



Совет руководителей  
статистических служб  
государств-участников  
Содружества  
Независимых  
Государств

# СОЗДАНИЕ ЦИФРОВОЙ ЭКОСИСТЕМЫ ДАННЫХ В РОССИЙСКОЙ ФЕДЕРАЦИИ

**Малков Павел Викторович**

руководитель Федеральной службы государственной статистики,  
Российская Федерация

Бишкек, 5 сентября 2019 года



## СУЩЕСТВУЮЩАЯ ЭКОСИСТЕМА

### ФЕДЕРАЛЬНАЯ СТАТИСТИКА

**65** федеральных органов  
исполнительной власти

Более 2000 баз данных  
федеральных органов  
исполнительной власти,  
представленных в более чем  
200 информационных системах

### РЕГИОНАЛЬНАЯ И МУНИЦИПАЛЬНАЯ СТАТИСТИКА

**85** регионов  
Более **20 000**  
муниципалитетов

Множество региональных  
и муниципальных  
информационных систем

### СПОСОБЫ СБОРА СТАТИСТИЧЕСКОЙ ОТЧЕТНОСТИ

Электронные формы  
Бумажные формы  
Спецоператоры  
Государственные  
информационные системы

- ▶ **МНОЖЕСТВЕННОСТЬ ПРОГРАММНЫХ И ТЕХНОЛОГИЧЕСКИХ РЕШЕНИЙ**
- ▶ **РАЗРОЗНЕННОСТЬ МЕТОДОЛОГИЧЕСКИХ ПОДХОДОВ**
- ▶ **НАГРУЗКА НА РЕСПОНДЕНТОВ:**

в среднем за год предоставляется форм отчетности:

**125** – предприятиями      **390** – организациями сферы социального обеспечения



## ВЫЗОВЫ ДЛЯ ОФИЦИАЛЬНОЙ СТАТИСТИКИ



- ◆ Слабая координация статистического учета на национальном уровне
- ◆ Непрозрачность процессов статистического планирования, сбора и обработки данных
  
- ◆ Высокая отчетная нагрузка на респондентов
- ◆ Низкое качество первичных данных
- ◆ Низкая оперативность и детализация
  
- ◆ Отсутствие компетенций по работе с альтернативными источниками данных
- ◆ Конкуренция со стороны производителей неофициальной статистики, использующих альтернативные источники данных



## ВЗАИМОИСКЛЮЧАЮЩИЕ ЗАДАЧИ: ВОЗМОЖЕН ЛИ БАЛАНС?

Достижение оперативности при сохранении и повышении качества

Увеличение объемов публикуемой информации при снижении нагрузки на бизнес



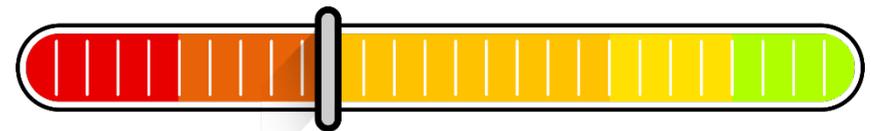
Качество

Скорость



Нагрузка на бизнес

Низкая детализация статистики

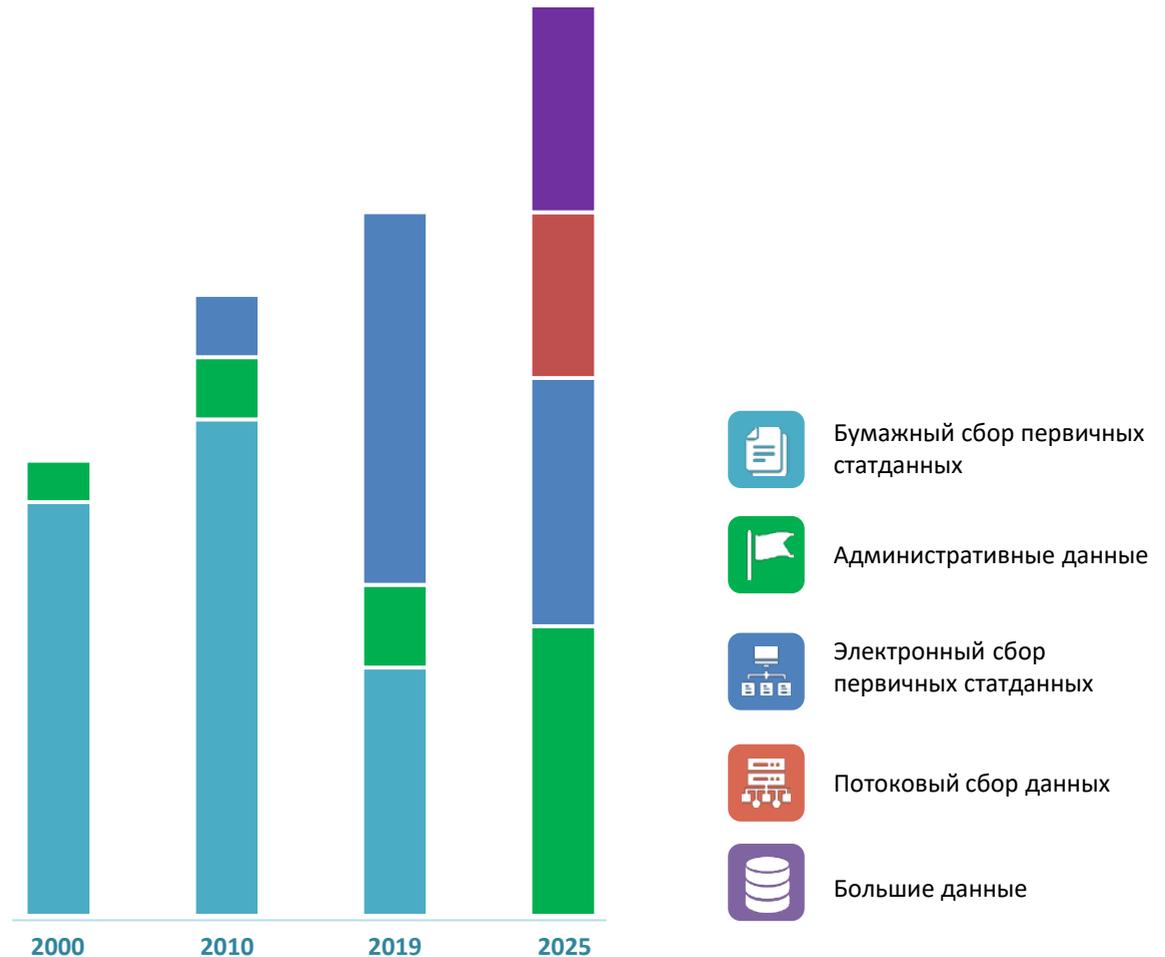


НОВЫЕ ВОЗМОЖНОСТИ ДЛЯ ДОСТИЖЕНИЯ БАЛАНСА – **ЦИФРОВАЯ ЭКОНОМИКА**



## ИЗМЕНЕНИЕ ПАРАДИГМЫ СТАТИСТИЧЕСКОГО НАБЛЮДЕНИЯ

ПЕРЕХОД ОТ  
ТРАДИЦИОННОЙ  
ОТЧЕТНОСТИ К НОВЫМ  
ИСТОЧНИКАМ ДАННЫХ





# СТРАТЕГИЯ РАЗВИТИЯ РОССТАТА ДО 2024 ГОДА

**НАША МИССИЯ: качественная и оперативная статистика  
при минимальной нагрузке на бизнес**



**ОБЕСПЕЧИТЬ**

потребителя востребованными и  
актуальными данными



**ОПТИМИЗИРОВАТЬ**

затраты на их получение



**ПОВЫСИТЬ**

внутреннюю эффективность  
статистических служб

**Статистика должна быть: независимой, открытой  
и соответствующей международным стандартам**



## ЦИФРОВАЯ АНАЛИТИЧЕСКАЯ ПЛАТФОРМА (ЦАП) – ЯДРО ИНФОРМАЦИОННО-ТЕХНОЛОГИЧЕСКОЙ ИНФРАСТРУКТУРЫ НАЦИОНАЛЬНОЙ ПРОГРАММЫ «ЦИФРОВАЯ ЭКОНОМИКА»



**ЦЕЛЬ** – сделать процесс сбора данных «незаметным» и необременительным для респондентов, а процесс распространения статистики – удобным для всех категорий пользователей



Автоматизация всех процессов статистического производства и снижение отчетной нагрузки на респондентов

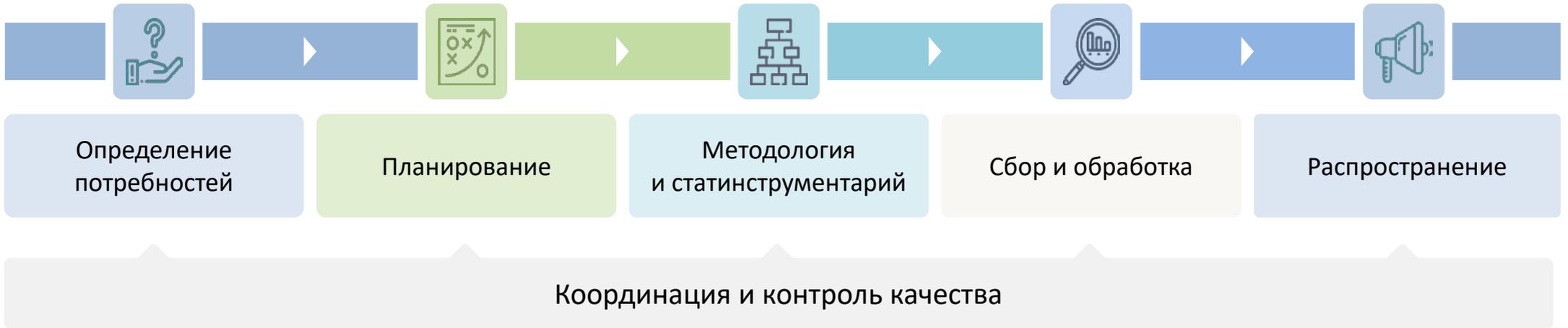


Росстат – инициатор и разработчик цифровой аналитической платформы



# МОДЕРНИЗАЦИЯ ПРОЦЕССА СТАТИСТИЧЕСКОГО ПРОИЗВОДСТВА

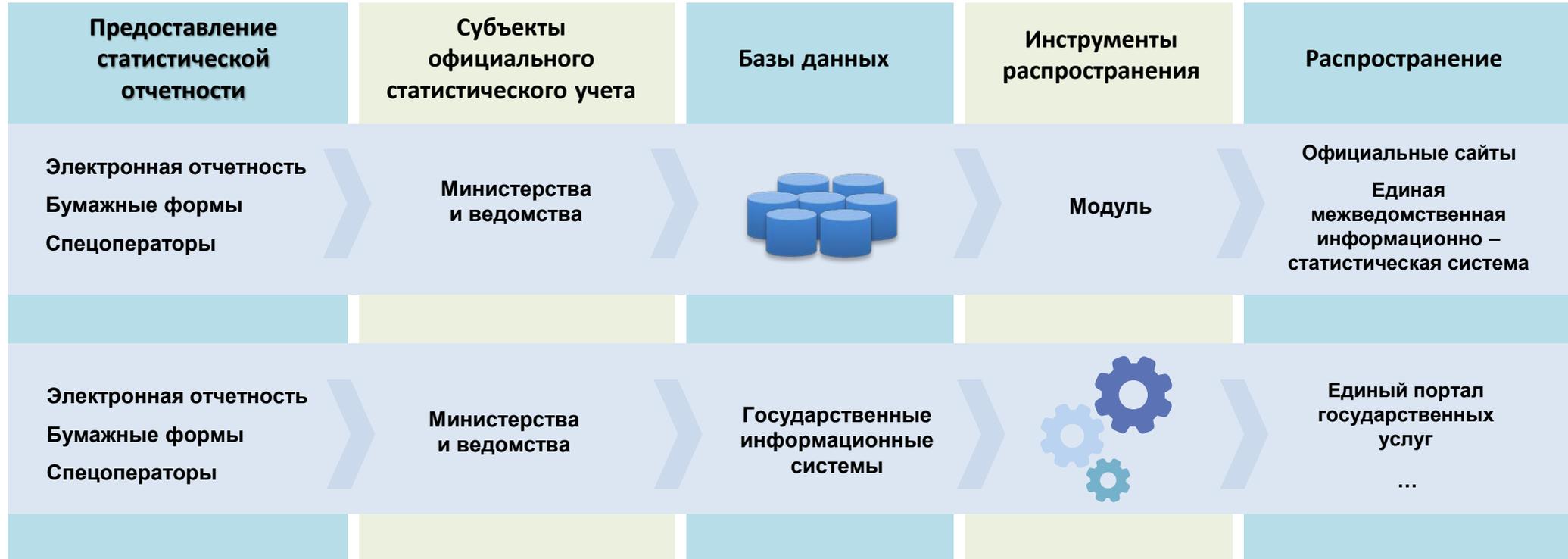
## ЦАП как единый инструмент производства статистики





# ИТ-АРХИТЕКТУРА СИСТЕМЫ ГОСУДАРСТВЕННОЙ СТАТИСТИКИ «КАК ЕСТЬ»

Респонденты



Множественность программных решений и рассогласованность подходов к каждому этапу статистического процесса



## ЦЕЛЕВАЯ МОДЕЛЬ

# ЦИФРОВАЯ АНАЛИТИЧЕСКАЯ ПЛАТФОРМА



### Источники данных

Административные данные, большие данные, респонденты



### Потребители данных

Органы власти, бизнес, эксперты, международное сообщество (SDMX, XBRL)



## КОМПОНЕНТ 1

### Формирование реестра форм статистического наблюдения и статистических показателей (РФиП)

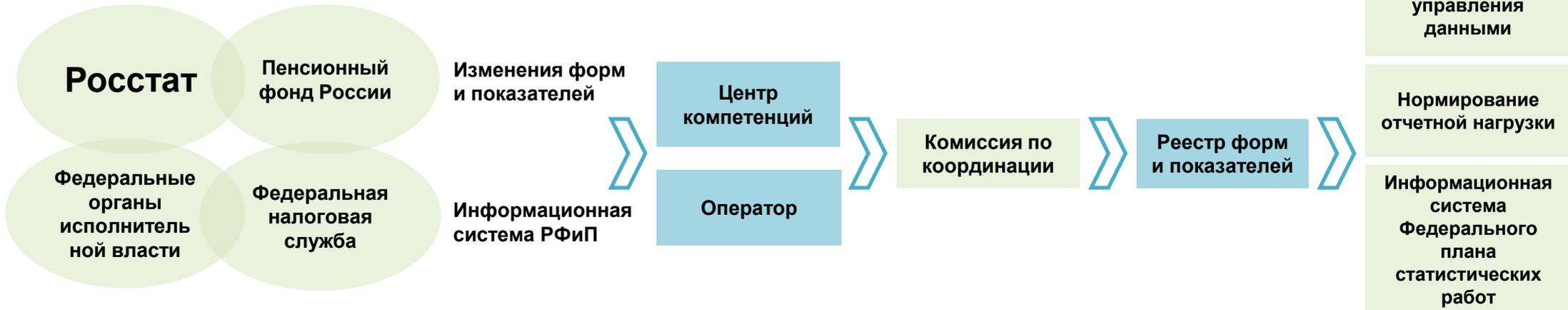
#### Вызовы:

- Избыточное количество форм отчетности
- Дублирование первичных данных до 60%
- Неэффективное применение первичных данных
- Нерегулируемое создание ведомственной, региональной отчетности

#### Решения:

- Запрет сбора отчетности по формам и показателям, которые не включены в реестр
- Обоснованное создание новых форм и показателей
- Определение и устранение дублирования форм и показателей
- Минимальный объем отчетной нагрузки, ее систематизация и нормирование
- Единая методология описания форм и показателей
- Однократный запрос первичных данных
- Увязка собираемой и распространяемой статистической информации

#### Результат (схема):





## КОМПОНЕНТ 2

### Формирование реестра объектов статистического наблюдения

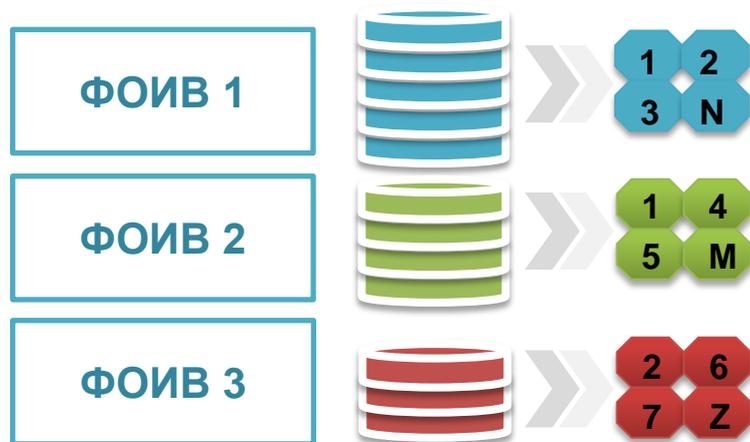
#### Вызовы:

- Множественность реестров респондентов
- Количество респондентов и их свойств не совпадают
- Принципиально разные подходы к формированию совокупности объектов наблюдения

#### Решения:

- Единый реестр респондентов с объединенными свойствами
- Автоматизированное формирование выборок респондентов по заданным критериям
- Сквозная привязка первичных данных к уникальному респонденту
- Аналитика по свойствам респондентов

#### Результат (схема):



#### Информационная система Реестр Респондентов

(на базе статистического регистра, Единого государственного реестра индивидуальных предпринимателей, Единого государственного реестра юридических лиц )





## КОМПОНЕНТ 3

### Формирование Федерального плана статистических работ (ФПСР)

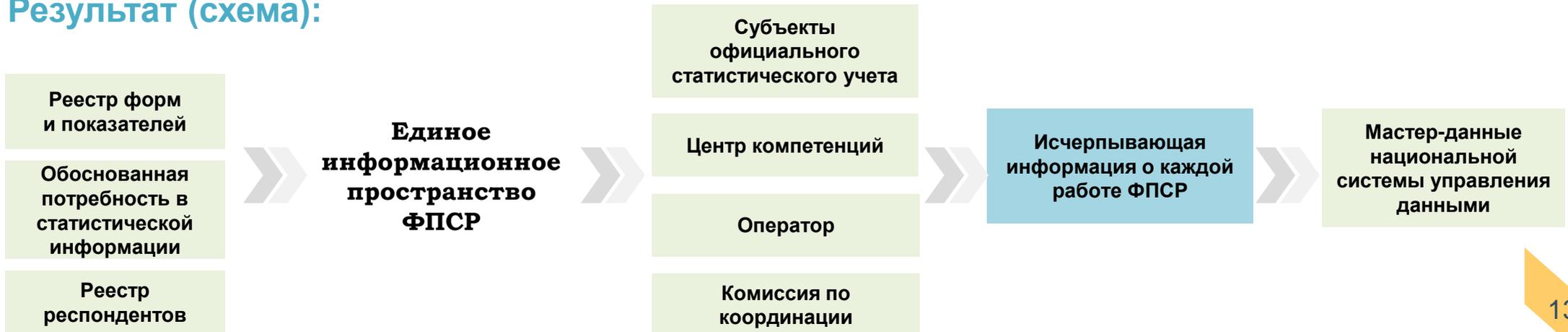
#### Вызовы:

- Нет взаимосвязи отчетности и работ ФПСР (отчетность  $\neq$  выходная информация)
- Не оптимальное расходование бюджета на стат. деятельность
- Низкая детализация распространяемой стат. информации
- Непрозрачность используемых ведомствами методологии и метаданных
- Нет инструментов оценки востребованности стат. информации

#### Решения:

- Единое информационное пространство планирования статистических работ
- ФПСР информативный для пользователя
- Оценка востребованности статистической информации
- Оценка нагрузки на бизнес на этапе планирования работ
- Осмечивание работ ФПСР
- Унифицированное методологическое описание стат. работ, включая метаданные

#### Результат (схема):





## КОМПОНЕНТ 4

### Единое хранилище первичных статистических данных

#### Вызовы:

- Нет единой базы первичных стат. данных
- Однократное использование первичных статистических данных
- Многократный ввод одних и тех же первичных стат. данных
- Каждый субъект учета собирает первичные стат. данные от респондента самостоятельно, по своим правилам

#### Решения:

- Многократное и совместное использование первичных стат. данных
- Устранение дублирование первичных стат. данных
- Возможность анализа первичных стат. данных всеми категориями пользователей (с учетом деперсонификации)
- Сбор данных через единое окно
- Сохранение всех первичных стат. данных

#### Результат (схема):





## КОМПОНЕНТ 5

### Инструмент для бизнеса по предоставлению статистической отчетности

#### Вызовы:

- Респондент не знает какую отчетность кому предоставлять
- Множество инструментов для представления отчетности
- Процесс автоматизирован частично
- Необходимо отслеживать частое изменение форм

#### Решения:

- Однократность сбора первичных статистических данных
- Адресный подход к сбору (Личный кабинет)
- Автоматизация процесса предоставления стат. отчетности
- Актуальность метаданных
- Унифицированный инструмент

#### Результат (схема):





## КОМПОНЕНТ 6

### Аналитика на базе первичных статистических данных (BI компонент)

#### Вызовы:

- Первичные стат. данные не используются повторно
- Несопоставимость объемов первичных и агрегированных стат. данных
- Аналитическая стат. информация не формируется
- Неудовлетворенность пользователей в детализированной информации

#### Решения:

- Многократность использования первичных статистических данных
- Формирование показателей на базе первичной информации по запросу (витрины данных)
- Детализированная статистическая информация для ведения бизнеса
- Доступ к деперсонифицированным данным
- Аналитические инструменты работы с данными, API

#### Результат (схема):





## РАЗРАБОТКА КОМПОНЕНТОВ ЦАП



- **Автоматизированная система статистического планирования**

- **Единый реестр респондентов**

- ◆ единая генеральная совокупность для всех органов власти
- ◆ единые правила и единый инструментарий формирования выборок

- **Единый реестр форм и показателей отчетности**

- ◆ автоматизированное выявление дублирования форм и показателей
- ◆ право респондента не предоставлять отчетность вне реестра

- **Единое хранилище первичных данных**

- ◆ «Единое окно» приема отчетности
- ◆ интеграция с государственной информационной системой федеральных органов исполнительной власти

- **Электронный сбор через «Единое окно»**

- ◆ обязательность предоставления отчетности в электронном виде
- ◆ единый порядок электронного сбора для всех федеральных органов исполнительной власти

- **Аналитические инструменты ЦАП (BI компонент)**

- ◆ формирование витрины данных по запросам пользователей
- ◆ API-доступ к Единому хранилищу данных

- **Потоковый сбор**

- **Единая централизованная технология сбора и обработки**

2020

2021

2022

2023

2024+





## SMART STATISTICAL SYSTEM



### КАЧЕСТВЕННАЯ И ОПЕРАТИВНАЯ СТАТИСТИКА

- ◆ Данные в режиме реального времени
- ◆ Предиктивная аналитика
- ◆ Прозрачность методологии
- ◆ Понятная статистика для всех
- ◆ Быстрое реагирование на новые вызовы
- ◆ Микроданные

### МИНИМАЛЬНАЯ НАГРУЗКА НА РЕСПОНДЕНТОВ

- ◆ Единая система статистического производства
- ◆ Максимальное использование административных данных
- ◆ Активное использование больших данных
- ◆ Поточковый сбор (M2M)