



Практики применения модели данных в государственном секторе и бизнесе:

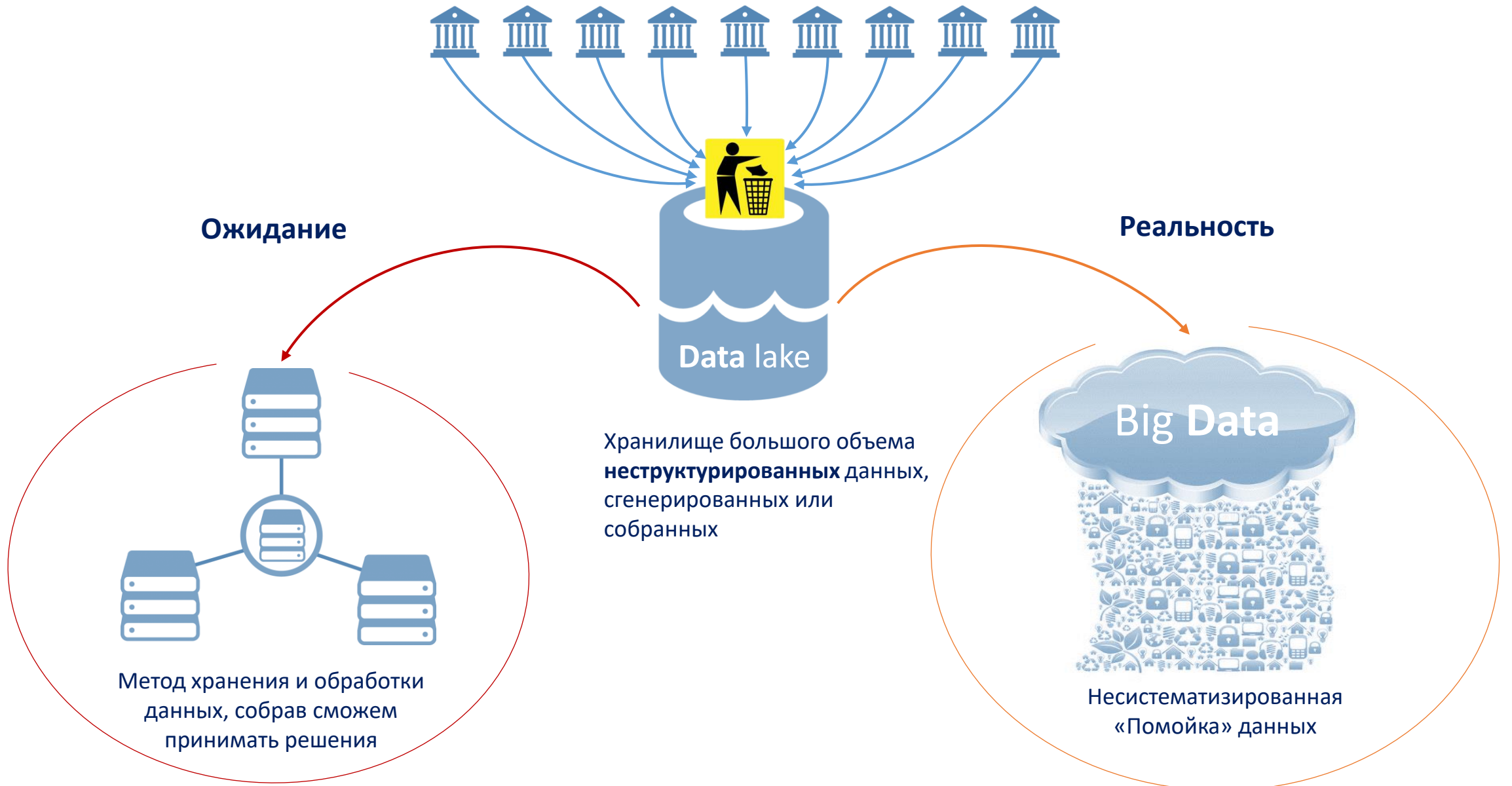
подходы к управлению данными и прикладные технологии

МАЛАХОВ АЛЕКСАНДР АНДРЕЕВИЧ

Начальник управления методологического
сопровождения работы с данными

Аналитического центра при Правительстве РФ

ЧТО ПЛАВАЕТ В ОЗЕРЕ ДАННЫХ?

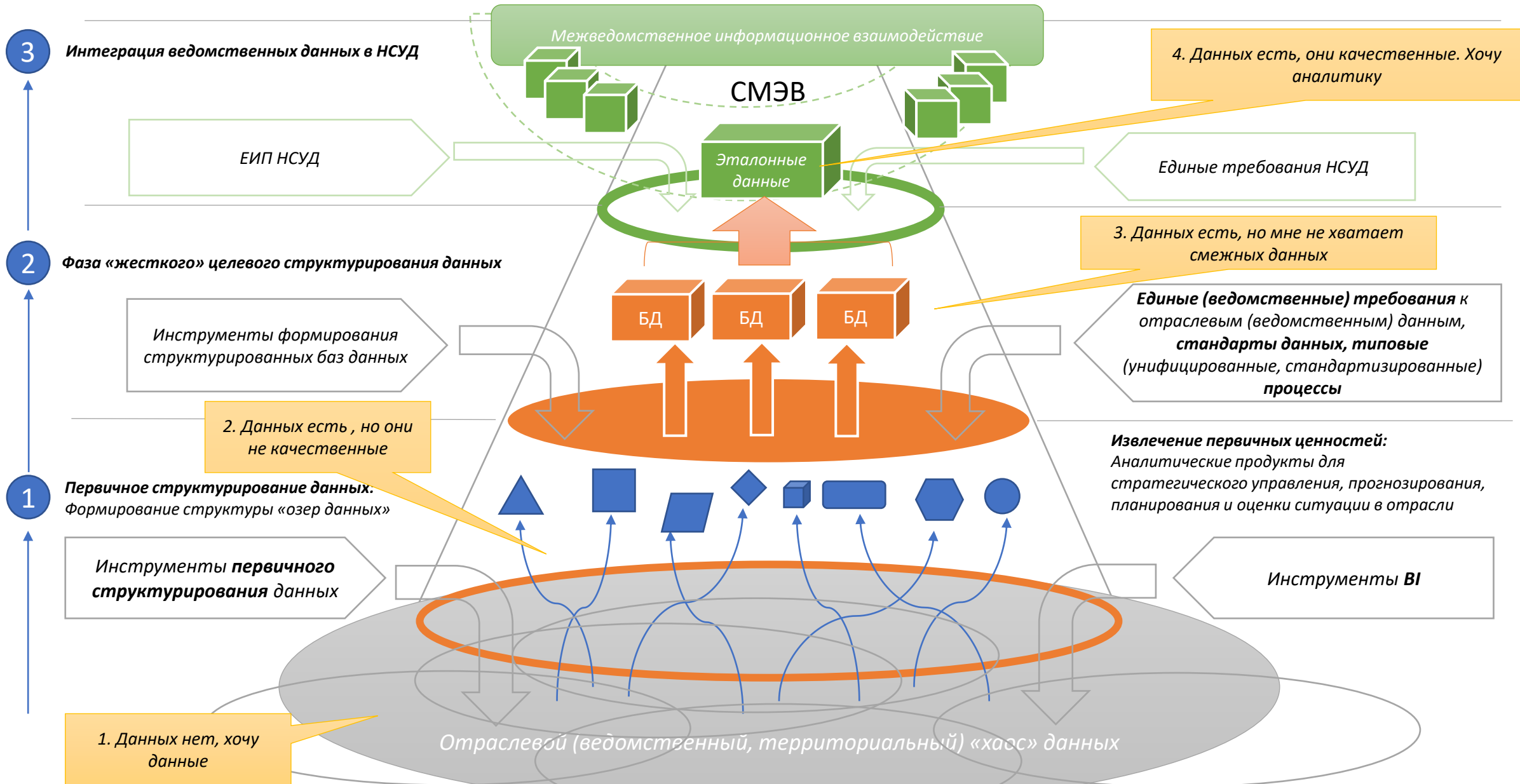


МОДЕЛЬ ЗРЕЛОСТИ ЦИФРОВОГО ГОСУДАРСТВА

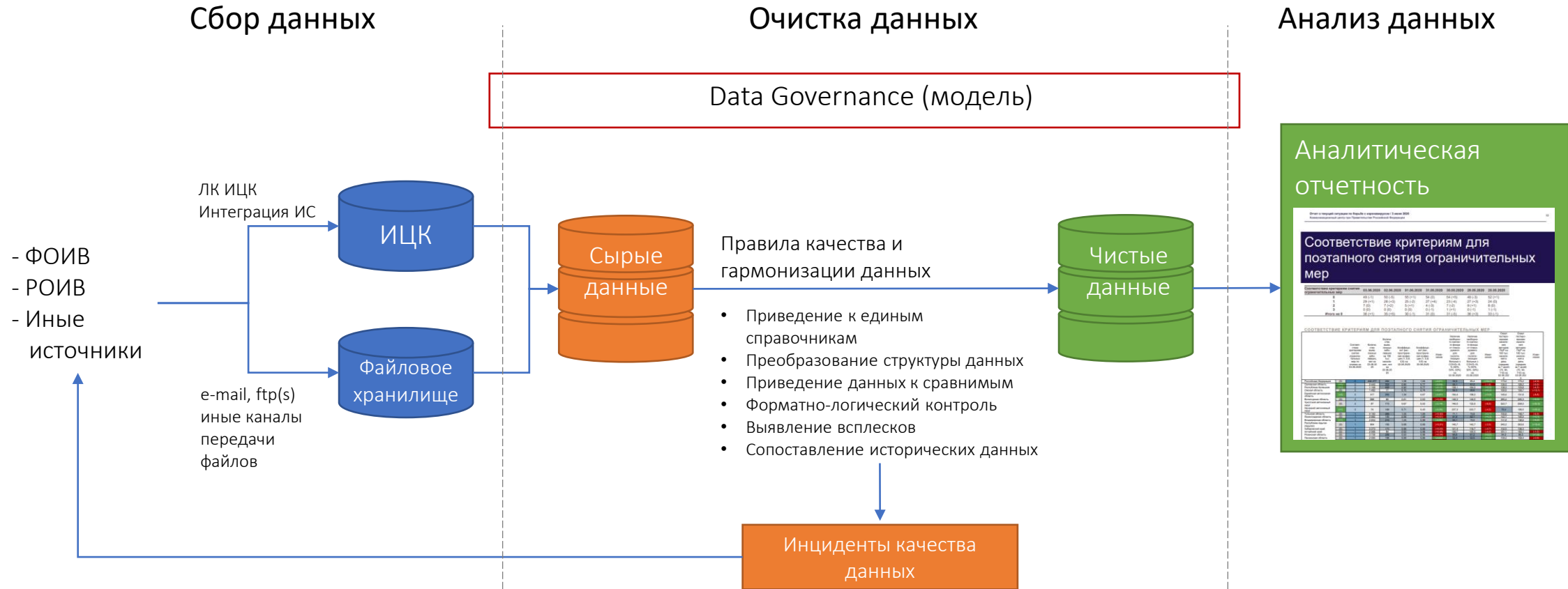
	E-Government		Open		Data-Centric		Fully Digital		Smart	
	01 Initial	02 Developing	03 Defined	04 Managed	05 Optimizing					
Value Focus	Compliance	Transparency	Constituent Value	Insight-Driven Transformation	Sustainability					
Service Model	Reactive	Intermediated	Proactive	Embedded	Predictive					
Platform	IT-Centric	Customer-Centric	Data-Centric	Thing-Centric	Ecosystem-Centric					
Ecosystem	Government-Centric	Service Co-creation	Aware	Engaged	Evolving					
Leadership	Technology	Data	Business	Information	Innovation					
Technology Focus	SOA	API Management	Open Any Data	Modularity	Intelligence					
Key Metrics	% of Services Online	Number of Open Datasets	% Improvement in Outcomes, KPIs	% of New and Retired Services	Number of New Service Delivery Models					

Source: Gartner

МОДЕЛЬ РАЗВИТИЯ СИСТЕМЫ УПРАВЛЕНИЯ ДАННЫМИ



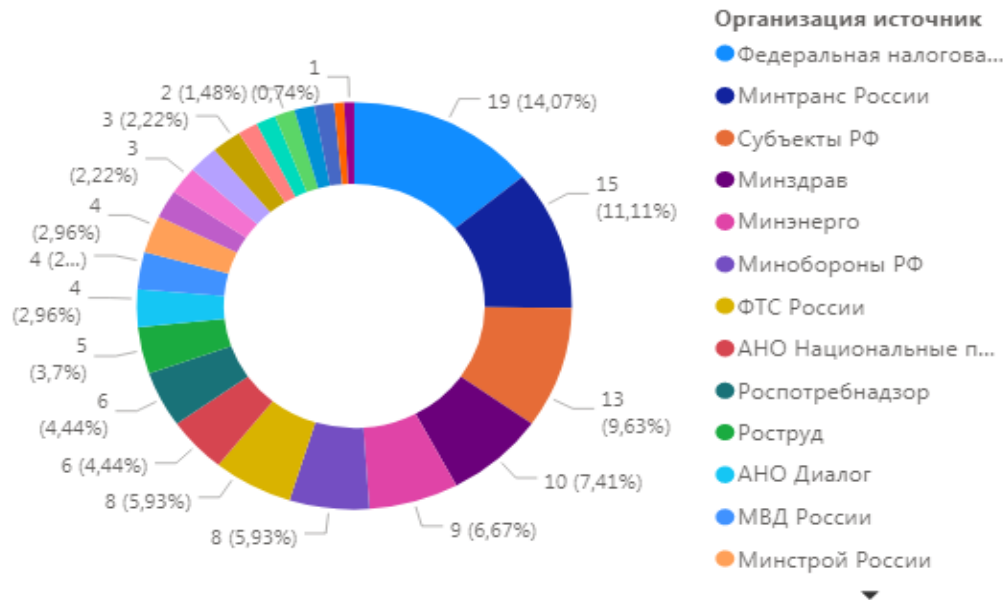
СБОР ДАННЫХ



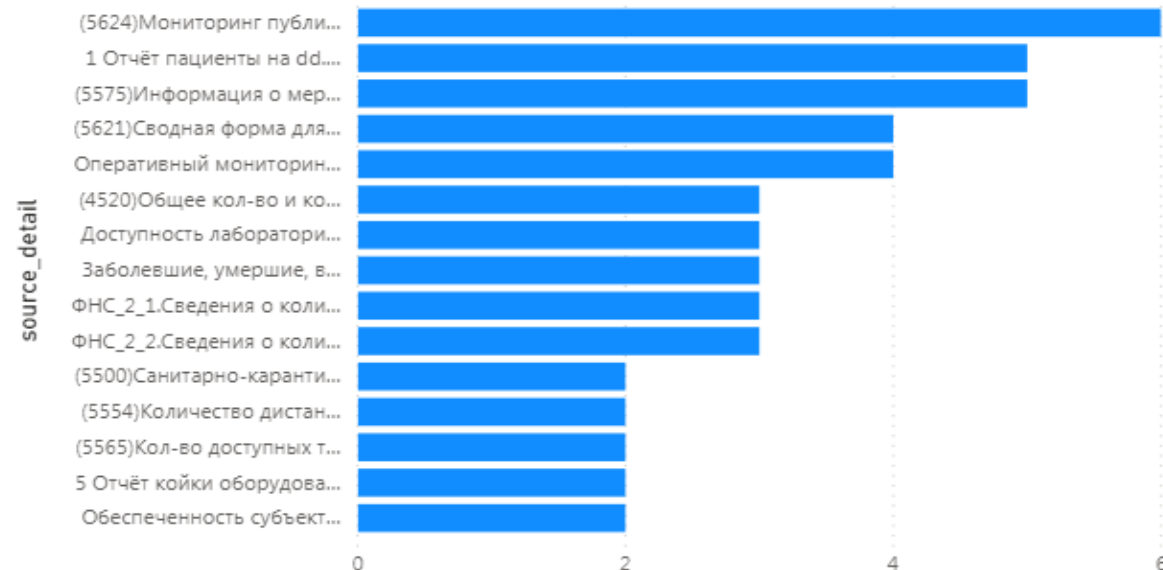
Необходимо преобразовать разрозненные данные в единый массив гармонизированных, качественных и сравнимых данных на основе которых можно формировать управленческие решения

ЭТАП 2. ГЛОССАРИЙ ДАННЫХ: АНАЛИТИКА ПО ПОКАЗАТЕЛЯМ

Количество Показателей по ФОИВ



Количество Показателей по Формам сбора

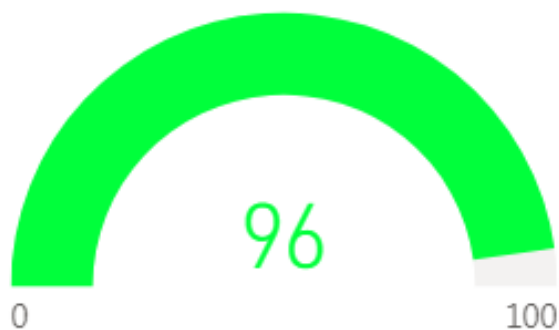


display_name	description	data_source	source_detail	data_quality
Число всех случаев заболевания пневмонией		Excel	1 Отчёт пациенты на dd.mm.yyyy.xlsx	Значения нарастающим итогом >= предыдущему отчёту; Сумма показателей итого по региону, входящих в округ = итого по округу, сумма показателя по регионам=итого по РФ; сумма показателей итого по округам=итого по рф;
76_2.Минтранс.Количество и объем продаж авиабилетов	1. Количество авиабилетов: - международных воздушных линий - внутренних воздушных линий 2. Объем продаж авиабилетов: - международных воздушных линий - внутренних воздушных линий	Excel	(5536)Количество и объем продаж авиабилетов (Минтранс 76_2)	1. Проверка на обязательность заполнения атрибута 2. Проверка на дублирование данных
76_1.Минтранс.Количество и объем продаж	1. Количество железнодорожных билетов: - дальнего следования	Excel	(5535)Количество и объем продаж железнодорожных (Минтранс 76_1)	1. Проверка на обязательность заполнения атрибута

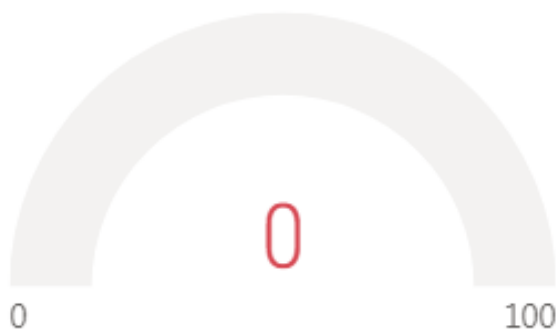
ЭТАП 2. ГЛОССАРИЙ ДАННЫХ: АНАЛИТИКА ПО ПОКАЗАТЕЛЯМ

Анализ качества показателей для мониторинга Covid-19

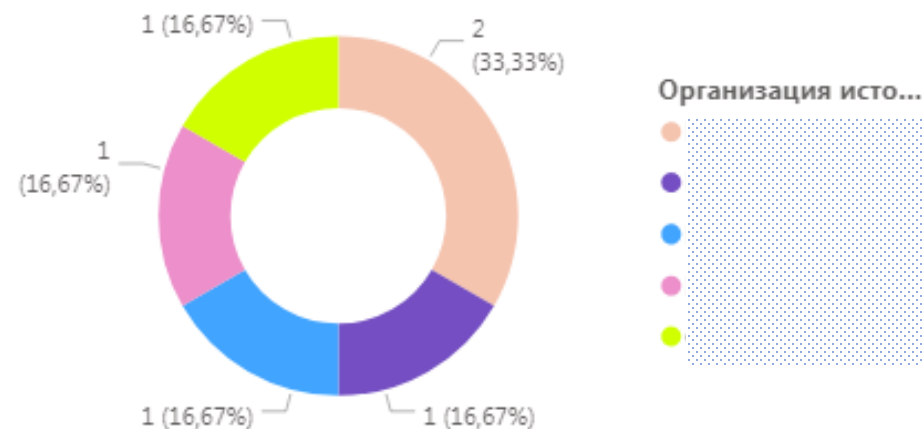
Полнота DQ (%)



Полнота Ответ-ный (%)



Должники по описанию DQ



display_name

Организация источник

data_quality

Потребность средств индивидуальной защиты и дезинфицирующих средств в Минобороны

Минобороны РФ

75.Минтранс.Индекс тарифов на услуги пассажирского транспорта (железнодорожного, воздушного, в разрезе субъектов РФ), % к предыдущему периоду

Минтранс России

Мониторинг фактического доведения до граждан и организаций предусмотренных федеральных льгот и мер поддержки (для субъектов Российской Федерации)

Минэкономразвития

Региональные планы первоочередных мероприятий (действий) по обеспечению устойчивого развития экономик

Минэкономразвития

Потребность средств индивидуальной защиты и дезинфицирующих средств в Роспотребнадзор

Роспотребнадзор

Объем импорта продуктов питания, товаров первой необходимости и лекарственных средств (по ТН ВЭД)

ФТС России

ЭТАП 3. РАЗРАБОТКА МОДЕЛИ ДАННЫХ

01.01.1900 – 31.12.9999 + все периоды

Скрытые атрибуты **ВИЗУАЛИЗАЦИЯ СВЯЗЕЙ**

ОТВЕТСТВЕННЫЙ ЗА ПРЕДОСТАВЛЕНИЕ

Организация источник: Федеральная налоговая служба (Ф...)
Ответственное лицо: [избран]

ПОКАЗАТЕЛИ КАЧЕСТВА ГОСУДАРСТВЕННЫХ ДАННЫХ

Способ контроля полноты данных: файл не выбран
Способ контроля достоверности данных: файл не выбран
Способ контроля актуальности данных: файл не выбран
Способ контроля уникальности данных: файл не выбран
Способ контроля согласованности данных: файл не выбран
Способ контроля полноты описания данных: файл не выбран

ЗАМЕЧАНИЯ К ПОКАЗАТЕЛЮ

ДИАГРАММА КЛАССОВ

Логическая модель: <http://vp.ce.a.gov.ru/public/content/CI...>
Логическая модель: <http://vp.ce.a.gov.ru/public/content/CI...>

ИСПОЛЬЗОВАНИЕ ПОКАЗАТЕЛЯ

АТРИБУТЫ БЕЗ ГРУППЫ

Периодичность мониторинга: Ежедневно
В разрезе: Суб..., РФ
Отрасли: Эко...
Правила качества для показателя:
1. Проверка на обязательность заполнения атрибута
2. Проверка на дублирование данных
3. Проверка атрибутов на соответствие справочным
Единица измерения: л
Единица изм. без справочника: [избран]

Наименование*: Объем потребления по топливу (в ...)
Описание: Объем потребления топлива в разбивке по регионам и видам (маркам) топлива
Область мониторинга: Covid-19
Куратор: [избран]
НПА: [избран]
Тип актуальности: [избран]

ИС ИСТОЧНИК: Excel
Название потока: ФНС_1_1. Объем потребления по т...
Id формы в ИЦК: [избран]

Код*: 102
Статус*: Новый

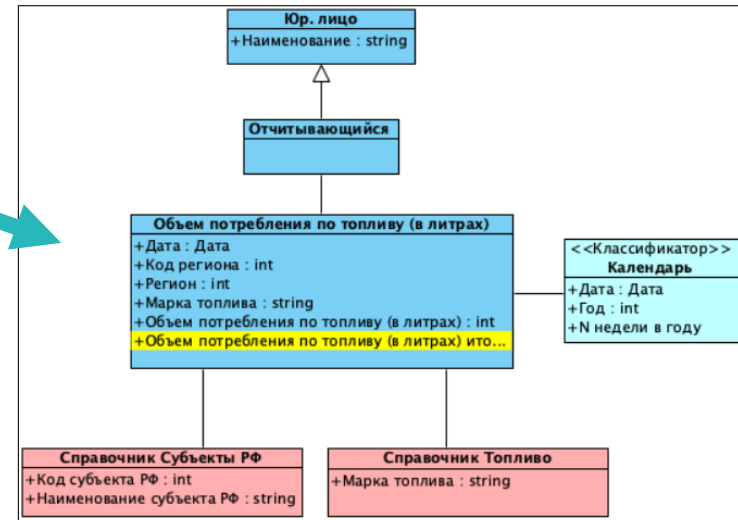
Не защищено | vp.ce.a.gov.ru/public/content/ClassDiagram_cfWymc6AU_W0oBRd.html

FNS_1_1

Мониторинг : Пакет

Диаграмма классов - Мониторинг COVID-19 [Ссылка](#)

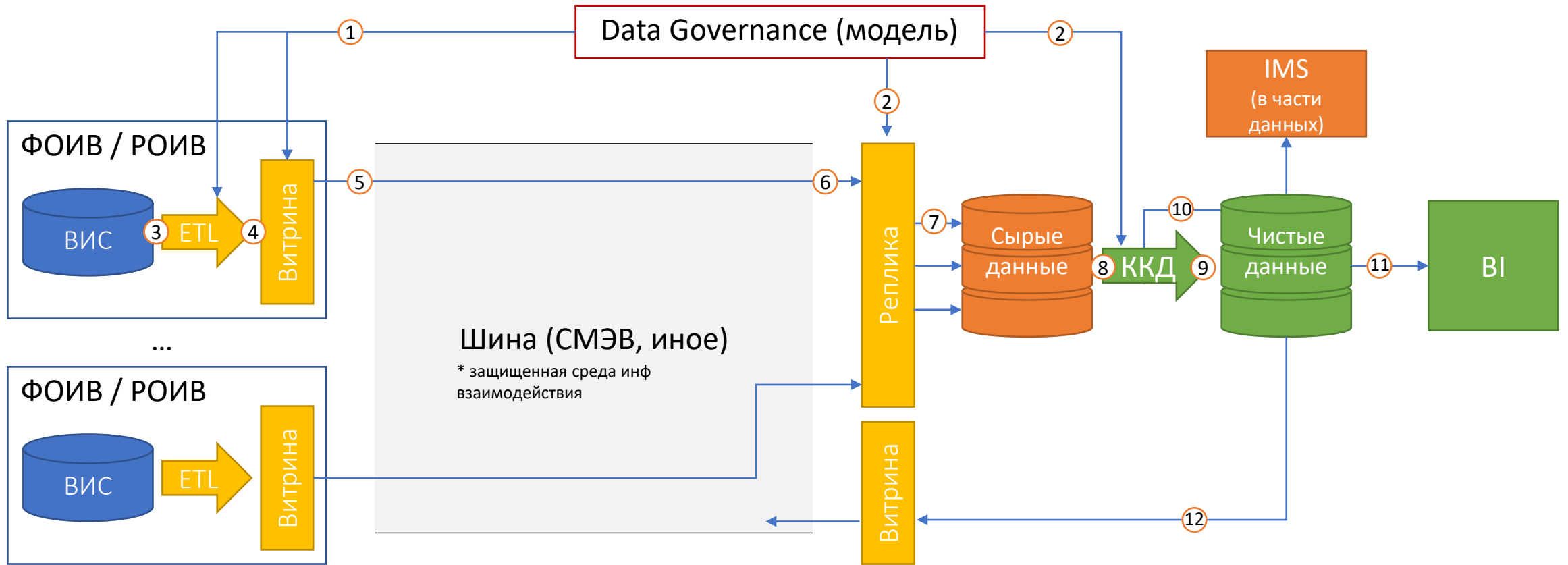
Перейти к: Please select a model element



Элементы модели

Имя	Описание
Справочник Топливо	: Класс
Справочник Субъекты РФ	: Класс
Календарь	: Класс
Объем потребления по топливу (в литрах)	: Класс
Юр. лицо	: Класс
Отчитывающийся	: Класс

ЦЕЛЕВАЯ МОДЕЛЬ ИНФОРМАЦИОННОГО ВЗАИМОДЕЙСТВИЯ



1. Метаописания форм и ФЛК передаются в ведомство для корректной отдачи данных
2. Метаописания форм передаются на реплику для корректного приема данных и в блок ККД для очистки данных
3. Ведомство настраивает ETL со своей стороны в соответствии с метаописанием
4. Данные попадают на витрину, где проходят первичный форматно-логический контроль. Это исключает передачу очевидно не корректных данных
5. Витрина передает данные и контролирует их получение репликой
6. Данные получают на реплике, реплика контролирует получение данных и их соответствие модели. Проверяется целостность переданных данных

7. Все полученные данные в неизменном виде с меткой времени и информацией об источнике передаются в хранилище сырых (грязных) данных
8. Над грязными данными осуществляется преобразование (гармонизация) и контроль качества данных
9. Данные прошедшие ККД перемещаются в хранилище чистых данных
10. По данным не прошедшим ККД формируются инциденты качества данных
11. Чистые данные передаются в BI для построения дашбордов
12. Чистые данные передаются на витрину для передачи иным заинтересованным органам

МОДЕЛЬ ГОСУДАРСТВЕННЫХ ДАННЫХ

Модель государственных данных (МГД) – описание государственных, муниципальных данных, связей между государственными, муниципальными данными в формате, пригодном для передачи, интерпретации и обработки с использованием средств автоматизации (ЗКП НСУД).

МГД создается с целью повышения эффективности создания, сбора и использования государственных данных как для предоставления государственных и муниципальных услуг и осуществления государственных и муниципальных функций, так и для обеспечения потребности физических и юридических лиц в доступе к информации (Концепция НСУД – РПРФ 1189-р от 03.06.2019 г.).

Важно!

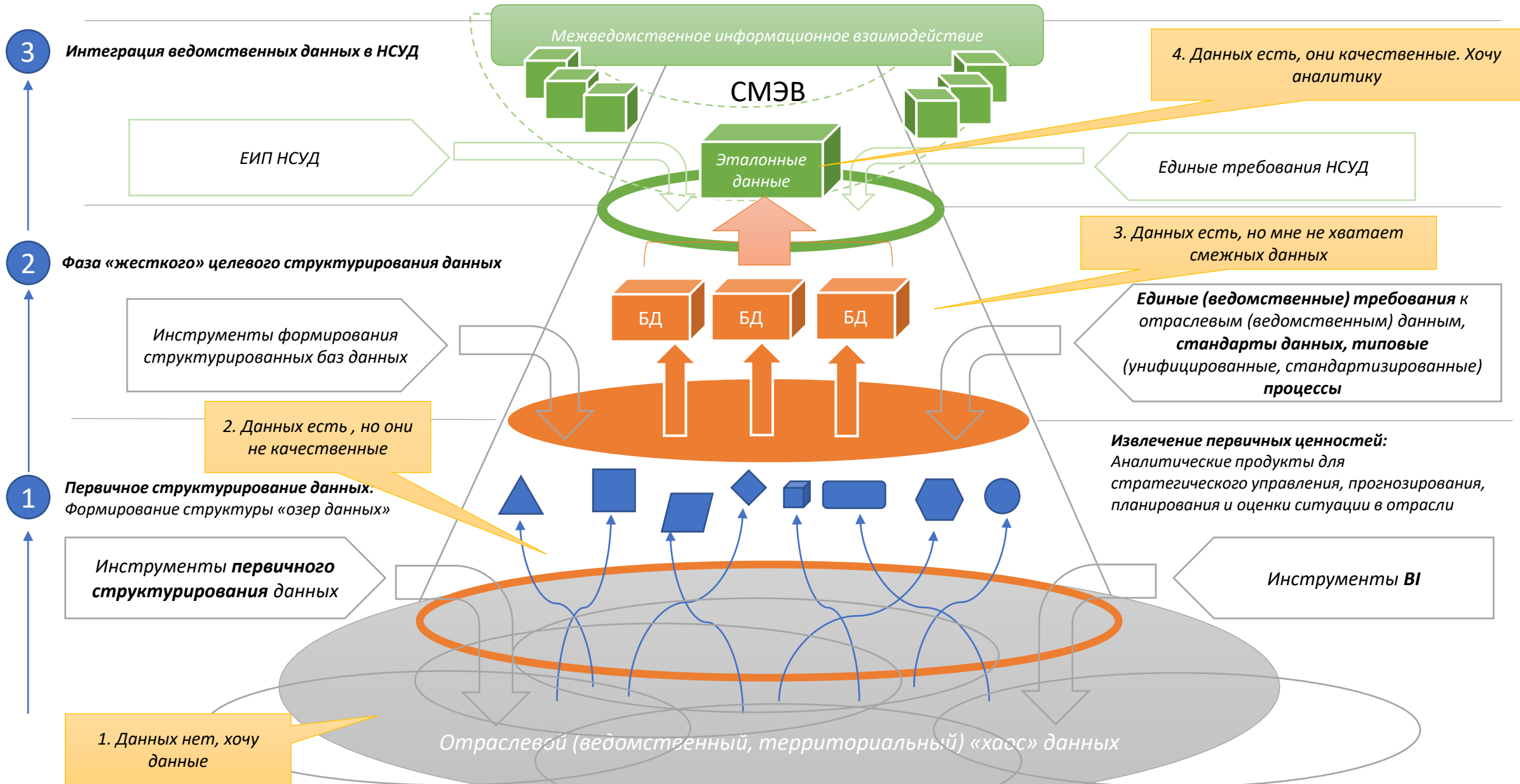
МГД

- и артефакт

- и процесс

одновременно

МОДЕЛЬ РАЗВИТИЯ СИСТЕМЫ УПРАВЛЕНИЯ ДАННЫМИ



СПАСИБО ЗА ВНИМАНИЕ!



МАЛАХОВ АЛЕКСАНДР АНДРЕЕВИЧ

Начальник управления методологического
сопровождения работы с данными
Аналитического центра при Правительстве РФ

Телефон: +7 (985) 296 65 71

E-mail: malakhov@ac.gov.ru

