

НАЦИОНАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ ДАННЫМИ  
ЕДИНЫЕ ТРЕБОВАНИЯ ПО УПРАВЛЕНИЮ ГОСУДАРСТВЕННЫМИ  
ДАННЫМИ

Часть 2. Требования по гармонизации и систематизации государственных данных

**Книга 2. Требования по систематизации государственных данных**

**Версия 1.0**

## СОДЕРЖАНИЕ

1.	ОБЩИЕ СВЕДЕНИЯ .....	5
1.1.	Нормативная база.....	5
1.2.	Назначение документа .....	5
1.3.	Общие положения.....	5
2.	ОПРЕДЕЛЕНИЕ КАТЕГОРИЙ ЗНАЧИМОСТИ.....	6
2.1.	Категории значимости ГИР .....	6
2.2.	Критерии и порядок категорирования.....	6
2.2.1	Определение количества и наименований видов государственных данных, входящих в состав ГИР .....	8
2.2.2	Определение уровня значимости видов государственных данных, содержащихся в ГИР .....	8
2.2.3	Определение категории значимости ГИР .....	8
3.	КЛАССИФИКАЦИЯ ГОСУДАРСТВЕННЫХ ДАННЫХ.....	8
4.	УНИФИКАЦИЯ СТРУКТУРНЫХ ЭЛЕМЕНТОВ ГОСУДАРСТВЕННЫХ ДАННЫХ И ФОРМАТОВ ИХ ПРЕДСТАВЛЕНИЯ.....	9
5.	ОНТОЛОГИЧЕСКИЙ ПОДХОД К УПРАВЛЕНИЮ ГОСУДАРСТВЕННЫМИ ДАННЫМИ.....	10
5.1.	Использование онтологического подхода в создании МГД.....	13
5.2.	Результат проектирования МГД .....	14
5.3.	Описание целевого состояния МГД .....	16

**ИЗМЕНЕНИЯ**

<b>Версия</b>	<b>Дата</b>	<b>Автор</b>	<b>Изменения</b>
1.0	07.11.2019	Карпенко А.Н. Воронов Ю.А.	Исходный документ

**ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ**

<b>Сокращение</b>	<b>Описание</b>
ГИР	Государственный (муниципальный) информационный ресурс
НПА	Нормативный правовой акт
НСУД	Национальная система управления данными
МГД	Модель государственных данных
РВД	Реестр видов государственных данных

## **1. ОБЩИЕ СВЕДЕНИЯ**

### **1.1. Нормативная база**

Требования по систематизации государственных данных (далее — требования) разработаны с учетом положений следующих нормативных правовых актов:

- Концепции создания и функционирования национальной системы управления данными, утвержденной распоряжением Правительства Российской Федерации от 3 июня 2019 г.;
- постановления Правительства Российской Федерации № \_\_\_\_ от \_\_\_\_ ноября 2019 г. «О формировании и ведении реестра видов государственных данных».

### **1.2. Назначение документа**

Документ содержит комплекс требований, направленных на обеспечение систематизации государственных данных участниками НСУД.

### **1.3. Общие положения**

Систематизация государственных данных (в контексте Требований) должна достигаться за счет выполнения комплекса требований в части:

- отнесения государственных информационных ресурсов к определенным категориям значимости с целью дифференциации требований к управлению государственными данными;
- отнесения государственных данных к определенному классу в зависимости от их специфики;
- унификации структурных элементов государственных данных и форматов их представления;
- применения онтологического подхода для управления государственными данными.

## **2. ОПРЕДЕЛЕНИЕ КАТЕГОРИЙ ЗНАЧИМОСТИ**

### **2.1. Категории значимости ГИР**

Категорирование ГИР осуществляется в целях дифференциации требований к управлению государственными (муниципальными) данными и представляет собой установление соответствия ГИР критериям значимости ГИР и содержащихся в них видов государственных (муниципальных) данных, присвоение таким ГИР одной из категорий значимости.

Устанавливаются три категории значимости ГИР: первая (наивысшая), вторая и третья.

### **2.2. Критерии и порядок категорирования**

Если иное не определено федеральным законодательством, операторы ГИР присваивают одну из трех категорий значимости ГИР, операторами которых они являются.

Категорирование осуществляется оператором ГИР, исходя из уровня значимости видов государственных данных, содержащихся в ГИР, в том числе:

1) социальной значимости, выражающейся в оценке возможного ущерба, причиняемого жизни или здоровью людей, возможности прекращения или нарушения функционирования объектов обеспечения жизнедеятельности населения, транспортной инфраструктуры, сетей связи, а также максимальном времени отсутствия доступа к государственной услуге для получателей такой услуги;

2) политической значимости, выражающейся в оценке возможного причинения ущерба интересам Российской Федерации в вопросах внутренней и внешней политики;

3) экономической значимости, выражающейся в оценке возможного причинения прямого и косвенного ущерба юридическим лицам, физическим лицам, бюджетам Российской Федерации;

4) экологической значимости, выражающейся в оценке уровня воздействия на окружающую среду;

5) значимости вида государственных данных для обеспечения обороны страны, безопасности государства и правопорядка.

При категорировании вида государственных данных рекомендуется использовать показатели и критерии, представленные в постановлении Правительства Российской Федерации от 8 февраля 2018 г. № 127 «Об утверждении

Правил категорирования объектов критической информационной инфраструктуры Российской Федерации, а также перечня показателей критериев значимости объектов критической информационной инфраструктуры Российской Федерации и их значений».

Кроме того, категорирование видов государственных данных осуществляется с учетом следующих показателей и критериев значимости видов государственных данных, формируемых с использованием ГИР:

Показатель	Значения показателя		
	III категория	II категория	I категория
<b>1. Востребованность</b>			
1.1 Суммарное количество фактических получателей видов государственных данных	не более 20	более 20, но не более 100	более 100
1.2 Суммарное количество запросов на предоставление видов государственных данных всеми получателями (в год)	Значения показателей должны быть уточнены по количеству запросов (например, за год или полугодие) к виду сведений в СМЭВ, соответствующему виду государственных по смыслу и составу показателей (атрибутов)		
<b>2. Масштабность</b>			
2.1 Уровень межведомственного использования	Муниципальный	Региональный	Федеральный

Категорирование ГИР осуществляется в следующем порядке:

- 1) определение количества и наименований видов государственных данных, входящих в состав ГИР;
- 2) определение уровня значимости видов государственных данных, содержащихся в ГИР;
- 3) определение категории значимости ГИР.

После завершения категорирования оператор ГИР принимает правовой акт, которым присваивает ГИР ту или иную категорию значимости, и направляет копию указанного акта в Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации.

### **2.2.1 Определение количества и наименований видов государственных данных, входящих в состав ГИР**

С целью последующего определения категории значимости ГИР, содержащего несколько видов государственных данных, оператор ГИР определяет количественный состав таких видов государственных данных.

### **2.2.2 Определение уровня значимости видов государственных данных, содержащихся в ГИР**

Для каждого вида государственных данных, входящих в состав ГИР, оператор ГИР в соответствии с определенными критериями определяет уровень значимости по каждому из заданных показателей.

В случае если какой-либо из заданных показателей уровня значимости неприменим для вида государственных данных, уровень значимости по соответствующему показателю не определяется.

Для каждого показателя значимости, для которого установлено более одного значения такого показателя (территория, количество людей), оценка производится по каждому из значений показателя значимости, а уровень значимости присваивается по наивысшему значению такого показателя.

В случае если вид государственных данных по одному из показателей значимости отнесен к первому уровню значимости, расчет по остальным показателям уровня значимости не проводится.

### **2.2.3 Определение категории значимости ГИР**

Категория значимости ГИР определяется в соответствии с уровнем значимости вида государственных данных, который в него входит.

При вхождении в ГИР нескольких видов государственных данных категория значимости ГИР должна определяться по виду государственных данных, имеющему наиболее высокий уровень значимости.

## **3. КЛАССИФИКАЦИЯ ГОСУДАРСТВЕННЫХ ДАННЫХ**

Государственные данные, входящие в состав ГИР, с целью их систематизации должны классифицироваться по следующим классификационным группам.

1) Уровень межведомственного использования данных ГИР:

- федеральный;
- региональный;
- муниципальный;
- данные на межведомственном уровне не используются.

2) Форма данных:

- первичные данные;
- агрегированные данные.

3) Уровень доступности данных:

- ограниченный доступ (персональные данные, данные, содержащие коммерческую тайну, служебную тайну, профессиональную тайну, тайну следствия и судопроизводства, сведения о сущности изобретения и/или иная конфиденциальная информация);
- доступ без ограничений (данные не содержат конфиденциальной информации).

4) Эталонность:

- эталонные;
- неэталонные.

5) Идентифицируемость:

- идентификатор (по значению показателя данных однозначно могут быть получены сведения по определенному объекту учета);
- не идентификатор.

#### **4. УНИФИКАЦИЯ СТРУКТУРНЫХ ЭЛЕМЕНТОВ ГОСУДАРСТВЕННЫХ ДАННЫХ И ФОРМАТОВ ИХ ПРЕДСТАВЛЕНИЯ**

Унификация структурных элементов государственных данных должна осуществляться операторами ГИР по следующим направлениям:

- использование указателей на эталонные государственные (муниципальные) данные вместо их воспроизведения в полном объеме;
- приведение наименования государственных данных на русском языке, указываемого в паспорте ГИР, в соответствие с наименованием тех же государственных данных, содержащихся в ГИР с эталонными государственными данными;

— приведение формата представления государственных данных в соответствие с форматами представления тех же государственных данных, содержащихся в ГИР с эталонными государственными данными.

В случае невозможности унификации наименования государственных данных, обусловленной технологической спецификой ГИР и требованиями НПА, оператор ГИР должен сформировать и представить для размещения в НСУД сведения, позволяющие установить соответствие структурных элементов государственных данных с эталонными государственными данными по следующей форме.

Наименования и коды ГИР	Технологическое наименование государственных данных	Наименование государственных данных на русском языке
Наименование (код в НСУД) ГИР, содержащего эталонными данными		
Наименование (код в НСУД) ГИР		

Сведения, представленные операторами ГИР по всем структурным элементам государственных данных, относящимся к неэталонным государственным данным, интегрируются в рамках НСУД в единый синонимический словарь наименований государственных данных.

Такой словарь должен содержать наименования государственных данных, содержащихся в ГИР, где такие данные имеют статус эталонных, и соответствующие им наименования государственных данных, содержащихся в других ГИР.

## **5. ОНТОЛОГИЧЕСКИЙ ПОДХОД К УПРАВЛЕНИЮ ГОСУДАРСТВЕННЫМИ ДАННЫМИ**

Нормативно закрепленное в Концепции национальной системы управления данными, утвержденной распоряжением Правительства Российской Федерации от 3 июня 2019 г. № 1189-р, требование онтологического единства государственных данных обеспечивается введением единого онтологического подхода при построении НСУД.

Онтологический подход выполняет задачу формального, подробного и систематизированного описания системы понятий в области управления

государственными данными с целью отражения семантики (смыслового содержания) этой области в структуре и составе государственных данных, содержащихся в государственных информационных ресурсах.

Онтологический подход обеспечивает эффективное человеко-машинное взаимодействие, предоставляя для работы с государственными данными единый понятийный аппарат, используемый как пользователями НСУД, так и информационно-технологическими элементами НСУД. При этом понятия реального мира адекватно отражаются в онтологии, представляя всем пользователям НСУД, а не только ИТ-специалистам, весь необходимый набор сведений для работы с государственными данными на русском языке.

При этом подходе модель государственных данных имеет в своем составе два взгляда, отражающие одну и ту же реальность государственного управления:

- онтологический взгляд, который понятен и обычному пользователю, и компьютеру, дает преимущества обращения к информационной системе на обычном языке, предметная область (государственное управление) описывается понятным исчерпывающим образом, пользователю точнее передаются смыслы, проще и легче происходит адаптация модели к неизбежным изменениям в сфере государственного управления;
- информационно-технологический взгляд, который понятен только компьютеру и ИТ-специалисту; польза состоит в том, что при этом рационально используется весь исторически наработанный массив данных в унаследованных государственных информационных ресурсах и системах.

По мере развития модели государственных данных и погружения исторически наработанного массива данных в онтологию оба подхода будут сближаться до полного объединения по принципу конвергенции технологий.

В рамках архитектуры НСУД онтологический подход обеспечивается следующими средствами:

- методология концептуального проектирования организационных и информационных систем управления;
- методология онтологического моделирования предметных областей информатизации;
- стек семантических технологий Консорциума W3C;
- инструментальное ПО на базе математического аппарата для синтеза концептуальных схем и формализованного определения понятий и связей;

- технология человеко-машинного воплощения онтологических схем;
- технология правового, организационного, семантического, онтологического синтеза;
- библиотека онтологических моделей;
- научно-методическое и учебное обеспечение онтологического подхода.

Онтология государственных данных включает:

- структуру, классификацию и типизацию основных сущностей;
- связи между концептами онтологии;
- контекст деятельности;
- правила (аксиомы, ограничения), принятые в области управления государственными данными.

Модель государственных данных, построенная на базе онтологии, кроме реализации принципов создания НСУД, обеспечит следующие преимущества:

- пользователи могут обращаться к МГД на обычном языке при поиске любых сведений, касающихся государственных данных; при этом однозначно передается смысл всех понятий в сфере цифрового государственного управления;
- предметная область (государственное управление) представляется любому пользователю НСУД понятным, точным и исчерпывающим образом;
- проще и легче происходит адаптация МГД к неизбежным изменениям в сфере государственного управления, средствами НСУД реализуется возможность быстро и с малой трудоемкостью вносить необходимые изменения в МГД с одновременным обеспечением:
  - требуемого качества данных, включая сохранение их целостности, непротиворечивости и актуальности;
  - удобства внесения в МГД новых объектов, типов данных и функций;
  - в перспективе создается основа для гарантированного получения доверенных сведений (данных, информации и знаний) о любом объекте государственного управления в любое время и в любом месте любым авторизованным пользователем НСУД.

### 5.1. Использование онтологического подхода в создании МГД

Являясь частью национальной программы «Цифровая экономика Российской Федерации» и Федерального проекта «Цифровое государственное управление», НСУД входит также в состав государственного регулирования процессов систематизации и гармонизации информации в Российской Федерации (изменения в Федеральный закон «Об информации, информационных технологиях и защите информации», статья 12.2 «Особенности государственного регулирования в сфере систематизации и гармонизации информации в Российской Федерации»), регулируя сферу управления государственными данными.

Онтологический подход в НСУД выполняет задачу формального, подробного и систематизированного описания области управления государственными данными с целью построения модели государственных данных, адекватной реальности.

В ходе проектирования формируются две части, взаимно дополняющие друг друга:

- 1) онтологическая часть МГД;
- 2) информационно-технологическая часть МГД (схема данных).

Онтологическая часть МГД заменяет собой концептуальный уровень, принятый в общепринятом подходе, обогащая и развивая его:

- семантикой (уточненный смысл понятий по результатам реконструкции, оценки и реинжиниринга данных);
- возможностями навигации по данным («карта данных» в форме графа, содержащего структуру онтологии из концептов и связей между ними);
- функциями единого толкового словаря-гlossария-тезауруса по предметным областям, охватываемым МГД;
- возможностями логического вывода (*semantic reasoner*).

Онтологическая часть МГД понимается как людьми, так и системами, служит смысловым навигатором для обычных пользователей и машиночитаемой моделью для информационных систем.

Схема данных включает:

- логическую модель данных, которая является совокупностью логических моделей ведомственных витрин данных и описывает структуру сведений, которые содержатся в витринах данных отдельных информационных ресурсов; каждая логическая модель представляет собой частный срез соответствующей части онтологической модели;

— физическую модель данных, которая описывает реализацию объектов логической модели на уровне конкретных баз данных ведомств.

Схема данных связана с РВД (продукт НСУД «Реестр видов государственных данных»).

Схема данных рассчитана на понимание ИТ-специалистами и является машиночитаемой частью МГД.

## **5.2. Результат проектирования МГД**

Результатами проектирования модели государственных данных являются следующие артефакты:

- онтологическая часть МГД; описание онтологической части осуществляется в соответствии с Соглашением о моделировании НСУД;
- информационно-технологическая часть МГД (схема данных); описание схемы данных осуществляется в соответствии с Соглашением о моделировании НСУД;
- единый тезаурус в сфере управления государственными данными (вводит в сферу управления государственными данными единую систему терминов и определений), представляется в форме документа;
- научно-практическое обоснование гармонизации и согласования понятий в сфере управления государственными данными в рамках единой онтологии с целью формирования концептуальных, логических и физических моделей данных государственных информационных ресурсов, представляется в форме документа.

Онтологическая часть МГД представляет собой единую онтологию государственных данных в форме ориентированного графа, описывающего:

- структуру, классификацию и типизацию всех необходимых и достаточных сущностей предметной области (далее — концептов онтологии);
- все необходимые и достаточные концепты онтологии, их семантику и атрибутивный состав;
- связи и отношения между концептами онтологии;
- контексты деятельности, в которые вовлечены концепты онтологии;
- аксиомы и ограничения, принятые в области управления государственными данными;

— правила логического вывода заключений на базе заданных аксиом, исполняемые посредством машины логических выводов (*semantic reasoner*);

— деловые правила (*business rules*), в том числе правила описания концептов, связей и отношений, технологические правила в МГД.

В процессе формирования архитектуры МГД от исходного, через транзитное до целевого состояния, состав онтологической части МГД изменяется:

1) исходная и транзитная архитектура онтологической части МГД включает в себя семантический и концептуальный уровни;

2) целевая архитектура онтологической части МГД содержит только семантический уровень, а именно, единую онтологию государственных данных, которая объединяет в себе и логический, и концептуальный уровни, гармонизированные в процессе улучшения качества государственных данных в ходе формирования МГД в период с 2020 по 2023 годы.

Семантический уровень исходной архитектуры МГД содержит ограниченную экспериментом онтологию, частично устраняющую ведомственную разобщенность; семантический уровень целевой архитектуры МГД отражает модель государственных данных цифровой экономики, в котором ведомственные данные интегрированы в онтологию и являются эталонными.

Концептуальный уровень МГД представляет собой слой объединения онтологии и схемы данных, является зоной конвергенции проектных подходов «сверху» и «снизу» на пути от эксперимента к целевой архитектуре. Исходный концептуальный уровень содержит результат аналитической работы по согласованию негармонизированных данных логического уровня с онтологическими моделями данных в рамках эксперимента, формируя исходную версию МГД. В целевой архитектуре концептуальный уровень становится не нужен, он интегрируется в онтологию государственных данных.

Управление концептуальным уровнем в исходной архитектуре МГД осуществляется средствами подсистемы ведения метаданных ЕИП НСУД, которая предоставляет следующие возможности:

1) ведение бизнес-словаря;

2) управление концептуальной моделью данных в части описания сущностей и их отношений в терминах бизнес-словаря, в том числе:

— описания предназначения сущностей,

— отображения систем, порождающих сущность (эталонных систем), в том числе используя реестр информационных ресурсов в составе РВД;

— отображения систем, использующих сущность, в том числе в соответствии с реестром информационных ресурсов в составе РВД;

- 3) управление доменами сущностей концептуальной модели;
- 4) ведение реестра сервисов в терминах концептуальной модели;
- 5) ведение реестра отчетов в терминах концептуальной модели.

Организационные правила управления государственными данными (например, порядки и регламенты, включая порядок представления государственных данных и порядок доступа к ним) не входят в МГД, но должны быть представлены в форме исполняемых бизнес-процессов в нотации *BPML*, связанных с онтологией государственных данных. Описание бизнес-процессов в нотации *BPML* осуществляется в соответствии с Соглашением о моделировании НСУД.

### 5.3. Описание целевого состояния МГД

В связи с инновационным характером разработки МГД необходимо задать основные ориентиры «образа будущего» для МГД, который планируется достичь в ходе реализации проекта НСУД.

Предлагаются три основных ориентира:

1) предполагается, что в пределе (в перспективе, за рамками дорожной карты НСУД до 2024 года) МГД будет содержать в себе «цифровой образ» реального мира, отражающий аспекты единого информационного пространства государства, необходимые и достаточные для его оптимального функционирования и развития в целях построения информационного общества Российской Федерации;

2) этот цифровой образ подлежит постоянной целенаправленной оптимизации инструментами НСУД;

3) в ходе оптимизации обеспечивается обратная связь, вырабатывается и поддерживается инструментами НСУД и МГД необходимый баланс информационных ресурсов, необходимых для функционирования информационного общества Российской Федерации.

При этом:

— формируется цифровой образ реального мира (модели объектов-субъектов, процессов, данных и отношений между ними) на том уровне

детализации, который необходим и достаточен, с точки зрения экспертного сообщества, для функционирования единого информационного пространства Российской Федерации;

— цифровой образ реального мира выражается в форме единой онтологии, в которой выделяются ключевые сущности и процессы; основной ключевой сущностью объявляется концепт онтологии «физическое лицо (гражданин Российской Федерации)»; основным ключевым процессом объявляется концепт онтологии «процесс экономического развития», имеющий целью обеспечить благополучие и благосостояние граждан Российской Федерации;

— главным критерием успешности деятельности государства объявляется высокое и растущее качество жизни граждан Российской Федерации;

— единая онтология подвергается процедурам оптимизации с целью достижения необходимого и достаточного множества сущностей и процессов, управление которыми со стороны государства необходимо для успешного функционирования и развития информационного общества Российской Федерации; при этом ведется постоянная синхронизация цифрового образа и реального мира, то есть осуществляются целенаправленные реформы.

В итоге обратная связь между изменениями в реальном мире и его цифровым образом становится систематической за счет оптимизации соответствующих алгоритмов, сформированных в рамках экосистемы НСУД, с учетом реализации национальной программы «Цифровая экономика Российской Федерации» и других государственных программ.