данных».

*Рекомендации и примеры выполнения мероприятия*

Пример типового процесса формирования и согласования заявки на изменение записи в справочнике представлен на Рис.8:

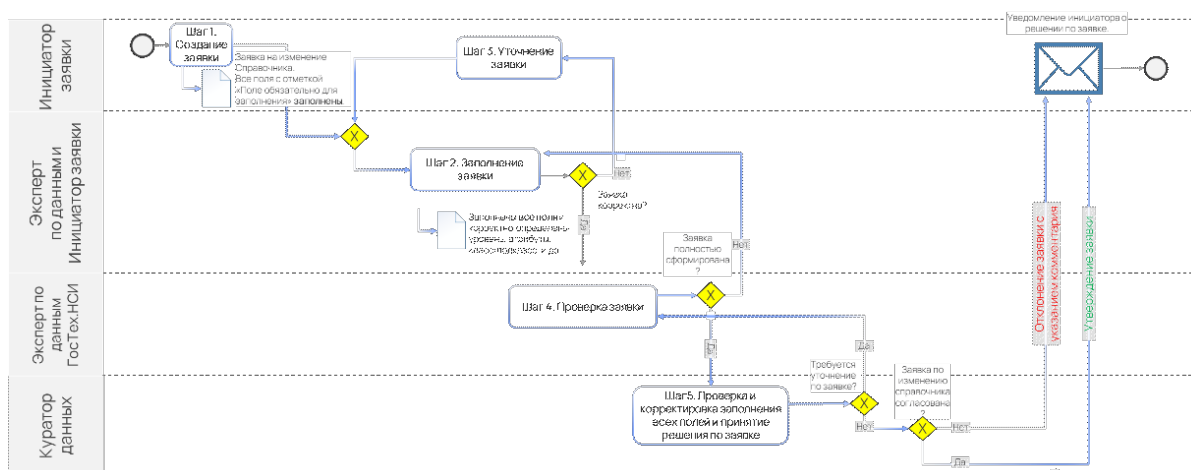


Рис. 8. Процесс формирования и согласования заявки на изменение справочника.

## 5. Нормализация основных и справочных данных

### Описание мероприятия

В рамках мероприятия необходимо:

- выполнить нормализацию записей в справочниках и реестрах основных данных для приведения значений к единому формату;
- исправить записи, не прошедшие форматно-логический контроль;
- поля записей, содержащие несколько значений, разделить на соответствующее число отдельных полей.

Внесение изменений в основные и справочные данные осуществляется в соответствии с мероприятием 4 текущего процесса.

### Исходная информация

- Локальные справочники ИС организации, загруженные в инструмент ГосТех «Сервис управления НСИ».
- Информация о реестрах и справочниках, актуализированная в хранилище метаданных.

### Результат выполнения мероприятия

Нормализованные основные и справочные данные.

### Рекомендуемые инструменты для реализации

- ГосТех «Сервис управления НСИ»;
- ГосТех «Сервис каталога данных».

### Рекомендации и примеры выполнения мероприятия

Пример стандартизации записей в справочнике: если в одной записи справочника адрес указан как «г. Москва, ул. Ленина, д.1, кв. 20», а в другом «Москва, ул. Ленина, 1, 20», то необходимо определить целевой формат значений (с указанием подписей «г.», «ул.», «д.» и прочих или без указания) и привести данные в справочнике к единому формату.

## 6. Контроль качества основных и справочных данных

### Описание мероприятия

В рамках мероприятия осуществляется контроль качества основных и справочных данных в соответствии с **процессом управления качеством данных**.

### Исходная информация

- Локальные справочники ИС организации, загруженные в инструмент ГосТех «Сервис управления НСИ».
- Информация о справочниках и реестрах, актуализированная в хранилище метаданных.

#### 10.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА

Для оценки эффективности процесса рекомендуется применять следующие показатели:

- средняя доля некорректных записей в реестрах от общего объема записей в реестрах (целевое значение: 0,01%).
- средняя доля некорректных записей в справочниках от общего объема записей в справочниках составляет (целевое значение: 0,01%).
- скорость актуализации записей в справочниках, ведущихся в инструменте ГосТех «Сервис управления НСИ» и в ЕСНСИ (целевое значение устанавливается организацией).

#### 10.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА

Для оценки уровня зрелости процесса рекомендуется применять шкалу оценок, указанную в Таблице 7.

Таблица 3. Описание уровня зрелости процесса

КРИТЕРИИ	Уровни зрелости процесса			
	1	2	3	4
Ключевые локальные справочники ИС организации загружены в инструмент ГосТех «Сервис управления НСИ»		✓	✓	✓
Информация о ключевых реестрах и справочниках загружена в хранилище метаданных		✓	✓	✓
Разработаны и используются регламенты ведения ключевых реестров / справочников организации		✓	✓	✓
Данные в реестрах / справочниках организации нормализованы, исключены дубликаты		✓	✓	✓
Все локальные справочники ИС организации загружены в инструмент ГосТех «Сервис управления НСИ»			✓	✓
Информация обо всех реестрах и справочниках загружена в хранилище метаданных			✓	✓
Разработаны и используются регламенты ведения всех реестров / справочников организации			✓	✓
Локальные справочники организации, имеющие аналоги в ЕСНСИ, актуализируются с использованием интеграции с ЕСНСИ			✓	✓
Осуществляется регулярный контроль качества основных и справочных данных		✓	✓	✓
В организации назначены сотрудники для ведения основных данных и НСИ		✓	✓	✓
Для каждого справочника организации назначен Владелец справочника		✓	✓	✓
Реализованы регламенты ведения справочников (основные данные и НСИ)			✓	✓
Выполняется постоянный мониторинг ведения основных данных и НСИ		✓	✓	✓

## 11. ПРОЦЕСС «УПРАВЛЕНИЕ КАЧЕСТВОМ ДАННЫХ»

### 11.1. ВВЕДЕНИЕ

#### Назначение и цели процесса

**Управление качеством данных** – набор подходов и задач, направленных на создание и использование наборов данных, отвечающих критериям качества данных.

Целями процесса управления качеством данных являются:

- предоставление потребителям данных, обладающих контролируемым уровнем качества;
- обеспечение организаций качественными данными для своевременного принятия решений;
- постоянное повышение качества обрабатываемых и предоставляемых данных;

Описание процесса управления качеством данных включает мероприятия, направленные на оценку и постоянное улучшение качества данных на протяжении всего их жизненного цикла, роли участников процесса, необходимые инструменты и результаты выполнения мероприятий.

#### Основные понятия

**Качество данных** - степень, в которой совокупность присущих данным характеристик отвечает установленным требованиям<sup>7</sup>.

**Контроль качества данных** – регулярное измерение (оценка) численных значений характеристик качества данных и сравнение их с установленными предельными значениями – критериями качества данных.

**Точки контроля** качества данных – этапы жизненного цикла данных, на которых осуществляется контроль характеристик качества данных.

**Правило контроля качества данных** - алгоритм оценки характеристики качества данных называется

Характеристики качества данных сгруппированы в **показатели качества данных**, приведенные в таблице 8.

Таблица 4. Показатели качества данных

ПОКАЗАТЕЛЬ	ОПРЕДЕЛЕНИЕ
Полнота	Насколько полно в системе представлены объекты (записи) и заполнены их атрибуты.
Достоверность	Степень соответствия данных требуемому формату и ограничениям.
Точность	Степень соответствия данных их истинным значениям или значениям, полученным из доверенного источника данных.
Актуальность	Степень соблюдения требований к срокам обновления и предоставления данных.
Консистентность	Непротиворечивость и связанность данных.

Примеры ошибок в данных, приводящих к снижению их качества, приведены на Рис. 9.

<sup>7</sup> ГОСТ Р ИСО 8000-2-2019 КАЧЕСТВО ДАННЫХ Часть 2 Словарь

Таблица 1

UID	Фамилия	Имя	Отчество	Дата рождения	СНИЛС
2d46b716-	Саакян	Раиса		26.05.1982	352-354-939-81
f8e885e3-	Лыков	Комстантин	Витальевич	18.04.1997	529-988-498-86
63199ac3-	Царева	Зинаида	Климентьевна	19.11.1957	388-545-368-44
25ddd64a-	Ооржак	Алдын-Херел	Аянович	31.02.1991	034-711-438-28
63199ac3-	Царева	Зинаида	Климентьевна	19.11.1957	388-545-368-44
d0c95b76-	Фатыхов	Айнур	Магомедович	10.08.2003	176-802-253-81

Таблица 2

СНИЛС	Тип пособия	Дата
352-354-939-81	12-13	04.03.2021
529-988-498-86	12-4	11.05.2019
388-545-368-44	5-109	22.11.2022
034-711-438-28	12-4	11.05.2001
867-90-037-70	8-3	04.11.2018
176-802-253-81	11-15	14.04.2021

Ошибки полноты 2: Не заполнен атрибут

Ошибка актуальности: Устаревшая дата

Ошибка точности: Недостовверное значение

Ошибка достоверности 1: Значение не соответствует ограничениям типа

Ошибка полноты 1: В таблице 1 нет соответствующей записи

Ошибка достоверности 2: Не соответствует правилам ФЛК

Ошибка консистентности: Дублирующееся значение

Рис. 9. Примеры ошибок в данных, ведущих к снижению их качества

## Рекомендуемые нотации и стандарты

В качестве одной из основ при применении процесса управления качеством данных к основным данным рекомендуется использовать ГОСТ Р ИСО 8000-100-2019.

### 11.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ

В рамках текущего процесса задействованы следующие роли:

- **Эксперт по качеству данных:** отвечает за определение характеристик и критериев оценки качества данных, а также правил контроля качества данных исходя из требований к качеству данных.
- **Куратор данных:** решает инциденты качества данных, поступившие в отношении данных, участвующих в межведомственном обмене. Осуществляет надзор за деятельностью экспертов по качеству данных.
- **Потребитель данных:** участвует в формировании требований к качеству данных.

### 11.3. ЦЕЛЕВОЙ ПРОЦЕСС

Целевой процесс управления качеством данных выполняется на уровне организации и состоит из следующих мероприятий:

- 1) Определение данных для контроля качества данных;
- 2) Определение требований потребителей к качеству данных;
- 3) Создание правил проверки качества данных;
- 4) Настройка регулярного запуска алгоритмов проверки качества данных и мониторинг результатов;
- 5) Создание инцидентов качества данных и управление их жизненным циклом;
- 6) Регулярная оценка и планирование инициатив по повышению качества данных.

Мероприятия 1. «Определение данных для контроля качества данных», 2. «Определение требований к качеству данных», 3. «Создание правил проверки качества данных», 4. «Первичная оценка качества данных» и 5. «Планирование мероприятий по повышению качества данных» **выполняются разово при начале работ** по управлению качеством данных и при существенных изменениях информационных систем организации.

Мероприятия 6. «Обеспечение регулярного мониторинга качества данных», 7. «Обеспечение решения инцидентов качества данных и претензий к качеству данных» и 8. «Регулярная оценка результатов, совершенствование и выполнение мероприятий управления качеством данных» **выполняются регулярно.**



Упрощенный пример выполнения мероприятий процесса качества данных приведен в конце раздела.

## **1. Определение данных для контроля качества данных**

### *Описание мероприятия*

Цель мероприятия: определить, контроль качества каких данных и на каких этапах необходимо осуществлять в приоритетном порядке.

Нет необходимости контролировать качество всех данных, обрабатываемых организацией на всех этапах их обработки. Достаточно обеспечить контроль качества только тех данных, которые необходимы для выполнения возложенных на организацию задач, а также тех данных, ведет в соответствии с требованиями НПА и/или предоставляет внешним потребителям в соответствии с межведомственными соглашениями.

При определении контролируемых данных необходимо выполнить следующие мероприятия:

- Определить, какие данные (записи, объекты, атрибуты), обрабатываемые организацией, наиболее удовлетворяют критериям, приведенным выше.
- Определить, точки контроля таких данных.

### *Исходная информация*

Для определения контролируемых данных используются:

- Требования нормативных правовых актов: контролю подлежат данные, за ведение которых организация несет ответственность согласно НПА;
- Межведомственные соглашения: контролю подлежат данные, предоставляемые организацией согласно межведомственным соглашениям;
- Требования внутренних потребителей данных: контролю подлежат данные, от качества которых зависит исполнение возлагаемых на организацию функций.
- Техническая документация на ИС организации и ЕЦП «ГосТех»: точки контроля качества данных определяются с учетом архитектуры ИС

### *Результат выполнения мероприятия*

- Контролируемые данные определены и отмечены в базовом сервисе ГосТех «Сервис контроля качества данных»;
- Контролируемые данные, участвующие в межведомственном информационном взаимодействии, отмечены в ФГИС ЕИП НСУД;
- Определены точки контроля качества для всех контролируемых данных.

### *Рекомендуемые инструменты для реализации*

Маркировка контролируемых данных выполняется:

- В базовом сервисе ЕЦП ГосТех «Сервис каталога данных» при организации процесса на уровне организации;
- В ФГИС ЕИП НСУД при организации процесса на уровне межведомственного взаимодействия.

## **2. Определение требований потребителей к качеству данных**

### *Описание мероприятия*

Цель мероприятия: определить и сформулировать требования к качеству контролируемых данных: измеряемые характеристики качества данных и критерии их качества.

Достижение целей мероприятия обеспечивается выполнением следующих мероприятий:

- Определить, какие требования к качеству данных организации установлены НПА и соглашениями об информационном взаимодействии.
- Определить, какие характеристики качества данных в каких точках контроля необходимо оценивать.

- Определить критерии качества данных<sup>8</sup> для всех характеристик всех контролируемых данных.

*Исходная информация*

- Сведения о контролируемых данных и точках контроля качества данных, полученные в результате выполнения мероприятия 1;
- Требования нормативных правовых актов и межведомственных соглашений в части качества данных (при наличии);
- Ведомственные документы, регламентирующие управление качеством данных (при наличии)

*Результат выполнения мероприятия*

Для контролируемых данных определены:

- показатели качества данных;
- характеристики качества данных;
- критерии качества данных.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис контроля качества данных».

### **3. Создание правил проверки качества данных**

*Описание мероприятия*

Цель мероприятия: сформулировать и реализовать правила проверки качества данных.

Достижение целей мероприятия обеспечивается выполнением следующих мероприятий:

- Определить, значения каких характеристик качества данных возможно рассчитывать автоматически с помощью программных средств, а для каких возможна только частичная автоматизация;
- Сформулировать правила контроля качества данных для всех характеристик всех контролируемых данных;
- Определить рекомендуемую периодичность проверок качества данных для каждого из правил контроля качества данных;
- Обеспечить реализацию в программных средствах правил проверки качества данных;
- Проверить корректность реализации правил проверки качества данных на реальных данных и/или специально подготовленных тестовых данных.

*Исходная информация*

- Требования к качеству данных, определенные при выполнении мероприятий 2 и с учетом Технической документации к средствам, на базе которых реализуются правила проверки качества данных: ИС организации, ЕЦП «ГосТех» и ЕИП НСУД.

*Результат выполнения мероприятия*

- В ИС организации, ЕЦП «ГосТех» и ЕИП НСУД реализованы корректные правила проверки качества данных;
- Сформулирован порядок экспертной оценки качества данных, автоматическая проверка которых на данном этапе невозможна, определены сроки и ресурсы, необходимые для данной оценки.

*Рекомендуемые инструменты для реализации*

- Для данных, размещаемых на витринах данных НСУД, контроль качества осуществляется с использованием проверок качества данных, создаваемых в ЕИП НСУД.

---

<sup>8</sup> Для критических данных критерии качества устанавливаются возможно более высокими, по возможности не более 0,1% - 1% некачественных данных

- Для остальных данных контроль качества осуществляется с использованием Базовый сервис ГосТех «Сервис контроля качества данных»,
- допускается применение и иных средств контроля качества, имеющихся в составе ИС организации.

#### **4. Настройка регулярного запуска алгоритмов проверки качества данных и мониторинг результатов**

##### *Описание мероприятия*

##### *Цели мероприятия:*

- Определить значения характеристик качества данных и удовлетворение критериев качества данных для всех контролируемых данных организации;
- обеспечить регулярный мониторинг качества данных организации.

##### *Для выполнения целей мероприятия необходимо:*

- обеспечить регулярное автоматическое исполнение проверок качества данных в программных средствах организации: расчет значений характеристик качества данных и сравнение их с критериями качества данных;
- обеспечить регулярную оценку качества данных тех характеристик качества данных, которые оцениваются в автоматизированном режиме;
- обеспечить регулярный анализ результатов выполнения критериев качества данных и формирование инцидентов качества данных в случае их невыполнения.

##### *Исходная информация*

- Результаты выполнения мероприятия 3. «Создание правил проверки качества данных»;
- Критерии качества данных, установленные в рамках мероприятия 2 «Определение требований потребителей к качеству данных»

##### *Результат выполнения мероприятия*

Документально закреплены процедуры контроля качества данных, предусматривающие, в том числе:

- порядок выполнения проверок качества данных организации;
- порядок получения и отработки претензий к качеству данных, формируемых внутренними пользователями организации;
- порядок информирования заинтересованных лиц о результатах проверок качества данных организации;
- порядок корректировки некачественных данных (см. также п.7);
- порядок формирования регулярной и разовой отчетности о результатах проверок качества данных организации;
- порядок внесения изменений в план мероприятий по повышению качества данных в информационных системах и витринах данных организации, разработанный в рамках п. 5.
- Процедуры контроля качества выполняются, ход и результаты их выполнения контролируются:
  - регулярно определяются значения всех характеристик качества данных;
  - значения характеристик качества данных регулярно проверяются на удовлетворение критериям качества данных.
- Рекомендуемые инструменты для реализации
  - ГосТех «Сервис контроля качества данных»;
  - Инструменты контроля качества данных ЕИП НСУД.

#### **5. Создание инцидентов качества данных и управление их жизненным циклом**

##### *Описание мероприятия*

Цель мероприятия: обеспечить оперативную обработку инцидентов качества данных и их решение в контролируемые сроки

Инциденты качества данных создаются:

- Непосредственно участником взаимодействия в ручном режиме;
- В результате обработки претензий физического или юридического лица, заявленных с использованием ЕПГУ;
- Автоматически в результате выполнения проверок контроля качества данных в ЕИП НСУД.

В целях обеспечения обработки и решения таких инцидентов качества данных, необходимо:

- Сформировать процедуры решения инцидентов качества данных, созданных в отношении данных организации;
- Обеспечить контроль исполнения процедур решения инцидентов качества данных.

#### *Исходная информация*

- Результаты выполнения мероприятия 5 «Настройка регулярного запуска алгоритмов проверки качества данных и мониторинг результатов запуска»;
- Инциденты качества данных, поступающие от потребителей данных.

#### *Результат выполнения мероприятия*

Сформированы и нормативно закреплены процедуры проверки и решения инцидентов качества данных, созданных в отношении данных организации, предоставляемых в рамках межведомственного информационного обмена. Процедуры должны предусматривать:

- порядок проверки и решения инцидентов качества данных;
- порядок внесения изменений в юридически значимые данные;
- порядок учета результатов проверки и решения инцидентов качества данных в рамках регулярного мониторинга качества данных.

#### *Рекомендуемые инструменты для реализации*

- Ситуационный центр – для получения дополнительных сведений об инцидентах качества данных.
- ФГИС «ЕИП НСУД» – для контроля решения автоматически созданных инцидентов качества данных.
- АРМ «Дата-стюард» – для локализации ошибок в данных, приведших к появлению инцидента качества данных.
- Стандартные или специализированные утилиты для корректировки данных организации;
- Базовый сервис ГосТех «Сервис контроля качества данных»;

## **6. Регулярная оценка и планирование инициатив по повышению качества данных**

### *Описание мероприятия*

Цель мероприятия: обеспечить непрерывность контроля и повышения качества данных организации

Непрерывность обеспечения качества данных достигается за счет регулярной оценки результатов процесса управления качеством данных. Изменение (совершенствование) мероприятий процесса управления качеством данных может потребоваться в случае возникновения любой из перечисленных ниже ситуаций:

- измеренные значения характеристик качества данных несколько раз подряд ниже целевых значений;
- изменился состав данных организации, подлежащих контролю качества;
- изменились требования к качеству данных;
- заметное и/или повторяющееся повышение числа претензий к качеству данных.

Для обеспечения непрерывного улучшения качества данных необходимо:

- Определить ситуации, в случае возникновения которых требуется выполнение мероприятий 1. «Определение критических контролируемых данных», 2. «Определение требований к качеству данных», 3. «Создание правил проверки качества данных» и 5. «Создание инцидентов качества данных и управление их жизненным циклом» .

- Выделить сотрудников, ответственных за регулярную оценку результатов контроля качества данных.

#### *Исходная информация*

Требования нормативных правовых актов, регламентирующие состав данных организации и состав данных, предоставляемых в рамках межведомственного информационного обмена.

#### *Результат выполнения мероприятия*

Порядок проведения регулярной оценки процесса управления качеством данных, содержащий:

- перечень ситуаций, в случае возникновения которых требуется перезапуск отдельных мероприятий процесса управления качеством данных;
- целевые показатели для регулярной оценки эффективности процесса управления качеством данных;
- процедуры проведения регулярной оценки процесса управления качеством данных
- в случае необходимости повышения качества данных, получаемых из внешних источников, в плане должны быть предусмотрены мероприятия по взаимодействию с источниками некачественных данных, направленные на повышение качества получаемых данных, включая заключение соглашений о поставке данных с требуемым уровнем качества данных (SLA).

#### *Рекомендуемые инструменты для реализации*

- Отсутствуют

### **11.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА**

Для оценки эффективности процесса рекомендуется применять следующие показатели:

1. Реализованы автоматические проверки качества в части полноты и ФЛК атрибутов баз данных ИС организации.
2. Реализованы автоматические проверки качества в части дублирования и уникальности не менее 50% атрибутов баз данных ИС организации.
3. Реализованы автоматические проверки качества в части соответствия данных баз данных ИС организации внешним эталонным данным для не менее 75% данных организации, связанных с эталонными данными.
4. Реализованы проверки актуальности данных организации.
5. Не менее 95% инцидентов качества данных решаются в установленные сроки.

### **11.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА**

Таблица 5. Описание уровня зрелости процесса

Критерии	Уровни зрелости процесса			
	1	2	3	4
Определен перечень контролируемых данных и сформированы требования к качеству критических данных.		✓	✓	✓
Проводится автоматическая оценка качества данных для отдельных наборов данных.		✓	✓	✓
Сформированы требования к контролю качества данных в разрезе не менее 5 показателей качества. Требования зафиксированы в соглашениях на поставку данных между ИС.			✓	✓
По всем контролируемым наборам данных осуществляется регулярный мониторинг качества данных при помощи автоматизированных инструментов.			✓	✓
Выполняется управление жизненным циклом инцидентов по данным.			✓	✓
Регулярно выполняется анализ первопричин недостаточного качества данных.				✓
Выполняется непрерывное улучшение текущего процесса управления				✓

качеством данных.				
Оценивается эффективность мероприятий по обеспечению качества данных; Качество данных, полнота и результативность мероприятий по его обеспечению подтверждаются независимым аудитом.				✓

## 12. ПРОЦЕСС «ОБЕСПЕЧЕНИЕ БЕЗОПАСНОСТИ ДАННЫХ»

### 12.1. ВВЕДЕНИЕ

#### Назначение и цели процесса

Безопасность информации (данных) - состояние защищенности информации (данных), при котором обеспечены ее конфиденциальность, доступность и целостность. Безопасность информации (данных) реализует механизмы защиты, которые обеспечивают:

- **конфиденциальность:** доступ к данным только авторизованных пользователей;
- **целостность:** достоверность и полноту данных и методов их обработки;
- **доступность:** доступ к данным и связанным с ними активам авторизованных пользователей по мере необходимости<sup>9</sup>.

Основными целями обеспечения безопасности данных являются:

- обеспечение конфиденциальности данных для предотвращения несанкционированного доступа и утечек;
- предотвращение нарушений целостности данных и несанкционированных изменений;
- предотвращение нарушений доступности данных для авторизованных потребителей;
- обеспечение соблюдения требований законодательства РФ в области информационной безопасности и защиты информации.

Цели процесса обеспечения безопасности данных достигаются за счет:

- разработки политик информационной безопасности, определяющих основные принципы и правила обеспечения безопасности данных;
- постоянного мониторинга и обнаружения потенциальных нарушений безопасности и проведения анализа инцидентов;
- ограничения доступа к данным на основе ролевой матрицы доступов.

Обеспечение безопасности данных на платформе «ГосТех» выполняется в том числе с учетом применением методических рекомендаций по обеспечению безопасности при разработке программного обеспечения с использованием компонентов единой цифровой платформы Российской Федерации «ГосТех».

При работе с информацией ограниченного доступа, в том числе персональными данными, полученными в результате обезличивания, должны соблюдаться требования законодательства об информации, информационных технологиях и о защите информации и законодательства Российской Федерации в области персональных данных.

### 12.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ

В рамках текущего процесса задействованы следующие роли:

- **Инженер данных:** отвечает за подготовку хранимых данных для последующего анализа.
- **Эксперт по информационной безопасности:** формирует требования к защите данных.
- **Куратор данных:** отвечает за идентификацию данных, подлежащих защите, с учетом требований НПА.
- **Эксперт по данным:** отвечает за маркировку данных, подлежащих защите, и их учет.

### 12.3. ЦЕЛЕВОЙ ПРОЦЕСС<sup>10</sup>

---

<sup>9</sup> ГОСТ Р ИСО/МЭК 17799–2005 «Информационная технология. Практические правила управления информационной безопасностью»

Целевой процесс обеспечения безопасности данных выполняется на уровне организации и состоит из следующих мероприятий:

**Формирования требований к защите информации, содержащейся в информационной системе:**

1) определение мероприятий, необходимых для обеспечения безопасности данных

**Разработка системы защиты информации информационной системы:**

1) определение требований по безопасности данных;

2) выявление и классификация данных, требующих защиты;

3) проектирование модели угроз и нарушителя;

4) проектирование ролевой матрицы доступов к данным;

**Внедрения системы защиты информации информационной системы:**

1) реализация мер защиты информации;

**Обеспечение защиты информации в ходе эксплуатации аттестованной информационной системы:**

1) аудит и логирование событий информационной безопасности;

2) контроль безопасности данных.

При управлении данными государственных информационных систем, размещаемых на единой цифровой платформе Российской Федерации «ГосТех», владелец указанных данных должен обеспечить реализацию Требований о защите информации, не составляющей государственную тайну, содержащейся в государственных информационных системах, утвержденных приказом ФСТЭК от 11 февраля 2013 г. № 17.

*Описание мероприятия*

В случае разработки / развития ИС организации необходимо выполнить мероприятия 2. – 8.

В случае изменения состава / структуры данных, обрабатываемых в организации, необходимо выполнить мероприятия 2, 3, 5, и 8.

Мероприятие 9 выполняется на постоянной основе.

*Исходная информация*

НПА, регламентирующие деятельность организации.

*Результат выполнения мероприятия*

Повышена эффективность обеспечения безопасности данных в организации.

*Рекомендуемые инструменты для реализации*

Приведены в описаниях соответствующих мероприятий.

## **Определение требований по безопасности данных**

*Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи:

- определить категории информации, для которой формируются требования по информационной безопасности:

- конфиденциальная информация;
- конфиденциальная информация с отдельной пометкой «персональные данные»;
- отраслевые тайны (медицинская тайна, тайна связи, налоговая тайна и другие).

- определить перечень НПА, которые применимы в части формирования требований к безопасности данных.

- сформулировать требования к безопасности данных в зависимости от категории данных.

---

<sup>10</sup> Целевой процесс учитывает требования приказа ФСТЭК от 11 февраля 2013 г. N 17 ОБ УТВЕРЖДЕНИИ ТРЕБОВАНИЙ О ЗАЩИТЕ ИНФОРМАЦИИ, НЕ СОСТАВЛЯЮЩЕЙ ГОСУДАРСТВЕННУЮ ТАЙНУ, СОДЕРЖАЩЕЙСЯ В ГОСУДАРСТВЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМАХ.



*Исходная информация*

- НПА в части хранения данных и безопасности данных;
- техническая документация на ИС организации;
- сведения о данных в **хранилище метаданных**.

*Результат выполнения мероприятия*

Определены требования к безопасности данных в зависимости от категории данных.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»

**Выявление и классификация данных, требующих защиты***Описание мероприятия*

В рамках мероприятия необходимо выполнить следующие задачи:

- распределить данные по категориям информации, определенным по результатам предыдущего мероприятия;
- провести поатрибутную разметку данных (классификацию данных с точки зрения информационной безопасности) в соответствии с категориями информации по уровню конфиденциальности в **хранилище метаданных**;
- определить данные, которые передаются вне периметра организации и оценить требования к безопасности передачи данных.

Процедуру классификации данных требуется производить регулярно, в случае наличия изменений состава и структуры данных.

*Исходная информация*

- требования к безопасности данных в зависимости от категории данных;
- сведения о данных в **хранилище метаданных**.

*Результат выполнения мероприятия*

- в **хранилище метаданных** отражена классификация данных с точки зрения информационной безопасности;
- осуществлена оценка требований к безопасности передачи данных вне периметра организации.

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных».

**Проектирование модели угроз и нарушителя***Описание мероприятия*

В рамках мероприятия необходимо разработать модель угроз и нарушителя в соответствии с действующими нормативно-методическими документами в области защиты информации и персональных данных.

*Исходная информация*

- Действующие нормативно-методические документы в области защиты информации и персональных данных;
- Сведения о данных, загруженные в **хранилище метаданных**, классифицированные на уровне атрибутов.

*Результат выполнения мероприятия*

Модель угроз и нарушителя

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных».

**Проектирование ролевой матрицы доступов к данным***Описание мероприятия*

В рамках мероприятия необходимо реализовать ролевую матрицу доступов к данным.

Типовая матрица доступов к данным ИС организации должна содержать определение ролевого доступа (ролевой модели) к данным на уровне доступа к конкретному атрибуту таблицы базы данных;

В рамках работ по проектированию ролевой матрицы доступов к данным требуется:

- сформировать перечень активных пользователей в разрезе информационных систем;
- сформировать ролевую матрицу на основании имеющихся данных (набор ролей в привязке к организационно-штатной структуре, выполняемым функциям) в разрезе информационных систем;
- сформировать матрицу конфликтов в случае наличия противоречий в рамках существующей ролевой модели (если она существует);
- актуализировать ролевую модель доступа к данным информационных систем, включая устранение конфликтов в существующей ролевой модели;
- установить срок действия прав доступа и разработать процедуру отзыва прав по истечению заданного срока;
- сформировать критерии доступа к данным для внешних потребителей.

Все действия, описанные в данном разделе, рекомендуется выполнять на регулярной основе, в случае создания или изменения объекта в базе данных.

#### *Исходная информация*

- требования к безопасности данных;
- модель угроз и нарушителя;
- сведения о данных, загруженные в хранилище метаданных, классифицированные на уровне атрибутов.

#### *Результат выполнения мероприятия*

- сформирована ролевая матрица доступов к данным;
- ролевая матрица доступов к данным настроена в ИС организации;
- сформированы критерии доступа к данным для внешних потребителей.

#### *Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»;
- Личный кабинет участников электронного взаимодействия (ЛК УВ).

## **Реализация мер защиты информации**

### *Описание мероприятия*

В рамках мероприятия необходимо определить и реализовать меры защиты информации в соответствии с действующими нормативно-методическими документами в области защиты информации и персональных данных.

#### *Исходная информация*

- действующие нормативно-методические документы в области защиты информации и персональных данных;
- техническая документация на ИС организации;
- сведения о данных, загруженные в **хранилище метаданных**, классифицированные на уровне атрибутов.

#### *Результат выполнения мероприятия*

Реализованные группы мер по обеспечению безопасности данных, в том числе актуализированная модель угроз и нарушителя информационной безопасности в случае изменения состава и/или структуры данных, обрабатываемых в организации.

#### *Рекомендуемые инструменты для реализации*

Комплекс средств защиты информации.

## **Аудит и логирование событий информационной безопасности**

### *Описание мероприятия*

В рамках мероприятия необходимо реализовать аудит и логирование событий в ИС организации, связанных с операциями с данными, включая предоставление доступа к данным по обращениям внешних организаций. Логированию, в частности, должны подлежать следующие события:

- факт обращения к базе данных до уровня физических таблиц/атрибутов (сессия, запрос к базе данных, учетная запись, под которой осуществляется доступ, IP-адрес клиента, с которого осуществляется запрос, имя/домен компьютера). В момент каждого обращения к БД должна осуществляться автоматическая сверка запрошенной информации с правилами, описанными в политике безопасности, а также в ролевой модели доступа к базе данных.

#### *Исходная информация*

Сведения о данных, загруженные в **хранилище метаданных**, классифицированные на уровне атрибутов

#### *Результат выполнения мероприятия*

Выполнена настройка / доработка ИС организации для реализации аудита и логирования событий, связанных с операциями с данными, включая предоставление доступа к ним.

#### *Рекомендуемые инструменты для реализации*

- ГосТех Базовые сервисы работы с данными (СУБД).

### **Контроль безопасности данных**

#### *Описание мероприятия*

В рамках мероприятия необходимо выполнять регулярный анализ безопасности данных. Анализ требуется осуществлять для следующих направлений, которые должно быть прописаны в политике безопасности организации:

- контроль учетных записей пользователей, посредством которых осуществляется доступ к БД, как минимум, требуется контролировать – срок действия записи, длину и сложность пароля учетной записи;
- контроль технических учетных записей, как минимум, требуется осуществлять контроль IP-адреса и домена/имени компьютера, с которого осуществляется запрос к БД, перечень команд и объектов, которые используются учетной записью;
- контроль и защита резервных копий: требуется контролировать место хранения резервных копий и ограничивать доступ до уровня администратора БД, права и полномочия которого должны быть зафиксированы в политике безопасности.

В режиме реального времени требуется:

- осуществлять выявление аномалий в части работы с данными (попытка поиска по персональным данным, конфиденциальной информации, информации ограниченного использования, факта использования технических учетных записей не с зарезервированного под администратора БД IP-адреса, компьютера);
- обрабатывать зафиксированные инциденты безопасности доступа к данным согласно требованиям, прописанным в политике безопасности;
- отправлять инциденты, являющимися аномальными с точки зрения количества запросов, нетиповым поведением пользователя или администратора БД, на ручной разбор специалисту информационной безопасности, отвечающему за данное направление.

Также необходимо осуществлять:

- мониторинг актуальности нормативных оснований для предоставления доступа к данным;
- мониторинг выполнения категоризации и классификации данных;
- мониторинг актуальности модели угроз и нарушителя;
- мониторинг полноты и актуальности матрицы доступов к данным;
- мониторинг реализации правил обеспечения безопасности конфиденциальных данных;
- мониторинг процесса регулярного анализа безопасности данных;

*Исходная информация*

- НПА, регламентирующие деятельность организации;
- модель угроз и нарушителя;
- ролевая матрица доступов к данным;
- журналы (логи) событий информационной безопасности.

*Результат выполнения мероприятия*

Отчет о результатах анализа информационной безопасности, включая перечень мероприятий по повышению безопасности данных

*Рекомендуемые инструменты для реализации отсутствуют*

**12.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА**

Для оценки эффективности процесса управления безопасностью данных рекомендуется применять следующие показатели:

- скорость предоставления доступа к данным с даты получения заявки на доступ;
- количество атрибутов в ИС организации, для которых выполнена категоризация данных с точки зрения ИБ;
- количество инцидентов безопасности данных за период;
- среднее время отработки инцидентов безопасности данных за период.

**12.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА**

Для оценки уровня зрелости процесса рекомендуется применять шкалу оценок, указанную в таблице 10.

*Таблица 6. Описание уровня зрелости процесса*

КРИТЕРИИ	УРОВНИ ЗРЕЛОСТИ ПРОЦЕССА			
	1	2	3	4
Реализованы меры защиты информации в ИС организации		✓	✓	✓
Выполнена классификация и категоризация данных организации не менее 10% от всего объема данных		✓	✓	✓
Описана и реализована ролевая модель доступа к данным в ИС организации		✓	✓	✓
Выполнена классификация и категоризация данных организации не менее 30% от всего объема данных			✓	✓
Выполнена классификация и категоризация данных организации не менее 70% от всего объема данных			✓	✓
Выстроен регулярный мониторинг безопасности данных			✓	✓
Выполнена классификация и категоризация данных организации не менее 95% от всего объема данных				✓

## 13. ПРОЦЕСС «АНАЛИЗ И ПРЕДСТАВЛЕНИЕ ДАННЫХ»

### 13.1. ВВЕДЕНИЕ

#### Назначение и цели процесса

**Анализ и представление данных** – это процесс сбора, обработки, и представления данных конечным потребителям, а также выстраивания процессов анализа и интерпретации информации для целей принятия решений.

Ключевая цель процесса – своевременно обеспечить работников органа или организации государственного сектора, подведомственных им организаций и потребителей данных из иных органов и организаций простым доступом к необходимой информации в нужном формате, которая позволит сформировать, обосновать и принимать решения в области государственного управления, сократив трудозатраты на работу с большим объемом данных. Внедрение данного процесса должно позволять своевременно увидеть приоритетную информацию для организации, региона или ведомства, и принять решение на основе этого знания.

Анализ и представление данных решает следующие задачи:

- Выявление потребностей и сбор требований к аналитике;
- Поиск, заказ и сбор данных;
- Выполнение обработки данных и расчета дополнительных показателей;
- Формирование аналитических отчетов и визуальных представлений качественной информации, которой можно доверять при принятии решений.

#### Основные понятия

**Аналитика** – набор аналитической информации и результатов расчетов в виде различных показателей с необходимым уровнем детализации и агрегации, который позволяет получить полную картину о предметной области.

**Визуализация данных** – представление аналитической информации с помощью диаграмм, графиков и других наглядных визуальных представлений, которые позволят конечному пользователю формировать выводы.

**Анализ** – это процесс изучения аналитической информации о предметной области или деятельности организации с целью выявления закономерностей или тенденций, формирования частных заключений и общих выводов.

Аналитика, представляемая пользователям для формирования заключений, выводов и принятия решений, делится на несколько категорий в зависимости от объема трудозатрат конечного пользователя на анализ, необходимый для формирования выводов и принятия решений:

- **Описательная аналитика:** описывает текущее состояние анализируемой области и позволяет ответить на вопрос «Что произошло?». Примером описательной аналитики может быть диаграмма, отображающая количество граждан, воспользовавшихся льготными услугами за каждый месяц последнего года. К данному виду аналитики также относится регулярное формирование отчетности по предметной области, в том числе для целей контроля выполнения заданных показателей. Является одним из самых часто используемых способов представления данных.

- **Диагностическая аналитика:** дополняет описательную аналитику деталями произошедшего события и возможными причинами его возникновения, с целью построения причинно-следственных связей при анализе, помогает выявить закономерности или корреляции. Позволяет ответить на вопрос «Почему это произошло?». Примером диагностической аналитики может быть информация о невыполнении показателей по строительству по причине отсутствия заключенных контрактов или отсутствия доведения финансирования. Для диагностической аналитики характерно использование аналитической информации об объекте исследования в разные моменты времени, а также анализ исторической последовательности событий.

- **Прогнозная аналитика:** помогает оценить вероятность возникновения ситуации в будущем и помогает ответить на вопрос «Что произойдет?». Прогнозная аналитика может быть реализована через построение прогнозных моделей, в том числе с применением методов

машинного обучения, применением математических моделей, или при помощи использования инструментов имитационного моделирования. Примером прогнозной аналитики может являться прогноз роста экономических показателей на основании исторических данных.

- **Рекомендательная аналитика:** представляет конечным пользователям варианты возможных решений при возникновении проблемных случаев, в том числе с указанием расчетов вероятности корректности и надежности предлагаемых решений, или выявить наиболее благоприятный сценарий решения ситуации. Позволяет ответить на вопрос «Что я должен сделать?» или «Какое решение будет самым правильным?». При определении пороговых значений и уверенности конечных пользователей в правильности работы рекомендательной аналитики – процесс принятия решений может быть автоматизирован. Примером рекомендательной аналитики может являться отмена финансирования мероприятий (автоматическая, в случае автоматизации процесса исполнения решения), которые с высокой долей вероятности не будут исполнены на отчетную дату.



Рис.10. Категории аналитики в зависимости от трудозатрат пользователя на анализ информации, необходимых для принятия решения

### Рекомендуемые нотации и стандарты

Для управления аналитикой и использованием данных рекомендуемые нотации отсутствуют.

Для целей формирования спецификаций требований к инструментам аналитики рекомендуется использовать принципы, аналогичные формированию спецификаций требований к программному обеспечению (аналог - спецификация программных требований), указанных в ГОСТ Р 51904-2002 «Программное обеспечение встроенных систем. Общие требования к разработке и документированию».

## 13.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ

В рамках текущего процесса задействованы следующие роли:

- **Инженер данных:** реализует процессы сбора и обработки данных, настраивает потоки данных при реализации процессов интеграции данных, разрабатывает модели данных в хранилище данных или аналитической СУБД, отвечает за эксплуатацию и поддержку работоспособности настроенных потоков данных и моделей данных.

- **Аналитик данных:** реализует расчеты аналитических метрик в Хранилище данных или аналитическом СУБД, настраивает аналитические инструменты, отвечает за формирование аналитики и ее представление конечным пользователям, отвечает за эксплуатацию и поддержку работоспособности аналитических расчетов и отображения данных. Может являться разработчиком аналитических информационных панелей в инструментах визуализации данных (BI инструментах).

- **Потребитель данных:** лицо, использующие данные для выполнения стоящих перед ним задач и/или принятия управленческих решений.

Реализация непрерывного функционирования текущего процесса возможна при построении организационной модели в следующих вариантах:

- **Организация централизованной аналитики:**

Создается специальная группа из инженеров данных и аналитиков данных, которая отвечает за следующие этапы работ: подготовка к анализу, сбор и обработка данных, формирование аналитики. Сбор требований выполняется аналитиками путем взаимодействия с Экспертами по анализу (отраслевыми экспертами) и руководством ОГВ. Сотрудники группы организуют сбор данных и формируют аналитику и отчеты. Эксперту по анализу (отраслевые аналитики) пользуются сформированной аналитикой и отчетами при выполнении этапа работ по анализу и интерпретации данных. Является самой распространенной и часто используемой организационной моделью.

- **Организация аналитики в режиме самообслуживания:**

Данная модель, аналогично предыдущей модели, предполагает, что подготовкой к анализу, сбором и обработкой данных, занимаются инженеры данных.

При этом данная модель подразумевает, что в этап работ по формированию аналитики вовлечены как аналитики данных, так и Эксперты по анализу (отраслевые аналитики). Аналитики данных полностью отвечают за подготовку аналитических расчетов в Хранилище данных или аналитическом СУБД. Аналитики данных настраивают типовые формы отображения данных в аналитических инструментах. При этом Эксперты по анализу (отраслевые аналитики) не работают в хранилище данных, но могут самостоятельно настраивать собственные, дополнительные формы отображения данных в аналитических инструментах. Для организации такой модели необходимо учитывать наличие у Экспертов по анализу (отраслевые аналитики) компетенций по работе с данными в инструментах визуализации или в других аналитических инструментах.

### 13.3. ЦЕЛЕВОЙ ПРОЦЕСС

Анализ и представление данных выполняется на уровне ОГВ и состоит из следующих мероприятий:

- 1) Определение целевых пользователей аналитических инструментов;
- 2) Сбор требований к аналитике, определение исходных данных и показателей для анализа;
- 3) Разработка макетов отображения данных;
- 4) Сбор данных, проектирование и разработка моделей данных;
- 5) Реализация аналитических инструментов и формирование аналитики;
- 6) Анализ и интерпретация данных;
- 7) Формирование и ведение реестра аналитической отчетности;
- 8) Регулярная оценка и планирование инициатив по развитию аналитики данных.

#### 1. Определение целевых пользователей аналитических инструментов

##### *Описание мероприятия*

Аналитик данных должен определить перечень целевых пользователей аналитических инструментов, кто будет непосредственно работать с аналитикой и проводить регулярный анализ с целью формирования выводов. Данные сотрудники являются функциональными заказчиками требований к аналитике. Для каждого целевого пользователя необходимо однозначно определить зоны ответственности по функциональным областям с учетом отраслевой специфики. Набор доступной аналитики может быть ограничен в зависимости от определенных зон ответственности целевых пользователей.

Среди целевых пользователей необходимо однозначно выделить лиц, принимающих решения в части согласования итогового перечня необходимых аналитик.

##### *Исходная информация*

- Отсутствует

*Результат выполнения мероприятия*

Перечень целевых пользователей, с указанием функциональных зон ответственности.

Перечень лиц, принимающих решения, в части согласования итогового перечня необходимых аналитик.

*Рекомендуемые инструменты для реализации*

- Интервью

*Рекомендации и примеры выполнения мероприятия*

При выполнении мероприятий рекомендуется изучить штатную структуру ОГВ или организации государственного сектора с целью формирования понимания зон ответственности.

Также рекомендуется провести интервью с целевыми пользователями касательно основного рода деятельности, для которой может потребоваться аналитика.

Примером выполнения мероприятий может быть заполненная таблица в формате:

*Таблица 7. Пример перечня целевых пользователей*

ФИО	Должность	Функциональная зона ответственности с учетом отраслевой специфики	Является ли лицом, принимающим решение?

## **2. Сбор требований к аналитике, определение исходных данных и показателей для анализа**

*Описание мероприятия*

Требования к аналитике должны быть сформированы для каждой анализируемой предметной области (область анализа) посредством перечисления необходимых данных для проведения исследований. При формулировании требований необходимо отталкиваться от целей анализа, стоящих перед Экспертом по анализу в рамках каждого исследования.

Необходимые данные для анализа должны быть детализированы на перечень показателей, которые планируется использовать для анализа. Показатели должны быть категоризованы на расчетные показатели (в том числе расчетные агрегированные данные) и исходные показатели источников данных, которые будут отображаться конечному пользователю без каких-либо изменений по сравнению с источником данных. Для каждого расчетного показателя необходимо сформулировать логику расчета показателя и согласовать ее с Экспертами по анализу.

На основании перечня показателей, которые будут использоваться для анализа, необходимо определить перечень необходимых исходных данных с указанием источника данных.

В дополнение к пользовательским сценариям также необходимо собрать нефункциональные требования, которые могут наложить ограничения на дальнейшую реализацию аналитики (например, ограничение по глубине требуемой истории первичных данных, ограничение по типу обрабатываемой информации и т. д.).

В рамках формирования требований к аналитике необходимо однозначно определить категорию аналитики, которой будет пользоваться Потребителями данных: описательная, диагностическая, прогнозная или рекомендательная аналитика.

*Исходная информация*

Перечень целевых пользователей.

*Результат выполнения мероприятия*

Перечень требований к аналитике, который включает:

- Область анализа;
- Необходимые данные для анализа;



- Цель анализа;
- Частота обновления данных;
- Категория аналитики;
- Перечень аналитических разрезов для детализации необходимых данных (детализация по: субъектам РФ, по месяцам / дням);
- Перечень показателей, которые планируется использовать для анализа, с указанием логики расчета;
- Перечень необходимых исходных данных с указанием источника данных.

#### Рекомендуемые инструменты для реализации

- Отсутствуют

#### Рекомендации и примеры выполнения мероприятия

При формировании требований рекомендуется сформулировать пользовательский сценарий (или пользовательские истории) по работе данными, поскольку от этого зависит последующее представление данных. Пользовательские сценарии являются средством представления функциональных требований к аналитике и помогают выяснить как именно аналитика сможет принести пользу Эксперту по анализу. Пользовательский сценарий (может быть сформулирован в структуре, которая позволяет получить ответы на следующие вопросы:

- Кто является целевым конечным пользователем аналитики;
- Какие именно данные аналитический инструмент должен отображать целевому конечному пользователю;
- Как конечный пользователь будет пользоваться данными и какую цель будет преследовать при анализе.

Примером сбора требований к аналитике может быть файл, составленный в структуре, представленной на рисунке ниже.

Область для анализа	Необходимые данные для анализа	Департамент-заказчик требований и ответственный	Цель анализа	Детализация	Необходимые исходные данные	Источник данных	Требуемая периодичность получения данных
Развитие сетей связи	Покрытие связью населенных пунктов: Уровень проникновения услуг связи и качество связи в населенных пунктах	Департамент государственного регулирования рынка телекоммуникаций: в части развития сетей связи	Выявление населенных пунктов, в которых отсутствует или требуется повышение качества связи;	Данные в разрезе: - по периодам (по времени) - субъектов РФ - территориальных единиц (города, села)	Расчетные данные % и зон покрытия в населенных пунктах Карта населенных пунктов	ИС РКН ИС Росреестра (ФИАС)	по месяцам по месяцам

Рис. 11. Пример сбора требований к аналитике

Описание показателей рекомендуется оформлять в паспорте показателя. Пример паспорта показателя с указанием логики расчета представлен в таблице ниже.

Таблица 8. Пример паспорта показателя с указанием логики расчета

Наименование показателя	Описание показателя	Логика расчета (формула)	Предмет расчета	Частота обновления	Тип показателя	Фильтры	Детализация	Методика агрегации
Количество чеков и ККТ за месяц	Общее количество кассовых чеков по всем ККТ, которые используются для расчетов в интернете	$= A * B / C$	Кассовый чек	Ежемесячно	Расчетный	<a href="#">Перечень фильтров №1</a>	В детализации по владельцам ККТ (ИП/ЮЛ)	Суммирование по разрезу

### 3. Разработка макетов отображения данных

#### Описание мероприятия

Для целей экономии будущих трудозатрат на реализацию инструментов визуализации Аналитик данных должен спроектировать макеты отображения данных, используя требования к аналитике и перечень показателей, необходимый Эксперту по анализу для проведения исследований. Макеты отображения данных являются демонстрационным, концептуальным набором представлений данных для конечного пользователя.

В случае, если инструмент визуализации (или другой аналитический инструмент) должен отображать данные по нескольким областям анализа – предварительно необходимо

спроектировать структуру связанных отображаемых макетов и элементов (например, структуру связанных информационно-аналитических панелей). Данная структура предполагает создание интуитивно понятной логики навигации и переходов между несколькими макетами в соответствии с пользовательскими сценариями и логикой анализа. Проектирование отдельных макетов необходимо начинать после проектирования общей структуры.

При проектировании макетов необходимо соблюдать принцип «от общего к частному», переходя от показателей верхнего уровня к показателям более низкого уровня и углубленной детализации, а также принципы простоты, наглядности и сокращения количества информации в излишней детализации.

Разработанные макеты отображения данных необходимо согласовать с Экспертом по анализу и подтвердить соответствие между логикой навигации по макетам и потребностями Эксперта по анализу.

#### *Исходная информация*

- Перечень требований к аналитике;
- Перечень показателей, которые планируется использовать для анализа.

#### *Результат выполнения мероприятия*

- Макет отображения данных.

#### *Рекомендуемые инструменты для реализации*

- ГосТех «Визуализация данных»

### **4. Сбор данных, проектирование и разработка моделей данных**

#### *Описание мероприятия*

- Инженер данных, используя перечень необходимых исходных данных с указанием источника данных и перечень целевых показателей, которые планируется использовать для анализа, проектирует архитектуру данных хранилища данных, его сегмента или СУБД (см. процесс «Управление архитектурой данных»).

- Инженеру данных необходимо детализировать необходимые исходные данные систем источников до уровня объектов (например, таблиц) и атрибутов с учетом архитектуры хранения данных в системе-источнике.

- Для сформированного перечня исходных данных Инженер данных определяет частоту и способ получения или интеграции данных с учетом требований Потребителя данных. Сбор исходных данных для аналитической обработки должен выполняться в хранилище данных. Для полученных первичных данных должна быть выполнена маркировка метаданных, требующих защиты, в соответствии с процессом «Управление безопасностью данных».

- Инженер данных (если требуется, при поддержке Аналитика данных) должен выполнить процедуры обработки данных, реализовать аналитические расчеты в соответствии с указанной логикой расчета показателей и подготовить данные для отображения конечным пользователям.

- При выполнении данных процедур необходимо соблюдать требования к защите данных, в том числе учитывать специфику расчета агрегированных значений по конфиденциальным данным ввиду наличия возможности расшифровать агрегированные данные на основании совокупной информации.

- Подготовленные данные для передачи в аналитический инструмент с целью последующего отображения конечным пользователям должны находиться в Хранилище данных в области слоя витрин (CDM) или в области аналитического слоя.

- Для передачи подготовленных данных из Хранилища данных в аналитический инструмент Инженеру данных необходимо настроить интеграцию данных.

#### *Исходная информация*

- Перечень необходимых исходных данных с указанием источника данных.
- Перечень показателей, которые планируется использовать для анализа, с указанием логики расчета.

*Результат выполнения мероприятия*

- Физическая реализация модели данных в СУБД или другом аналитическом инструменте.
- Настроенные потоки передачи данных из Хранилища данных в аналитический инструмент.

*Рекомендуемые инструменты для реализации*

- Базовые сервисы «ГосТех» по работе с данными (СУБД).

**5. Реализация аналитических инструментов и формирование аналитики***Описание мероприятия*

После получения данных из Хранилища данных в инструмент визуализации Аналитику данных необходимо сформировать аналитику и настроить отображение данных в соответствии с макетом отображения данных.

Аналитик данных может реализовать дополнительные расчеты на уровне инструмента визуализации данных с использованием доступных языков программирования. Вся логика реализованных аналитических расчетов показателей должна быть задокументирована в Базовом сервисе ГосТех «Сервис каталога данных».

После настройки аналитических инструментов и формирования аналитики необходимо совместно с Экспертом по анализу провести ряд пробных исследований, чтобы убедиться в соответствии настроенных отображений требуемой логике анализа предметной области.

*Исходная информация*

- Макет отображения данных.
- Настроенные потоки передачи данных из Хранилища данных в аналитический инструмент.

*Результат выполнения мероприятия*

- Сформированная аналитика (настроенные отображения данных) в инструменте визуализации данных или другом аналитическом инструменте.

*Рекомендуемые инструменты для реализации*

- Базовый сервис ГосТех «Сервис визуализации данных»;
- Другие аналитические инструменты (при необходимости).

*Рекомендации и примеры выполнения мероприятия*

При настройке отображения данных в инструменте визуализации рекомендуется выводить подсказки конечным пользователям по отображаемым показателям и, при необходимости, ссылаться на логику расчета показателей.

Также рекомендуется в аналитическом инструменте указывать периметр данных, который подлежит анализу, и дату последнего обновления данных.

**6. Анализ и интерпретация данных***Описание мероприятия*

Потребители данных проводят анализ предметной области при помощи сформированной аналитики и настроенных отображений данных, в соответствии с ранее сформулированными целями анализа, интерпретирует информацию и делает выводы.

Выводы, которые формируются для руководства, должны приводить к логическому заключению и должны подтверждаться цифрами. Выводы могут быть сформированы в виде предложений по принятию решений для руководства, сфокусированы на ранее неизвестных трендах или подтверждать предположения.

*Исходная информация*

- Сформированная аналитика (настроенные отображения данных) в инструменте визуализации данных или другом аналитическом инструменте.

*Результат выполнения мероприятия*

- Сформированные выводы по результатам проведения анализа, в том числе для поддержки принятия решений на основании данных.

- Аналитическая справка или отчет для руководства (при необходимости).

*Рекомендуемые инструменты для реализации*

- Базовый сервис ГосТех «Сервис визуализации данных»;

*Рекомендации и примеры выполнения мероприятия*

Для проведения анализа предметной области рекомендуется сформировать гипотезы по исследованию данных и разработать сценарий анализа в виде логических рассуждений с учетом отраслевой специфики.

Выводы рекомендуется формулировать в первую очередь с ключевых тезисов, подтверждая их аналитикой во вторую очередь: не рекомендуется в основную часть аналитических справок для руководства включать последовательность сделанных логических рассуждений (от начала рассуждений - до выводов).

Примером выполнения мероприятий может быть сформированный вывод в части необходимости отзыва лицензии на деятельность регулируемой организации ввиду неоднократных нарушений, отмена или изменение программ финансирования по причине недостижения ожидаемых целей на протяжении определенного периода и т. д.

## **7. Формирование и ведение реестра аналитической отчетности**

*Описание мероприятия*

Для целей повышения доступности аналитики для других подразделений или заинтересованных сторон ОГВ или организации государственного сектора, в рамках которого была сформирована аналитика, необходимо внести информацию о ней в реестр аналитики базового сервиса ГосТех «Сервис каталога данных».

При формировании реестра необходимо зафиксировать:

- Описание сформированной аналитики;
- Перечень используемых показателей и ссылку на логику их расчета;
- Эксперта по анализу (или Отраслевого аналитика), кто пользуется информацией;
- Аналитика данных, кто вовлекался в формирование аналитики;
- Использованные источники данных;
- Режим обновления данных (дата, на которую данные актуальны);
- Количество уникальных пользователей аналитики в неделю / месяц / квартал.

*Исходная информация*

- Сформированная аналитика (настроенные отображения данных) в инструменте визуализации данных или другом аналитическом инструменте.

*Результат выполнения мероприятия*

- Перечень аналитики сформирован и загружен в базовый сервис ГосТех «Сервис каталога данных».

*Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных».

*Рекомендации и примеры выполнения мероприятия*

*Отсутствуют*

## **8. Регулярная оценка и планирование инициатив по управлению аналитикой данных**

*Описание мероприятия*

Аналитику данных необходимо на периодической основе оценивать востребованность сформированной аналитики посредством определения и отображения текущих показателей удовлетворенности конечных пользователей.

Для этого необходимо собирать следующие показатели:

- Число конечных пользователей аналитического инструмента и/или запросов в сутки, неделю или месяц;
- Мониторинг используемости информационно-аналитических панелей (дашбордов) в инструменте визуализации;
- Число зарегистрированных пользователей, а также число подключенных или активных пользователей за отчетный период или в среднем за сутки / рабочий день;
- Регулярность обновления данных.

Дополнительно рекомендуется проводить сбор обратной связи от конечных пользователей в части удобства функционала и наличия новых требований к существующей аналитике или запросов на новую аналитику.

При выявлении аналитики, которая утратила свою актуальность – необходимо согласовать данный факт с Экспертом по анализу и выполнить архивацию данных.

Также необходимо на периодической основе актуализировать данные о существующей, актуальной аналитике в базовом сервисе ГосТех «Сервис каталога данных» в соответствии с параметрами п7. Формирование и ведение реестра аналитической отчетности.

#### *Исходная информация*

- Перечень аналитики сформирован и загружен в базовый сервис ГосТех «Сервис каталога данных».

#### *Результат выполнения мероприятия*

- Перечень аналитики актуализирован в базовом сервисе ГосТех «Сервис каталога данных».

#### *Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных».

#### *Рекомендации и примеры выполнения мероприятия*

- Отсутствуют

### **13.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА**

Для оценки эффективности процесса рекомендуется применять следующие показатели:

- Рост количества потребителей данных витрин данных ОГВ;
- Рост количества аналитик / аналитических отчетов на основе витрин данных;
- Скорость реализации аналитических отчетов удовлетворяет потребителей данных
- Отсутствие и/или минимизация неиспользуемых аналитических отчетов;
- Отсутствие дублирующих аналитических отчетов.

### **13.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА**

Таблица 13. Описание уровня зрелости процесса

АНАЛИЗ И ПРЕДСТАВЛЕНИЕ ДАННЫХ	
УРОВЕНЬ ЗРЕЛОСТИ	ОПИСАНИЕ УРОВНЯ ЗРЕЛОСТИ
Уровень 1	Ни одно из мероприятий процесса не выполняется.
Уровень 2	Перечисленные ниже мероприятия выполнены частично:

АНАЛИЗ И ПРЕДСТАВЛЕНИЕ ДАННЫХ	
УРОВЕНЬ ЗРЕЛОСТИ	ОПИСАНИЕ УРОВНЯ ЗРЕЛОСТИ
	<ul style="list-style-type: none"> <li>- Определение целевых пользователей аналитических инструментов;</li> <li>- Сбор требований к аналитике, определение исходных данных и показателей для анализа;</li> <li>- Разработка макетов отображения данных;</li> <li>- Сбор данных, проектирование и разработка моделей данных;</li> <li>- Реализация аналитических инструментов и формирование аналитики;</li> </ul>
Уровень 3	<p>Выполнены все мероприятия, перечисленные ниже:</p> <ul style="list-style-type: none"> <li>- Определение целевых пользователей аналитических инструментов;</li> <li>- Сбор требований к аналитике, определение исходных данных и показателей для анализа;</li> <li>- Разработка макетов отображения данных;</li> <li>- Сбор данных, проектирование и разработка моделей данных;</li> <li>- Реализация аналитических инструментов и формирование аналитики;</li> </ul> <p>Существует сформированная аналитика в аналитическом инструменте. Обновление данных в аналитическом инструменте реализуется вручную или автоматически. Анализ и интерпретация данных не проводятся.</p>
Уровень 4	<p>Выполнены все мероприятия, перечисленные ниже:</p> <ul style="list-style-type: none"> <li>- Определение целевых пользователей аналитических инструментов;</li> <li>- Сбор требований к аналитике, определение исходных данных и показателей для анализа;</li> <li>- Разработка макетов отображения данных;</li> <li>- Сбор данных, проектирование и разработка моделей данных;</li> <li>- Реализация аналитических инструментов и формирование аналитики;</li> </ul> <p>Существует сформированная аналитика в аналитическом инструменте. Обновление данных в аналитическом инструменте реализуется автоматически. Регулярно проводится анализ и интерпретация данных.</p>
Уровень 5	<p>Выполняются все мероприятия процесса, в том числе:</p> <ul style="list-style-type: none"> <li>- Сформированная аналитика описана в базовом сервисе ГосТех «Сервис каталога данных».</li> <li>- Выполняется регулярная оценка и планирование инициатив по управлению аналитикой данных.</li> </ul>

## **14. ПРОЦЕСС «УПРАВЛЕНИЕ МЕТАДААННЫМИ И ПРОИСХОЖДЕНИЕМ ДАННЫХ»**

### **14.1. ВВЕДЕНИЕ**

#### **Назначение и цели процесса**

Управление метаданными – процесс управления «данными о данных», направленный на упорядочивание описания данных, их проектирование и хранение, а также использование данных, включая информацию об их свойствах, атрибутах и отношениях.

Управление метаданными является одним из ключевых процессов управления данными, который должен выполняться на всех этапах работы с данными.

Основными целями процесса управления метаданными являются:

- формирование виртуальной модели (онтологии) предметных областей и доменов государственного управления, за счет обеспечения формального, подробного и всеобъемлющего описания указанной модели;
- обеспечение онтологического единства государственных данных за счет применения настоящего стандарта;
- обеспечение компактного, систематизированного представления логических моделей ведомственных информационных ресурсов и информационных систем за счет увязки логических моделей данных ведомственных информационных ресурсов и информационных систем с онтологией государственного управления;
- обеспечение совместного однозначного понимания и использования структуры информации людьми или программными агентами за счет разработки онтологического описания предметной области с использованием формализованного языка и включением машинно-интерпретируемых формулировок основных понятий предметной области и отношений между ними;
- формирование необходимых предпосылок для создания непротиворечивых, полных и целостных комплексов нормативных правовых документов за счет формализации и автоматизации процесса создания и внесения изменений в нормативные правовые документы благодаря использованию формальной (концептуальной) модели предметной области;
- обеспечение быстрого и комфортного поиска средств межведомственного взаимодействия, а также источников данных, необходимых для оказания государственных услуг и выполнения государственных функций, с использованием естественного языка;
- оптимизация процессов и процедур получения данных в рамках межведомственного информационного обмена за счет обеспечения перехода к использованию семантических протоколов для запроса данных (например, СМЭВ-QL)

Цели процесса управления метаданными достигаются за счет:

- определения структуры данных: создания и ведения описаний данных, их свойств, атрибутов, отношений и структур;
- анализа происхождения данных: определения источников и получателей данных, описания движения данных между информационными системами и базами данных;
- встраивания описания данных в карту данных: увязки логических моделей данных с сущностями (понятиями) концептуальной модели данных;
- расширения концептуальной модели данных: пополнения концептуальной модели данных профильного домена (предметной области);
- обеспечения соответствия требованиям законодательства и стандартам: установления правил для описания и использования метаданных в соответствии с требованиями законодательства и стандартов;
- управления изменениями: управления версиями и изменениями в метаданных, включая обновление, модификацию и удаление метаданных.

### Основные понятия

**Метаданные** — это информация, которая описывает / логически увязывает другие данные, что помогает понять происхождение, структуру, предметную область и контекст данных.

Метаданные в информационных системах организации можно разделить на 3 категории:

- **Операционные метаданные (каталог и карта данных)** – метаданные, включающие концептуальную и логическую модели данных, связи между ними, а также методологию и подходы формирования концептуальной и логической модели данных;
- **Технические метаданные** – метаданные, описывающие технические компоненты информационных систем, атрибуты физической модели данных и связи между ними, правила физического преобразования данных в ETL / ELT процессах;
- **Функциональные метаданные** – метаданные, представляющие собой сведения о процессах, выполняемых в информационных системах организации.

Метаданные описывают именованные, логически связанные массивы данных (далее — **контейнеры данных**). Как правило, метаданные содержат общие сведения о контейнере данных (наименование, описание, категории, другие характеристики) и описание его структуры (атрибутивного состава). Примерами контейнеров данных являются таблицы баз данных, а также средства межведомственного взаимодействия (виды сведений СМЭВ 3 и регламентированные запросы СМЭВ 4).

В свою очередь метаданные, описывающие каждый из типов контейнеров данных, организуются в виде **контейнеров метаданных**, каждая запись которого соответствует контейнеру данных заданного типа. Примерами контейнеров метаданных, реализованных в ЕИП НСУД, являются:

- контейнер информационных систем, содержащий метаданные, описывающие структуры данных БД ИС организации;
- контейнер витрин данных, содержащий метаданные, описывающие структуры (модели) данных витрин данных НСУД.

В совокупности контейнеры метаданных составляют **каталог данных**.

Помимо метаданных, входящих в каталог данных, метаданными также считается **концептуальная модель данных** (далее — **КМД**), описывающая **онтологию** предметных областей государственного управления. КМД включает сведения о понятиях (сущностях), описывающих **субъекты и объекты государственного управления**, такие как «Физическое лицо», «Паспорт гражданина Российской Федерации», «Транспортное средство». Близкие по смыслу понятия объединяются в иерархию. Для каждого понятия КМД возможно описать его атрибуты и связи с иными понятиями КМД.

Для каждой предметной области (домена) может формироваться свой фрагмент КМД, который впоследствии встраивается в общегосударственную (глобальную) КМД.

Для увязки каталога данных с КМД осуществляется сопоставление метаданных логического уровня каталога данных (наименований атрибутов или контейнеров) с понятиями или атрибутами понятий КМД. В результате образуется **ведомственная карта государственных данных**.

Все метаданные, описывающие данные организации, хранятся в **хранилище метаданных** организации, реализованном с использованием инструмента ГосТех «Сервис каталога данных».

### Рекомендуемые нотации и стандарты

Для предварительного описания концептуальной модели данных рекомендуется использовать нотацию OWL (англ. Web Ontology Language) — язык описания онтологий для семантической паутины.

В стандартах серии ГОСТ Р ИСО/МЭК 11179 определяются виды и качество метаданных, необходимые для описания данных, и специфицируется управление и администрирование этих метаданных в регистре метаданных.



## 14.2. РОЛИ И ОРГАНИЗАЦИОННАЯ МОДЕЛЬ

Участниками процесса управления метаданными и происхождением данных являются организация – Куратор данных, а также Центр компетенции:

- **Аналитик данных** (организация): Ведение операционных и функциональных метаданных в хранилище метаданных организации;
- **Архитектор данных** (организация): Определение типов контейнеров данных ИС организации и требований к их описанию;
- **Куратор данных** (организация): Организация экспертной работы по формированию метаданных организации;
- **Инженер данных** (организация): Формирование требований к описанию технических метаданных ИС организации, ведение технических метаданных ИС организации;
- **Эксперт по данным** (Центр компетенции): Согласование метаданных организации, размещаемых в ЕИП НСУД, экспертиза локального фрагмента КМД организации, встраивание локального фрагмента КМД в глобальную КМД.

## 14.3. ЦЕЛЕВОЙ ПРОЦЕСС

Целевой процесс управления метаданными выполняется на уровне организации и Центра компетенции и состоит из следующих мероприятий:

- 1) Определение мероприятий, необходимых для реализации управления метаданными;
- 2) Ведение каталога данных;
- 3) Ведение карты данных;
- 4) Синхронизация метаданных с ЕИП НСУД.

### 1. Определение мероприятий, необходимых для реализации управления метаданными

#### *Описание мероприятия*

В рамках мероприятия необходимо:

- 1) Проверить, загружена ли в **хранилище метаданных** информация о всех данных, обрабатываемых в организации, а также **технические и функциональные метаданные**.

В случае отсутствия описания каких-либо данных в **хранилище метаданных** выполнить мероприятие **2** текущего процесса.

- 2) Проверить, учтена ли информация о всех данных, обрабатываемых в организации, в локальной карте данных, загруженной в **хранилище метаданных**.

В случае отсутствия учета каких-либо данных в ведомственной карте данных выполнить мероприятие **3** текущего процесса.

- 3) В случае изменения состава данных, обрабатываемых в организации, выполнить мероприятия **2** и **3** текущего процесса.

- 4) В случае, если операционные метаданные, загруженные в **хранилище метаданных**, не синхронизированы с **ЕИП НСУД**, выполнить мероприятие **4** текущего процесса.

#### *Исходная информация*

- НПА, регламентирующие деятельность организации;
- техническая документация на ИС организации;
- метаданные организации, загруженные в **хранилище метаданных**;
- операционные метаданные организации, загруженные в инструмент **ЕИП НСУД**.

#### *Результат выполнения мероприятия*

- информация о всех данных, обрабатываемых в организации, включая локальную карту данных, загружена в **хранилище метаданных**;
- **технические и функциональные метаданные** загружены в **хранилище метаданных**;
- **операционные метаданные** организации синхронизированы с **ЕИП НСУД**.

#### *Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»;
- ФГИС ЕИП НСУД.

## 2. Ведение каталога данных

### *Описание мероприятия*

В рамках мероприятия необходимо:

1) Определить типы контейнеров данных, представленных в ИТ-ландшафте организации, для которых требуется сформировать метаданные. Все контейнеры данных ИС и хранилища данных организации, которые находятся в продуктивной эксплуатации, должны быть описаны в хранилище метаданных. Для контейнеров, находящихся в средах разработки и тестирования, такое описание не является обязательным.

Ниже представлен список наиболее распространённых контейнеров данных, для которых требуется сформировать метаданные:

#### **В периметре организации:**

- таблицы СУБД;
- витрины данных;
- регламентированные запросы;
- виды сведений СМЭВ 3;
- показатели;
- наиболее важные данные журналирования и мониторинга действий пользователей и функционирования ИС.

#### **При межведомственном обмене:**

- БД ИС;
- наборы данных НСУД;
- витрины данных НСУД;
- регламентированные запросы;
- виды сведений СМЭВ 3.

2) Определить требования к составу метаданных.

При необходимости перечень требований к составу метаданных может быть расширен и дополнен для обеспечения адекватного описания специфических для своей предметной области данных.

3) Сформировать **хранилище метаданных** организации с использованием инструмента **ГосТех «Сервис каталога данных»** в соответствии с требованиями к составу метаданных, определенных в рамках предыдущей задачи;

4) Заполнить **хранилище метаданных**, осуществив сбор метаданных из ИС и/или хранилища данных организации с использованием автоматизированных средств и в ручном режиме.

Данные, для которых формируются метаданные, должны быть описаны на физическом и логическом уровне.

### *Исходная информация*

- НПА, регламентирующие деятельность организации;
- техническая документация на ИС организации.

### *Результат выполнения мероприятия*

- Сформировано хранилище метаданных организации;
- В хранилище метаданных организации загружены метаданные ИС и хранилища данных организации.

### *Рекомендуемые инструменты для реализации*

- Базовый сервис ГосТех «Сервис каталога данных»
- ГосТех. Сервис Управление данными

### *Рекомендации и примеры выполнения мероприятия*

Рекомендуемый перечень состава метаданных приведен в таблице ниже:

Таблица 9. Рекомендуемый перечень состава метаданных

№	ПЕРЕЧЕНЬ МЕТАДАНЫХ	ОПИСАНИЕ
1	Уникальный идентификатор объекта данных	Уникальный идентификатор, служебное поле
2	Краткое наименование объекта данных	Краткое наименование объекта данных, в том числе расшифровка аббревиатур
3	Полное наименование объекта данных	Полное название на русском языке
4	Описание	Краткое описание смысла объекта данных
5	Тип	Атрибут; Показатель – сформированный на базе одного элемента данных; Расчетный показатель – сформированный на базе нескольких элементов данных /систем источников
6	Тип данных	Тип данных, в котором хранится объект данных (текст / число / дата / др.)
7	Домен	Домен, в рамках которого ведется объект данных
8	Псевдонимы	Название объекта данных в других системах/отчётах
9	Расположение	Место хранения объекта данных (таблица / витрина /др.)
10	Владелец / Куратор	Должность и ФИО сотрудника, который отвечает за корректность и актуальность объекта данных, включая методологию
11	Эксперт по показателю (*для показателей)	Должность и ФИО сотрудников, которые могут выступать экспертами по показателю
12	Дата создания	Дата создания объекта данных
13	Дата изменения	Дата последнего изменения объекта данных
14	Регулярность обновления	Периодичность обновления объекта данных
15	Методология расчета (*для показателей)	Описание формулы расчета показателя с указанием кодов показателей, используемых в расчетах
16	Единица измерения (*для показателей)	Единица измерения, в которой выражается показатель
17	Разрядность	Разрядность показателя
18	Нормативный документ	Указывается в случае наличия нормативного документа, который регламентирует методику ведения и расчёта показателя
19	Влияет на показатели	Показатели, на которых влияет значение показателя, список в формате уникальный идентификатор - краткое наименование
20	Конфиденциальность	Описание конфиденциальности данных и возможности их использования
21	Источник	Система-источник
22	Статус	Отображает статус объекта данных (черновик, на согласовании, актуальный, неактуальный)

### 3. Ведение карты данных

#### *Описание мероприятия*

В рамках мероприятия необходимо:

1) Сформировать в **хранилище метаданных** локальный фрагмент КМД, описывающий предметную область (домен), к которому относятся данные организации.

Формирование локального фрагмента КМД выполняется с учетом понятий глобальной КМД в **ЕИП НСУД**.

2) Сформировать ведомственную карту данных, выполнив привязку метаданных логического уровня **каталога данных** (наименований атрибутов или контейнеров) с понятиями или атрибутами понятий локального фрагмента КМД.

#### *Исходная информация*

- Метаданные, загруженные в хранилище метаданных;
- Концептуальная модель данных в ЕИП НСУД.

#### *Результат выполнения мероприятия*

- В **хранилище метаданных** сформирован локальный фрагмент КМД и словарь терминов на ее основе.

- В **хранилище метаданных** сформирована локальная карта данных организации.

#### *Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»
- ФГИС ЕИП НСУД

#### 4. Синхронизация метаданных с ЕИП НСУД

##### *Описание мероприятия*

В рамках мероприятия необходимо выполнить синхронизацию операционных метаданных с ЕИП НСУД, для чего:

- передать в ЕИП НСУД локальный фрагмент КМД и ведомственную карту данных организации. До реализации на стороне ЕИП НСУД специализированного API, обеспечивающего автоматизированную передачу соответствующих сведений, допускается передача указанных сведений иными способами (посредством электронной почты и т. п.);
- Эксперту по данным Центра компетенции согласовать предоставленные сведения, обеспечить встраивание локального фрагмента КМД в глобальную КМД, а ведомственной карты данных — в карту государственных данных.

##### *Исходная информация*

- ведомственная карта данных, загруженная в хранилище метаданных;
- глобальная карта данных в ЕИП НСУД.

##### *Результат выполнения мероприятия*

- Метаданные организации синхронизированы с глобальной КМД и картой данных.

##### *Рекомендуемые инструменты для реализации*

- ГосТех «Сервис каталога данных»;
- ФГИС ЕИП НСУД.

#### 14.4. ОЦЕНКА ЭФФЕКТИВНОСТИ ПРОЦЕССА

Для оценки эффективности процесса рекомендуется применять следующие показатели:

- доля контейнеров данных, сведения о которых внесены в **хранилище метаданных** (целевое значение: 100 %);
- доля контейнеров данных, сведения о которых включены в ведомственную карту данных (целевое значение: 100 %);
- доля операционных метаданных, синхронизированных с глобальной картой данных в **ЕИП НСУД** (целевое значение: 100 %).

#### 14.5. ОЦЕНКА УРОВНЯ ЗРЕЛОСТИ ПРОЦЕССА

Для оценки уровня зрелости процесса рекомендуется применять шкалу оценок, указанную в Таблице 15.

Таблица 10. Описание уровня зрелости процесса

КРИТЕРИИ	УРОВНИ ЗРЕЛОСТИ ПРОЦЕССА			
	1	2	3	4
В <b>хранилище метаданных</b> загружено описание данных, используемых для межведомственного информационного обмена		✓	✓	✓
В <b>хранилище метаданных</b> загружена локальная карта данных, используемых для межведомственного информационного обмена		✓	✓	✓
В <b>хранилище метаданных</b> загружено описание всех данных организации			✓	✓
В <b>хранилище метаданных</b> загружена локальная карта всех данных организации			✓	✓
В <b>хранилище метаданных</b> загружены <b>технические и функциональные метаданные</b>			✓	
Локальный фрагмент КМД встроен в глобальную КМД, ведомственная карта данных встроена в глобальную карту государственных данных			✓	✓

КРИТЕРИИ	УРОВНИ ЗРЕЛОСТИ ПРОЦЕССА			
	1	2	3	4
Осуществляется регулярный контроль полноты и актуальности метаданных организации				✓
Разработаны и внедрены регламенты управления метаданными				✓

#### **IV. ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ ИИ В РАМКАХ ЗАДАЧ ПО РАБОТЕ С ДАННЫМИ**

Искусственный интеллект (ИИ) - комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение, поиск решений без заранее заданного алгоритма и достижение инсайта) и получать при выполнении конкретных практически значимых задач обработки данных результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека<sup>11</sup>.

Для успешной реализации задач с применением искусственного интеллекта требуются качественные и достаточные данные для обучения моделей машинного обучения и ИИ. Управление данными является основой для успешного функционирования и применения технологий искусственного интеллекта в государственном управлении, которая позволяет сформировать необходимые данные для дальнейшего применения технологий искусственного интеллекта.

В России утверждена Национальная стратегия развития искусственного интеллекта на период до 2030 года, определенная указом Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации». В рамках реализации стратегии принят федеральный проект «Искусственный интеллект».

Реализация методических рекомендаций по управлению данными государственных информационных систем позволяет получить результат в направлении мероприятий федерального проекта по повышению доступности и качества данных, необходимых для для функционирования ИИ на единой цифровой платформе Российской Федерации «ГосТех».

---

<sup>11</sup> ГОСТ Р 59277-2020 Системы искусственного интеллекта. Примечание — Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе, в котором используются методы машинного обучения), процессы и сервисы по обработке данных, анализу и синтезу решений.

## ТЕРМИНЫ И ОПРЕДЕЛЕНИЯ

ТЕРМИН	ОПРЕДЕЛЕНИЕ
АС	Автоматизированная система;
ГИС	Государственная информационная система;
ЗИ	Защита информации
ИС	Информационная система;
ИТ	Информационные технологии
КПЭ	Ключевой показатель эффективности
НСИ	Нормативно-справочная информация;
ОГВ	Государственные органы и внебюджетные фонды
СУБД	Система управления базами данных
ЕИП НСУД	Федеральная государственная информационная система «Единая информационная платформа национальной системы управления данными»;
Актуальность данных	Совокупность характеристик, соответствующих ситуации по срокам или графикам получения, синхронизации или обновления данных;
Аналитика данных	Составное понятие, охватывающее получение, сбор, проверку и обработку данных, включая их количественную оценку, визуализацию и интерпретацию;
Архитектор данных	Аналитик, отвечающий за архитектуру и интеграцию данных на уровне организационной единицы или функционального подразделения. В зависимости от профиля архитекторы данных могут специализироваться на построении хранилищ данных, витрин данных, процессов интеграции и т. п.
Архитектура данных	Модели данных, стандарты на инструменты, соглашения об именовании систем;
Витрина данных НСУД	Комплекс программных и технических средств, обеспечивающий загрузку, хранение и предоставление государственных данных из информационных систем органов и организаций государственного сектора;
Глоссарий ГИС	Перечень терминов и описаний, используемых в процессах, и ГИС ОГВ, сгруппированные по предметным областям, и отношение между ними, обеспечивающее их единообразное понимание, и реализованный на основе семантического слоя ГИС;
Данные	Представление информации в формальном виде, пригодном для передачи, интерпретации или обработки;
Дедуплицированные данные	Набор данных, содержащий только уникальные значения, в котором отсутствуют повторяющиеся значения данных;
Директор по данным	Организационная роль, исполнитель которой отвечает за управление данными на уровне организации;
Домен (домен деятельности)	Область деятельности государственных органов и внебюджетных фондов, принадлежащая одной отрасли экономики и социальной сферы, имеющая общие сегменты (профили) физических или юридических лиц, формируемая с учетом клиентских путей;
Жизненный цикл данных	Последовательность этапов, через которые проходят данные от начального этапа формирования до момента уничтожения;
Информационная система	Совокупность содержащейся в базах данных информации и обеспечивающих ее обработку информационных технологий и технических средств;
Информационные технологии	Процессы, методы поиска, сбора, хранения, обработки, предоставления, распространения информации и способы осуществления таких процессов и методов;
Качество данных	Степень, с которой набор характеристик, присущих данным, отвечает

ТЕРМИН	ОПРЕДЕЛЕНИЕ
	требованиям;
Консистентность данных	Данные, проверенные на соответствие значениям элементов областей или множеств допустимых значений, которые могут определяться как наборы (например, через справочные таблицы) или интервалы допустимых значений, или же через проверочные правила. Области допустимых значений определяются с учетом типа, формата, разрядности и точности/погрешности измерения ожидаемых на входе величин;
Концептуальная модель данных	Представляет собой минимально необходимое множество терминов глоссариев и семантических связей между ними, определяющих всю используемую информацию;
Куратор ГИС	Представитель оператора ГИС;
Куратор данных	Работник ОГВ, наделенный полномочиями по получению данных, по руководству процессом, в результате которого создаются данные, по предоставлению данных, по организации обработки данных;
Руководитель домена	Лицо не ниже заместителя руководителя федерального органа исполнительной власти, являющегося профильным в домене;
Куратор предметной области	Работник ОГВ, уполномоченный выделять ресурсы и организовывать работы по обеспечению качества интегрированных данных в определенной предметной области, и обеспечивающий описание и связанность данных предметной области ОГВ;
Куратор предметной области домена	Работник федерального органа исполнительной власти, являющегося профильным в домене, ответственный за описание и связанность предметной области домена;
Куратор процесса	Работник ОГВ, ответственный за организацию работ и достижение целевых результатов деятельности ОГВ в рамках Процесса;
Логическая модель данных	Перечень сущностей в рамках предметной области, связи между сущностями, их атрибуты и соглашение о порядке именовании атрибутов;
Метаданные	Данные, определяющие и описывающие другие данные;
Модель данных	Описание государственных данных, включая описание состава, структуры и взаимосвязи государственных данных, а также их обладателей;
Модель данных ГИС	Совокупность концептуальной, логической и физической моделей данных ГИС;
Обмен данными	Мероприятия по переносу или копированию данных из одной ИС в другую ИС, передача данных третьим лицам или получение данных от третьих лиц;
Обладатель информации (или Обладатель данных)	Лицо, самостоятельно создавшее информацию либо получившее на основании закона или договора право разрешать или ограничивать доступ к информации, определяемой по каким-либо признакам. Обладателем информации может быть гражданин (физическое лицо), юридическое лицо, Российская Федерация, субъект Российской Федерации, муниципальное образование. От имени Российской Федерации, субъекта Российской Федерации, муниципального образования правомочия обладателя информации осуществляются соответственно государственными органами и органами местного самоуправления в пределах их полномочий, установленных соответствующими нормативными правовыми актами.
Операционная модель	Совокупность участников, ролей, процессов, процедур и регламентов, функционирование которых позволяет достигнуть конечных результатов;
Оператор информационной системы	Гражданин или юридическое лицо, осуществляющие деятельность по эксплуатации информационной системы, в том числе по обработке информации, содержащейся в ее базах данных. Если иное не установлено федеральными законами, оператором информационной



ТЕРМИН	ОПРЕДЕЛЕНИЕ
	системы является собственник используемых для обработки содержащейся в базах данных информации технических средств, который правомерно пользуется такими базами данных, или лицо, с которым этот собственник заключил договор об эксплуатации информационной системы.
Организации	Федеральные органы исполнительной власти, органы исполнительной власти субъектов Российской Федерации, органы местного самоуправления, государственные внебюджетные фонды, а также государственные корпорации, государственные компании, публично-правовые компании и иные организации, наделенные в соответствии с федеральными законами полномочиями по осуществлению от имени Российской Федерации государственного управления в установленной сфере деятельности и обработки информации в информационных системах
Показатель качества данных	Качественная или количественная оценка степени соответствия определенного набора данных критериям качества данных;
Пользователь платформы «ГосТех»	Государственные органы и внебюджетные фонды, иные лица, уполномоченные в соответствии с нормативными правовыми актами Российской Федерации на осуществление мероприятий по созданию, развитию, эксплуатации государственных информационных систем, обеспечивающие создание, развитие, эксплуатацию государственных информационных систем на платформе «ГосТех» и (или) использование цифровых продуктов платформы «ГосТех»
Поставка данных	Передача, в том числе путем предоставления доступа, заказчику или, по заданию заказчика, иному лицу, массивов данных и прав на их использование;
Потребитель данных	Лицо, использующее данные для выполнения стоящих перед ним задач и/или принятия управленческих решений;
Пороговое значение	Допустимое отклонение, используемое для оценки качества данных формы с помощью численного теста.
Функциональная область деятельности домена	Набор данных, описывающий данные, связанные единым функциональным смыслом;
Процесс	Совокупность взаимосвязанных и(или) взаимодействующих видов деятельности, использующих входы для получения намеченного результата
Сервис	Программа из системы программного обеспечения компьютера, которая предоставляет ответы на запросы от других программ данной системы, которые часто расположены на других удаленных связанных компьютерах;
Согласованность данных	Непротиворечивость значений внутри набора данных, отсутствие расхождений в значениях между наборами данных и корректность определения связей по значениям между всеми наборами данных;
Соответствие данных	Степень близости данных к «реальности». Рассчитывается методом сравнения полученных данных с некими эталонными данными из источника, считающегося заведомо достоверным и содержащего достаточно точные данные или надежного источника справочных данных;
Точность данных	Точность соответствия между значением свойства и истинным значением;
Физическая модель данных	Модель, определяющая способы размещения данных в среде хранения и способы доступа к этим данным, которые поддерживаются на физическом уровне;
Целостность (связность) данных	Полнота, точность и согласованность совокупности имеющихся данных. В узком прикладном понимании контроля качества данных под целостностью понимается либо целостность данных на уровне ссылок (наличие во всех парах логически связанных объектов данных общего для обоих объектов ссылочного ключа), либо внутреннюю связность набора данных, то есть отсутствие в нем логических пустот (недостающих элементов).
Субъект (гражданского) права	Гражданин (физическое лицо), юридическое лицо, Российская Федерация, субъект Российской Федерации: республика, край, область, город федерального

ТЕРМИН	ОПРЕДЕЛЕНИЕ
	значения, автономная область, автономный округ, а также городское, сельское поселение и другое муниципальное образование (см. Гражданский кодекс Российской Федерации (часть первая) от 30.11.1994 N 51-ФЗ)
Объект (гражданского) права	Вещи (включая наличные деньги и документарные ценные бумаги), иное имущество, в том числе имущественные права (включая безналичные денежные средства, бездокументарные ценные бумаги, цифровые права); результаты работ и оказания услуг; охраняемые результаты интеллектуальной деятельности и приравненные к ним средства индивидуализации (интеллектуальная собственность); нематериальные блага (см. Гражданский кодекс Российской Федерации (часть первая) от 30.11.1994 N 51-ФЗ)
Структурированные данные	Данные, имеющие стандартизированный формат и одинаковые атрибуты для всех значений данных. Примеры структурированных данных: базы данных, результаты заполнения веб-формы, электронные таблицы, витрины данных.
Неструктурированные данные	Данные, не имеющие стандартизованного формата. Примеры неструктурированных данных: текстовые файлы, видеофайлы.
Полуструктурированные данные	Данные, не имеющие стандартизованного формата, но включающие метаданные, например теги и другие маркеры. Примеры полуструктурированных данных: JSON, XML, веб-файлы.

# ПРИЛОЖЕНИЕ 1. РЕФЕРЕНСНАЯ АРХИТЕКТУРА ХРАНИЛИЩА ДАННЫХ ОРГАНИЗАЦИИ

## РЕФЕРЕНСНАЯ АРХИТЕКТУРА ХРАНИЛИЩА ДАННЫХ ОГВ

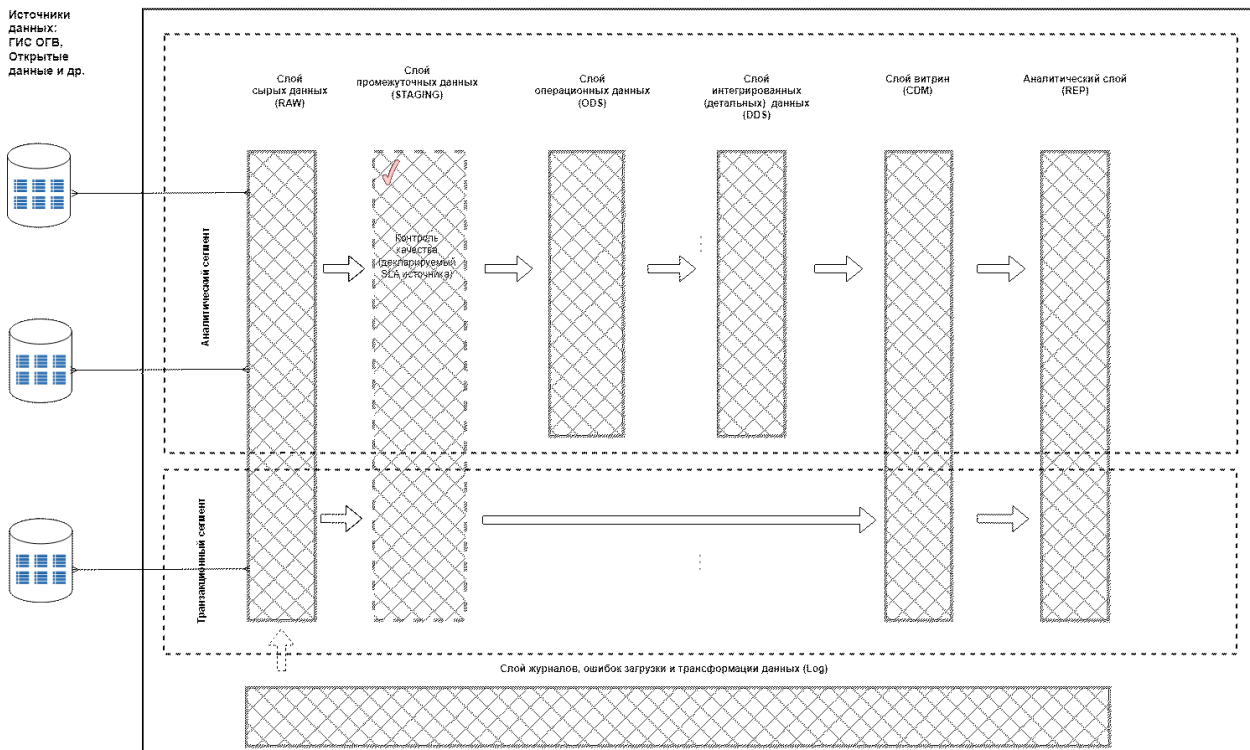


Рис.11. Референсная архитектура хранилища данных организации.

При решении задач организациям целесообразно использовать хранилище данных, в котором реализованы слои работы с данными и два сегмента аналитический и транзакционный. Для реализации типовой обработки данных на уровне хранилища данных организации должны быть сформированы следующие слои данных:

– Слой «сырых данных» RAW предназначен для хранения первичных данных из источников в неизменном виде, логов и других технических данных.

*Назначение:* Сбор данных из источников «как есть», загрузка дельты, преобразование в устойчивый к изменениям формат (формат данных в источнике может меняться).

– Слой операционных данных ODS содержит данные учетных систем, реестров, регистров, справочников, данных из внешних источников без истории. (Операционный слой может содержать данные источника за день, месяц, год и т. д., период хранения зависит от источника и требований потребителя данных. Данные ODS соответствуют декларируемому SLA по качеству данных источника данных. В ODS данные каждого источника представлены в непротиворечивой и связной логической модели данных источника.

*Назначение:* сформировать набор сущностей каждого источника и распределить данные по этим сущностям. Предоставить потребителям данных стандартный интерфейс доступа к операционным данным источников вне зависимости от особенностей хранения.

– Интегрированный (детализированный) слой данных (DDS) – ядро хранилища данных организации, которое обеспечивает:

- устойчивость к изменению требований потребителей данных и изменениям в концептуальной модели данных организации (возникновение новых сущностей, изменение кардинальности связей между сущностями),
- масштабируемость и унификацию (вплоть до автоматизации) построения витрин. DDS наполняется данными из операционного слоя и хранит детальную историю

изменений сущностей логической модели данных организации и связей между ними. Интегрированный (детализированный) слой данных содержит данные: извлеченные из различных источников; представленные в единой и непротиворечивой логической модели данных организации; унифицированные посредством единого НСИ; дедуплицированные и связанные, в том числе и между предметными областями; обладающие измеримым и контролируемым качеством данных. Слой DDS формируется на основе гибких методологий.

*Назначение:* Консолидация данных между различными источниками в рамках единой логической модели данных организации с учетом историчности сущностей, и связей между ними. Источник данных для слоя витрин.

Предоставляет стандартный интерфейс доступа к сущностям логической модели данных организации. Реализация произвольных ad-hoc запросов аналитиков.

– Слой витрин (CDM) - денормализованный слой данных, содержит витрины (подмножества данных, структурированные специально для нужд узкоспециализированных категорий потребителей) для внешних и внутренних потребителей, в том числе витрины данных НСУД и витрины различных внутренних источников данных организации.

CDM - наполняется данными из интегрированного слоя данных. Для обеспечения высокого SLA по доступности данных витрина может наполняться данными напрямую из источника (смотри транзакционный сегмент хранилища данных на схеме), такие витрины будем называть транзакционными.

*Назначение:* Предоставление данных потребителям для анализа в удобном для потребителя контексте и форме (широкая плоская таблица, OLAP-куб), оптимизация доступа на чтение данных из витрин.

– Аналитический слой (REP) содержит агрегированные и/или трансформированные вторичные представления данных (отчеты, статистика, датасеты, агрегаты, результаты работы аналитических сервисов), необходимые для работы и/или являющиеся результатами работы прикладных аналитических сервисов.

*Назначение:* Предназначен для создания и хранения агрегатов, статистики и отчетности, размеченных и неразмеченных датасетов для ИИ.

Слой журналов и ошибок (Log) технологический слой данных содержит логи и ошибки инструментов используемых в хранилище данных организации.

*Назначение:* слой предназначен для восстановления данных и выполнения действий при сбоях программного или аппаратного обеспечения, а также регулярного анализа возникших сбоев.

– Слой промежуточных данных (Staging) – технологический слой данных необходим для промежуточной обработки данных, может находиться между различными основными слоями хранилища данных ОГВ. Необходимость создания таких слоев (их в хранилище может быть несколько) диктуется требованиями к данным к примеру проверка декларируемого SLA по качеству источника данных (см схема ), проверка SLA по качеству потребителя данных после создания витрины данных, фиксация реакции на качество данных, улучшение качества данных, консолидация данных, обогащение поступающих данных технической информацией дата загрузки и код источника. Причин возникновения технологического слоя в архитектуре хранилища может быть много. Состав и количество технологических слоев определяется исходя из разумной целесообразности и требований к обработке данных. Необходимо отметить, что промежуточные слои данных необходимо регулярно «чистить» (удалять накопившиеся за период данные уже переданные в основные слои хранилища).

– *Назначение:* Выполнение промежуточных операций с данными между основными слоями хранилища данных.

## ПРИЛОЖЕНИЕ 2. РЕАЛИЗАЦИЯ ХРАНИЛИЩА ДАННЫХ ОГВ С ИСПОЛЬЗОВАНИЕМ ТЕХНОЛОГИЧЕСКИХ КОМПОНЕНТОВ «ГОСТЕХ»

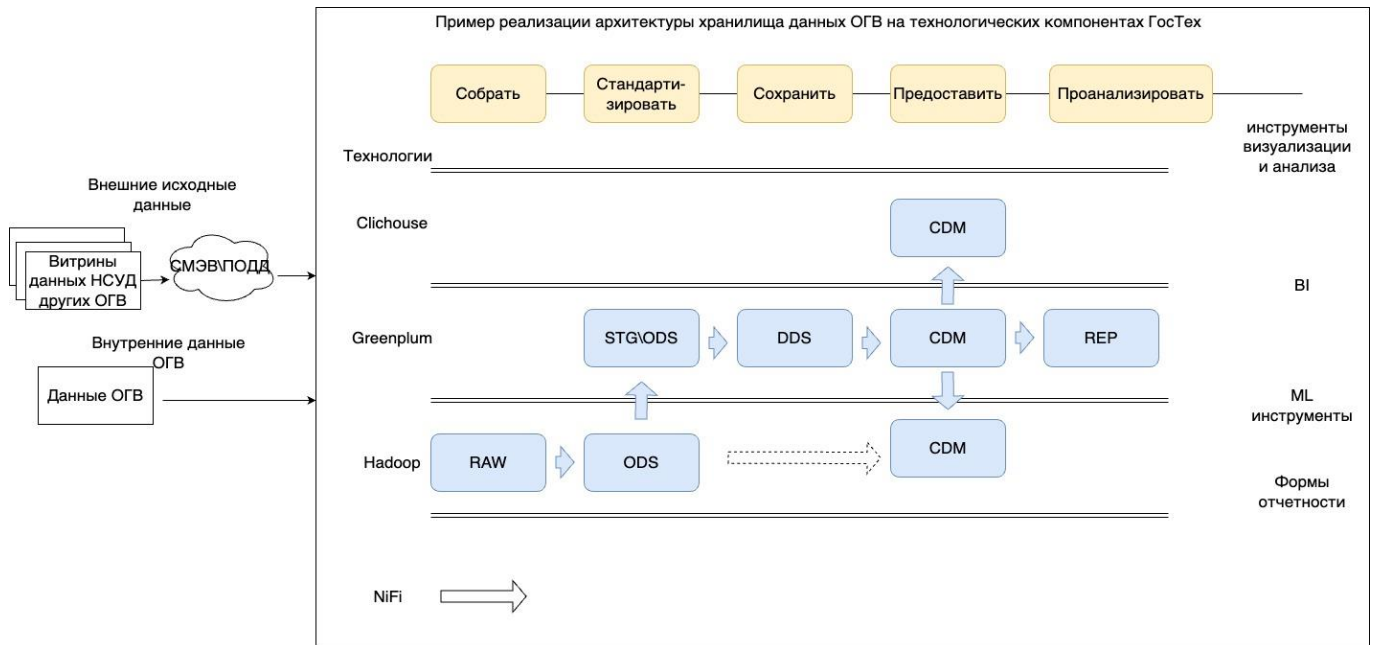


Рис. 12. Пример реализации архитектуры хранилища данных ОГВ с использованием технологических компонентов цифровой платформы «ГосТех».

### ПРИЛОЖЕНИЕ 3. РАСПРЕДЕЛЕНИЕ ОТВЕТСТВЕННОСТИ ЗА ВЫПОЛНЕНИЕ МЕРОПРИЯТИЙ ПО УПРАВЛЕНИЮ ДАННЫМИ НА ПЛАТФОРМЕ ГОСТЕХ

*Таблица 16. Распределение ответственности за реализацию мероприятий по управлению данными на платформе «ГосТех».*

О – Ответственный за реализацию задачи, И – Исполнитель задачи, Э – Эксперт, оказывающий консультационную поддержку, У – Участник, информируемый о принимаемых решениях и ходе выполнения мероприятий, С- участник, Согласующий принимаемые решения по развитию управления данными.

Процессы управления данными	Мероприятия по управлению данными	Заказчик (ОГВ)	Оператор платформы «ГосТех»	Исполнитель (Организации)
Управление жизненным циклом данных	Определение мероприятий, необходимых для реализации управления жизненным циклом данных	У, О	Э	И
	Определение обрабатываемых данных и требований к инфраструктуре для их обработки	У, О	С	И
	Управление созданием / получением данных	О	Э	И
	Определение потребителей данных	О	Э	И
	Управление трансформацией, хранением, использованием данных	У, О	Э	И
	Управление передачей данных	У, О	Э	И
	Управление архивацией и удалением данных	У, О	Э	И
Управление архитектурой данных	Анализ существующей ИТ-архитектуры и архитектуры данных организации	У, О	Э	И
	Документирование	У, О	Э	И

<b>Процессы управления данными</b>	<b>Мероприятия по управлению данными</b>	<b>Заказчик (ОГВ)</b>	<b>Оператор платформы «ГосТех»</b>	<b>Исполнитель (Организации)</b>
	существующей архитектуры данных, создание модели данных организации			
	Оценка текущей архитектуры	<b>У, О</b>	<b>Э, С</b>	<b>И</b>
	Разработка целевой архитектуры данных и планирование перехода к ней	<b>У, О</b>	<b>Э, С</b>	<b>И</b>
	Реализация целевой архитектуры данных	<b>У, О</b>		<b>И</b>
	Сопровождение архитектуры данных	<b>У, О</b>	<b>Э</b>	<b>И</b>
<b>Управление интеграцией данных</b>	Определение мероприятий, необходимых для интеграции данных	<b>О</b>	<b>Э</b>	<b>И</b>
	Учет потоков данных	<b>У, О</b>	<b>Э</b>	<b>И</b>
	Проектирование потоков данных	<b>О</b>	<b>Э, С</b>	<b>И</b>
	Реализация потоков передачи данных	<b>О</b>	<b>Э</b>	<b>И</b>
<b>Управление основными данными и НСИ</b>	Оценка доступности и качества основных данных и НСИ	<b>У, О</b>	<b>Э</b>	<b>И</b>
	Инвентаризация основных и справочных данных	<b>У, О</b>	<b>Э</b>	<b>И</b>
	Интеграция справочных данных с ЕСНСИ	<b>У, О</b>	<b>-</b>	<b>И</b>
	Ведение основных и справочных данных	<b>У, О</b>	<b>-</b>	<b>И</b>
	Нормализация основных и справочных данных	<b>У, О</b>	<b>Э, У</b>	<b>И</b>
	Контроль качества основных и справочных данных	<b>О</b>	<b>Э, У</b>	<b>И</b>

Процессы управления данными	Мероприятия по управлению данными	Заказчик (ОГВ)	Оператор платформы «ГосТех»	Исполнитель (Организации)
<b>Управление качеством данных</b>	Определение контролируемых данных	О	-	И
	Определение требований потребителей к качеству данных	О	-	И
	Создание правил проверки качества данных	О	-	И
	Настройка регулярного запуска алгоритмов проверки качества данных и мониторинг результатов запуска	Э, У, О	-	И
	Создание инцидентов качества данных и управление их жизненным циклом	Э, У, О	-	И
	Регулярная оценка и планирование инициатив по повышению качества данных	Э, У, О	У, Э	И
<b>Обеспечение безопасности данных</b>	Определение мероприятий, необходимых для обеспечения безопасности данных	Э, У, О	Э	И
	Определение требований по безопасности данных	О, У	Э	И
	Выявление и классификация данных, требующих защиты	О, У	Э	И
	Проектирование модели угроз и нарушителя	О, У	С	И
	Проектирование ролевой матрицы доступов к данным	О, У	С	И
	Реализация мер	О, У	-	И



Процессы управления данными	Мероприятия по управлению данными	Заказчик (ОГВ)	Оператор платформы «ГосТех»	Исполнитель (Организации)
	защиты информации			
	Аудит и логирование событий информационной безопасности	О	-	И
	Контроль безопасности данных	О	Э, У	И
<b>Анализ и предоставление данных</b>	Определение целевых пользователей аналитических инструментов	Э, У, О	Э	И
	Сбор требований к аналитике, определение исходных данных и показателей для анализа	О		И
	Разработка макетов отображения данных	О		И
	Сбор данных, проектирование и разработка моделей данных	Э, У, О	С	И
	Реализация аналитических инструментов и формирование аналитики	Э, У, О		И
	Анализ и интерпретация данных	У, О	Э, С	И
	Формирование и ведение реестра аналитической отчетности	О		И
	Регулярная оценка и планирование инициатив по развитию аналитики данных	Э, У, О	У, Э	И
		Определение мероприятий,	Э, У, О	У, Э

<b>Процессы управления данными</b>	<b>Мероприятия по управлению данными</b>	<b>Заказчик (ОГВ)</b>	<b>Оператор платформы «ГосТех»</b>	<b>Исполнитель (Организации)</b>
	необходимых для реализации управления метаданными			
	Ведение каталога данных	Э, У, О	У, Э	И
	Ведение карты данных	Э, У, О	У, Э	И
	Синхронизация метаданных с ЕИП НСУД	Э, У, О	-	И

## **ПРИЛОЖЕНИЕ 4. УПРАВЛЕНИЕ ДАННЫМИ НА ОСНОВЕ ПРЕДМЕТНЫХ ОБЛАСТЕЙ**

### **Описание предметных областей данных органов государственного управления**

Управление данными на платформе «ГосТех» осуществляется на основе предметных областей на трех уровнях, соответствующих универсальному иерархическому принципу, определенному в Политике управления данными:

- Уровень «ГосТех»;
- Уровень Домена;
- Уровень пользователя платформы «ГосТех».
- Предметные области на каждом уровне должны соответствовать следующим условиям:
  - Каждая предметная область может быть отнесена к предметной области более высокого уровня иерархии (например, предметные области Пользователя платформы «ГосТех» включаются в предметные области профильных Доменов);
  - Каждая предметная область должна иметь описанную концептуальную, логическую и физическую модель данных;
  - У предметных областей на каждом уровне должны быть определены кураторы, уполномоченные выделять ресурсы и организовывать работы в части:
    - Обеспечения описания концептуальной, логической и физической модели данных предметной области; обеспечения связности моделей данных предметных областей более низкого уровня иерархии, входящих в данную предметную область
    - Определения основных данных, поддержания в актуальном состоянии описания систем источников и потребителей данных предметной области.
    - Обеспечения связности и качества данных предметной области.
    - Управления метаданными в части моделей данных и глоссария.

#### **Уровень ГосТех**

В целях обеспечения перехода на клиентоцентричный принцип создания и функционирования государственных информационных систем на платформе «ГосТех», на уровне ГосТех предметные области данных соответствуют доменам деятельности (доменам) и включают наборы данных Пользователей платформы «ГосТех», обеспечивающие предоставление ценности для физических и юридических лиц в рамках домена.

Структура предметных областей соответствует перечню доменов, определенному в Методических рекомендациях по проектированию и утверждению целевой архитектуры домена с использованием платформы «ГосТех»

Кураторами предметных областей данных (доменов) на уровне «ГосТех» должны быть назначены отраслевые директоры по данным доменов.

Кураторов предметных областей данных на уровне «ГосТех» назначают Руководители соответствующих доменов, определенные в соответствии с Методическими рекомендациями по проектированию и утверждению целевой архитектуры домена с использованием единой цифровой платформы «ГосТех».

#### **Уровень Домена**

На уровне домена выделяется текущая и целевая структура предметных областей данных. Текущая структура предметных областей данных определяется при формировании домена и соответствует набору описанных предметных областей информационных систем Пользователей платформы «ГосТех», относящихся к домену. Целевая структура предметных областей данных определяется при проектировании домена и соответствует перечню выделенных функциональных

областей в соответствии с Методическими рекомендациями по проектированию и утверждению целевой архитектуры домена с использованием единой цифровой платформы «ГосТех». Переход к управлению данными на основе целевой структуры предметных областей данных осуществляется в соответствии с дорожной картой реализации целевой архитектуры домена, составляемой при проектировании домена согласно Методическим рекомендациям по проектированию и утверждению целевой архитектуры домена с использованием единой цифровой платформы «ГосТех».

Кураторами отдельных предметных областей данных на уровне Домена должны быть назначены уполномоченные сотрудники органов государственной власти.

В общем случае куратором отдельной предметной области данных в домене должен быть назначен сотрудник Пользователя платформы «ГосТех» – куратора домена.

### **Уровень Пользователя платформы «ГосТех»**

На уровне Пользователя платформы «ГосТех» структура предметных областей данных определяется директором по данным Пользователя платформы «ГосТех» и соответствует перечню описанных предметных областей данных информационных систем Пользователя платформы «ГосТех».

Для каждой предметной области данных Пользователя платформы «ГосТех» должны быть определены связанные предметные области данных на уровне Домена и уровне ГосТех, а также описаны концептуальная, логическая и физическая модель данных.

Кураторами предметных областей данных на уровне Пользователя платформы «ГосТех» назначаются соответствующим приказом работники Пользователя платформы «ГосТех». Куратор предметной области данных Пользователя платформы «ГосТех» назначается директором по данным Пользователя платформы «ГосТех».