



Казначейство России

Методологические подходы к созданию, ведению, изменению и применению классификаторов, реестров и иных информационных ресурсов и созданию единой информационной среды в сфере систематизации и кодирования информации

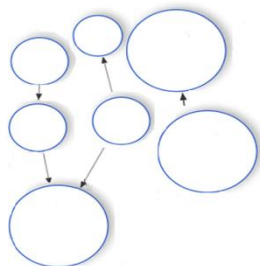
Заместитель руководителя
А.Т. Катамадзе

г. Москва

06.02.2015 г.

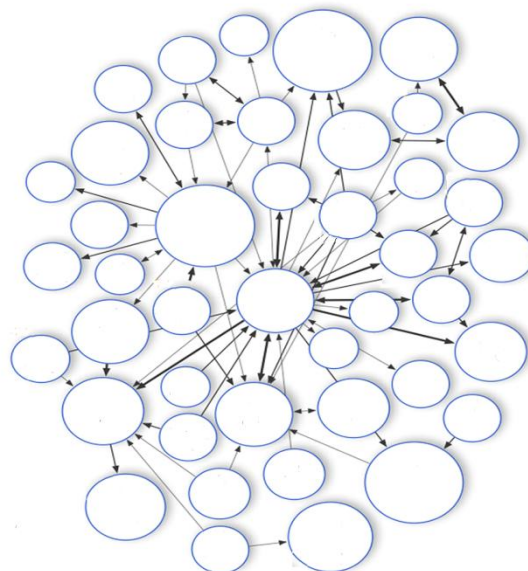


До электронного
правительства



**Потребность
в локальном
взаимодействии**

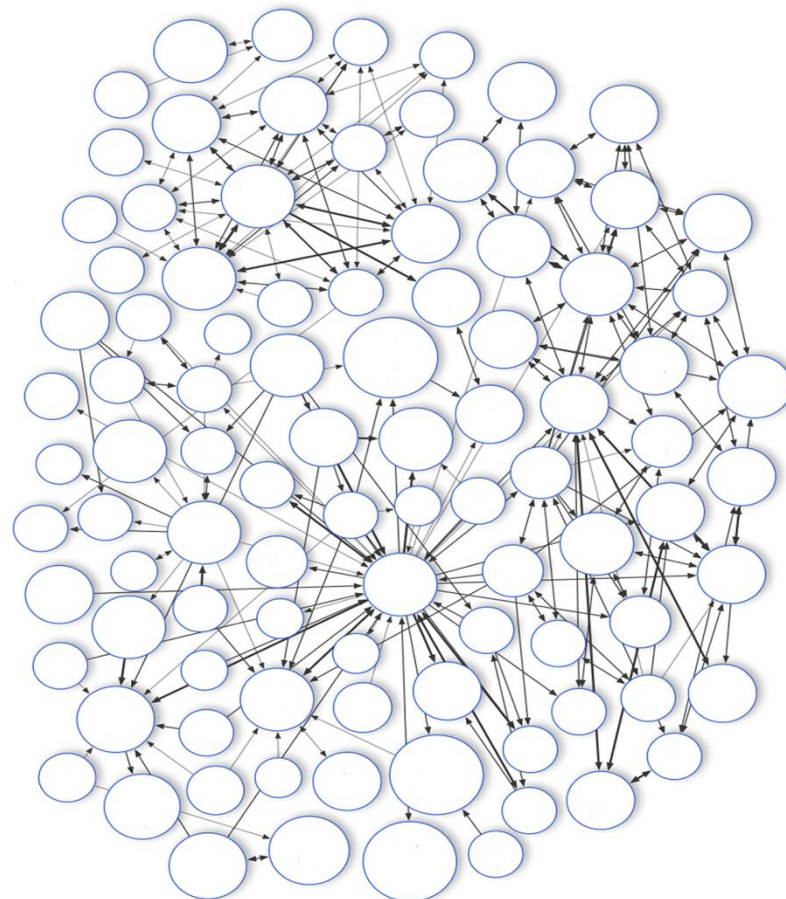
Текущая ситуация



**Потребность в обмене данными
каждого со многими:**

- > 300 ИР федерального уровня
- > 17 000 видов данных
- > 2 млрд. реестровых записей

Будущее



**Нарастание потребности в обмене
данными всех со всеми, в т.ч. с базами
данных 85 субъектов РФ, > 24 тыс. МО**



При выполнении гос. функций формируется множество реестров, регистров, кадастров и пр. Это привело к образованию баз данных с разрозненными сведениями об одних и тех же объектах и субъектах правоотношений.

В итоге:

- Данные не достоверны - ошибки при многократном вводе
- Данные не актуальны - где-то обновлены, где-то нет в зависимости от заявителя
- Первичные документы проверяет каждый ФОИВ – избыточно, риски ошибок
- Данные без идентификаторов - усложняет обмен
- Интеграция друг с другом всех ресурсов – дорого, не масштабируемо

Задача: обеспечить обмен актуальными данными между всеми инф. ресурсами на постоянной основе.

Нужны новые подходы!



План мероприятий по формированию методологии систематизации и кодирования информации... утвержденный поручением А.В. Дворковича от 31 июля 2014 г. № АД-П10-5785 (далее – План мероприятий)



Концепция методологии утверждена Правительством РФ



Концепция методологии систематизации и кодирования информации, а также совершенствования и актуализации общероссийских классификаторов, реестров и справочников (далее – Концепция методологии), утвержденная распоряжением Правительства Российской Федерации от 10 мая 2014 г. № 793-р.

✓ **Разработан проект Концепции создания единой информационной среды**



Концепция создания единой информационной среды (далее - Концепция) в сфере систематизации и кодирования информации (пункт 1.2 Плана мероприятий).

✓ **Разработан проект постановления Правительства Российской Федерации**



Постановление Правительства Российской Федерации «О порядке создания, ведения, изменения и применения отдельных информационных ресурсов» (далее – Постановление) (пункт 1.1 Плана мероприятий).

✓ **Разработан План мероприятий к проекту Постановления**

| № п/п | Наименование мероприятия | Исполнители | Срок | Ответственный |
|-------|--------------------------|---|--------------------|---------------|
| 1 | Выполнение поручения | Министерство Российской Федерации по развитию информационных ресурсов | до 31 июля 2014 г. | Министр |
| 2 | Выполнение поручения | Министерство Российской Федерации по развитию информационных ресурсов | до 31 июля 2014 г. | Министр |
| 3 | Выполнение поручения | Министерство Российской Федерации по развитию информационных ресурсов | до 31 июля 2014 г. | Министр |
| 4 | Выполнение поручения | Министерство Российской Федерации по развитию информационных ресурсов | до 31 июля 2014 г. | Министр |

План мероприятий по реализации постановления Правительства Российской Федерации «О порядке создания, ведения, изменения и применения отдельных информационных ресурсов».

Конкретные правовые и технологические решения в постановлениях!



Постановление № 928 «О БГИР» - единственный документ.
Однако в нормативной правовой базе:

- Отсутствуют критерии определения источника эталонных данных
- Отсутствуют требования к созданию и ведению ресурсов для обеспечения обмена данными
- Реестр паспортов БГИР не позволил унифицировать источники данных и их форматы
- Автоматизация процесса обмена данными между ресурсами предусматривается «каждый» с «каждым», что существенно усложняет обмен

**Не обеспечен обмен актуальными данными
одновременно между всеми ресурсами на постоянной основе**

Нужны новые правовые и технологические решения!



В проекте постановления о порядке создания, ведения, изменения и применения отдельных информационных ресурсов установлены единые требования:

- Виды ресурсов
- Правило определения источников данных (где эталон)
- Правила создания и ведения ресурсов
- Новая технология обмена данными между ресурсами

Далее - о каждом!



Информационный ресурс – «информация, содержащаяся в ГИС, а также иные имеющиеся в распоряжении гос. органов сведения и документы. Информация, содержащаяся в ГИС, является официальной» (п. 9 статьи 14 Закона № 149-ФЗ «Об информации...»).

Определены два вида ресурсов:

- Базовые и производные ресурсы – включают данные о конкретных объектах учета (ООО «Ромашка», «Петров П.С.»). Участвуют в обмене данными. Базовые ресурсы – содержат эталонные данные. Производные ресурсы – не содержат эталонные данные.
- Классификаторы - группируют объекты по общему признаку/типу («товары», «ОПФ» и пр.) Применяются для структурирования данных в базовых и производных ресурсах. Общероссийские классификаторы – являются обязательными для использования в базовых и производных ресурсах. Ведомственные справочники – для локальных задач.



Предложены универсальные критерии определения эталонных данных. Эталонные данные должны отвечать следующим требованиям:

- Формирование осуществляется ФОИВ в соответствии с его полномочиями впервые
- Изменение правового статуса объекта учета влечет необходимость соответствующих изменений данных
- При формировании проверяются все необходимые первичные документы

Дублирующиеся данные вносятся только с использованием эталонных данных.
Требовать от заявителя первичные документы в целях их формирования запрещено – новое правило



Новые требования к ресурсам:

- Должен быть сформирован в виде ГИС (база данных)
- Должен применять идентификатор своих эталонных данных
- Должен включать идентификаторы эталонных данных других ресурсов
- Должен применять классификаторы
- Включение дублирующихся данных с использованием эталонных должно осуществляться на постоянной основе и автоматизировано

Новая технология - Реестр видов данных, Система



I

Разработать единые форматы данных и внедрить их во все ресурсы.
Обмен самими данными между ресурсами сохранить по принципу «каждый с каждым», не связывая данные ресурсов.

II

Связать данные по ID и осуществлять обмен через единую систему.
Затем перейти к унификации форматов там, где есть обмен данными.

Реестр видов данных – необходим для I, II подходов



1

Отв. ОИВ
ЗАГС ПФР ФМС ФНС

1. Формируют перечень данных, которые содержатся в ИР для включения в заявку,
2. Подают через Систему заявку в РВД.

2

Оператор РВД

1. Определяет совместно с ОИВ виды данных (эталонные, дублирующиеся, источники, форматы и пр.),
2. Распределяет данные в РВД. Включает описание форматов эталонных данных, в т.ч. справочников,
3. Обеспечивает доступ ОИВ к РВД в целях формирования НПА и создания (ведения ИР).

| № зап. | Эталонные данные | | | Дублирующиеся данные | | | |
|--------|------------------------|----------------------|-----------------------|----------------------|------------------------|--------------------------------|---------------------------------|
| | ИР-эталон | ID ИР-эталона | Наим-ие этал. данного | База СНИЛС (ПФР) | Реестр паспортов (ФМС) | ЕГРЮЛ (ФНС) | Реестр лицензий в области связи |
| 1 | ЗАГС | № свид-ва о рождении | ФИО рожд. | ФИО застрах. | ФИО гражданина | ФИО ген. директора | ФИО ген. дир. лиценз. |
| 2 | | | пол рожд. | пол застрах. | пол гражданина | | |
| 3 | | | дата рождения | дата рожд. застрах. | дата рожд. гражданина | дата рожд. ген. директора | |
| 4 | | | место рождения | место рожд. застрах. | место рожд. гражданина | | |
| 5 | Реестр паспортов (ФМС) | № паспорта | дата выдачи паспорта | | | дата выдачи паспорта ген. дир. | |
| 6 | | | выдавший орган | | | выдавший орган | |
| 7 | | | код подразделения | | | код подразделения | |
| 8 | ЕГРЮЛ (ФНС) | ИНН | наим. ю. л. | | | | наим. ю. л.-лиценз. |
| 9 | ГАР (ФНС) | № ГАР | адрес | | адрес ф.л. | адрес ю.л. | адрес ю.л.-лиценз. |

Наборы видов данных

3

Наборы видов данных используются:

- при утверждении НПА об ИР,
- для ведения самого ИР.

Вывод:

РВД – вспомогательный инструмент для корректного формирования мета-данных. Вместе с тем, для обмена данными о конкретных объектах нужны дополнительные технологии

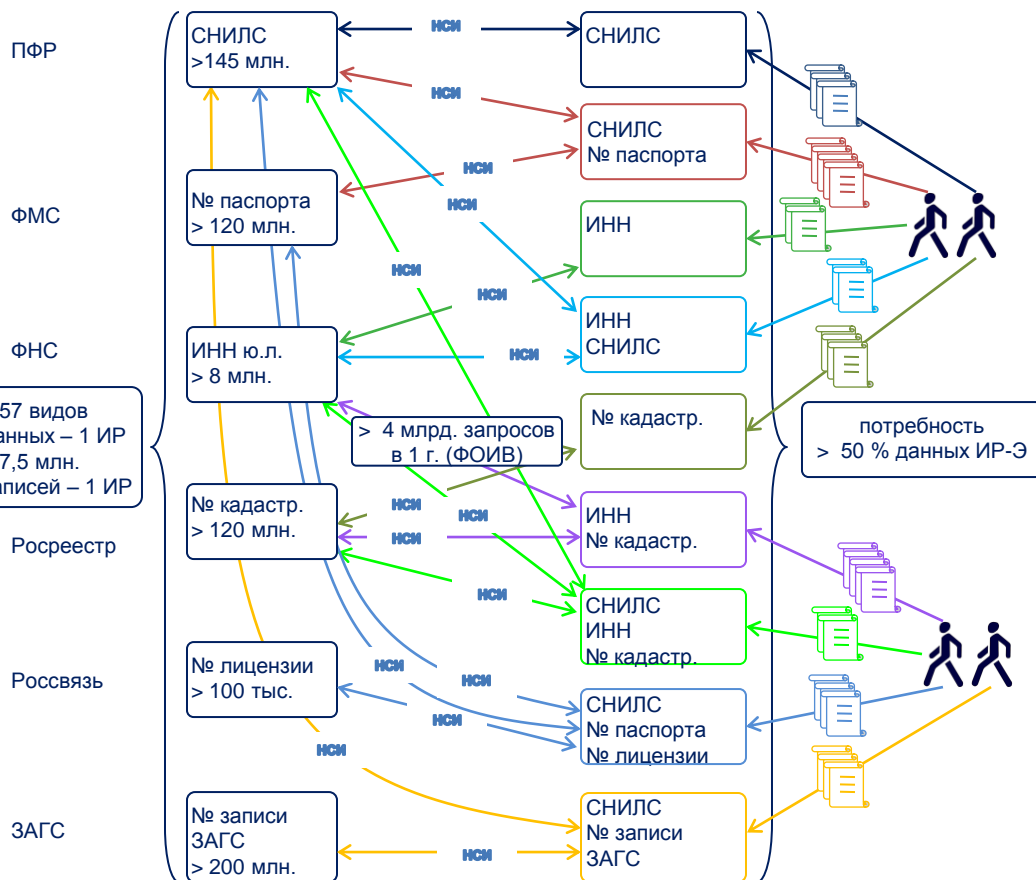
Обмен данными: текущая ситуация – запрос-ответ



ИР- эталоны

> 300 ИР

ИР- потребители



Запрос-ответ не позволяет организовать обмен актуальными данными, так как:

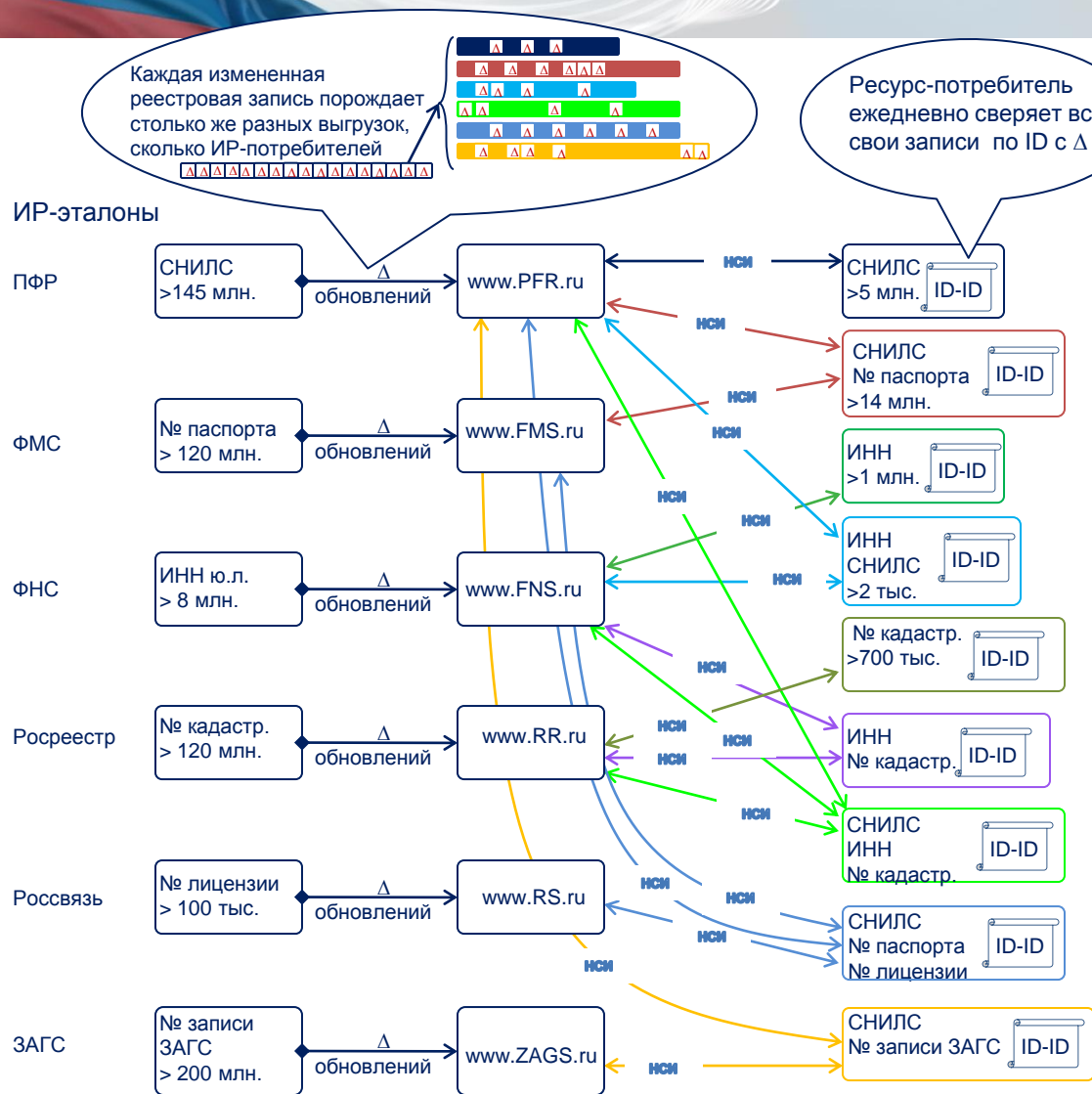
- в ресурсах нет эталонных ID, поэтому актуализация эталонных данных зависит от заявителя
- обмен данными каждого со многими не позволяет унифицировать форматы
- в условиях разной потребности в данных взаимодействие ресурсов «каждый с каждым» не позволяет управлять подпиской на обновления

Вывод:

- данные не достоверны без обращения заявителя – начисление пенсий и льгот умершим
- каждый проверяет первичные документы – избыточная функция
- ресурсы-эталон перегружены большим числом повторяющихся запросов
- нет возможности унифицировать справочники
- интеграция всех ресурсов со всеми по отдельности – дорого, не масштабируемо

Нужны новые подходы к организации обмена данными!

Обмен данными: вариант I – запрос-ответ к Δ обновлений



Запрос-ответ к Δ обновлений предполагает, что:

- ресурсы-эталоны должны обеспечить хранилище для ежедневных выгрузок Δ обновлений
- ресурсы-эталоны должны генерировать множество отчетов-выгрузок по каждой записи в зависимости от потребителя
- ресурсы-потребители должны ежедневно сверять все записи с выгрузками Δ обновлений, так как нет единого хранилища связей ID
- де-факто ресурсы будут иметь разные форматы по схожим сущностям, так как нет единой системы обмена данными
- необходимо обеспечить включение эталонных ID в ресурсы

Вывод:

- реализация возможна, но требует значительных доработок ресурсов-участников обмена и высокой организации владельцев ресурсов при обмене актуальными данными
- нет возможности администрировать всех в части реализации подхода

Целесообразен переход к единой технологии для соблюдения общих правил!

Обмен данными: вариант II - единая система

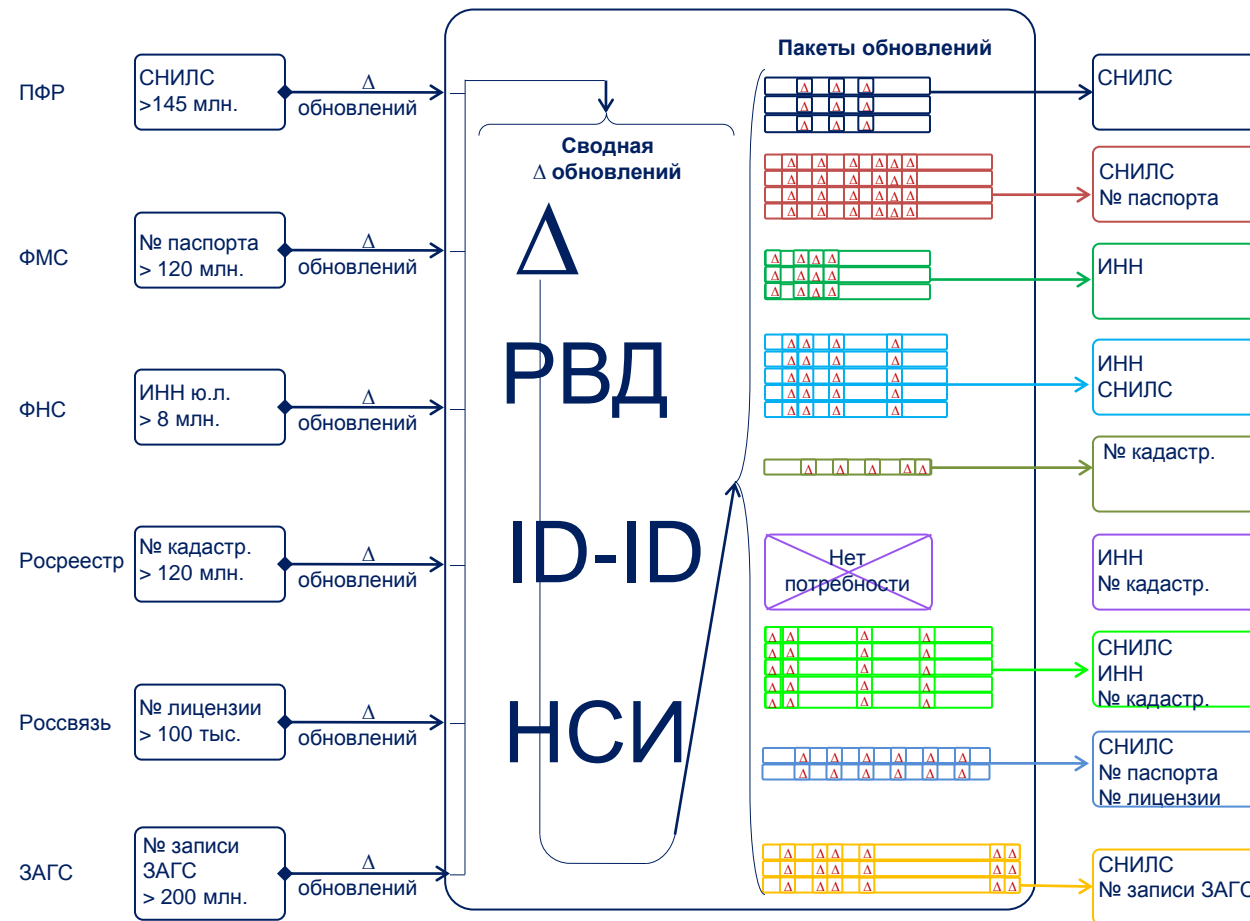


ИР-эталоны

Единая система

ИР-потребители

Обмен данными через единую систему предполагает:



- хранение ежедневных Δ обновлений в одной системе, а не каждым ресурсом
- генерацию отчетов-выгрузок по записям в зависимости от потребности через единую модель данных, а не каждым
- хранение связей ID, что позволяет очищать отчеты от неактуальных данных – не нужно всем ежедневно сверять все записи
- переход на унифицированные форматы, там где имеет место обмен данными
- упрощение процесса включения эталонных ID в ресурсы. Возможно обогащение данных

Вывод:

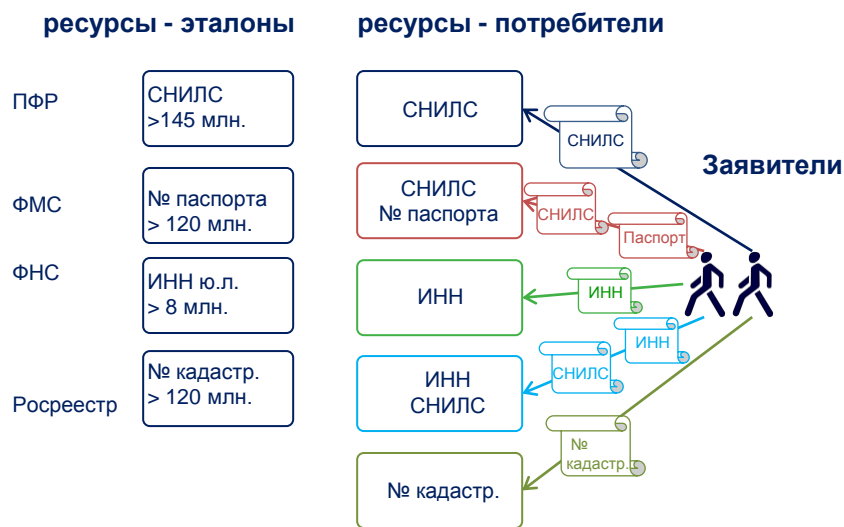
- единая система разгрузит ОИВ от непрофильной деятельности (сервисы выгрузок, Δ обновлений, отчеты потребителям и пр.) и позволит сосредоточить работу над улучшением качества эталонных данных
- уполномоченному органу необходимо создать систему и обеспечить взаимодействие с ней участников обмена

Подход направлен на оптимизацию сбора, обработки информации и обеспечение актуальности данных в ресурсах!

Подходы к включению эталонных идентификаторов в информационные ресурсы



Базовый для I и II вариантов: по мере обновления данных включать эталонные ID в ресурсы



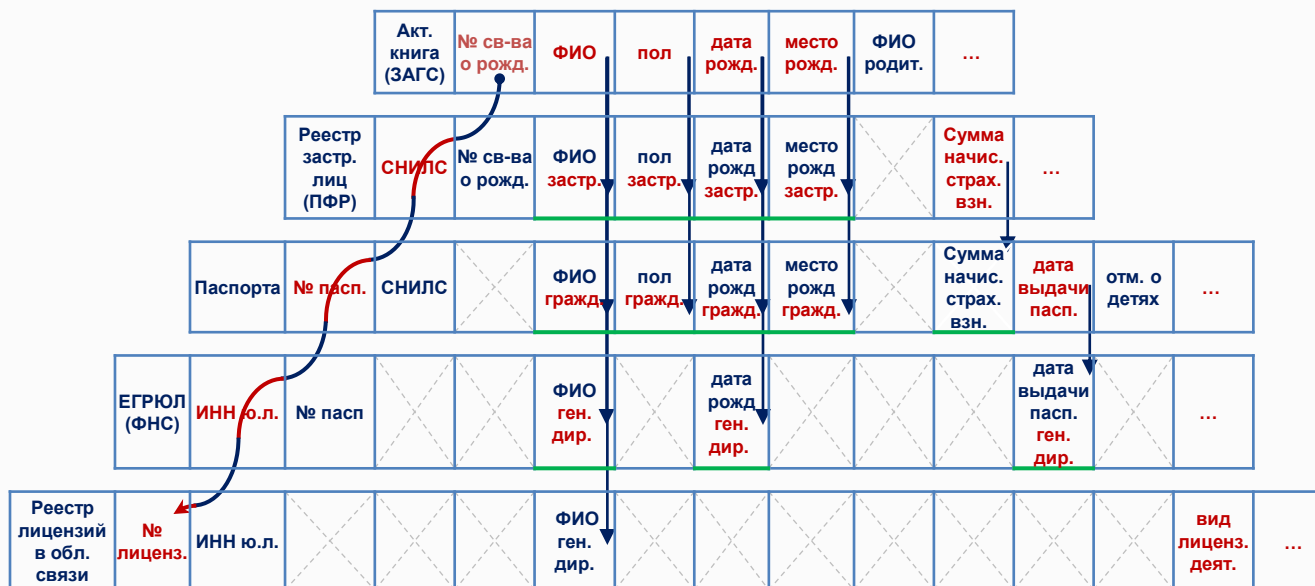
Дополнительный для II варианта: «обогащение данных» за счет доступа ко всем ресурсам



Для чего необходимо хранить связи идентификаторов в единой системе



Идентификатор каждого последующего ресурса определяет новый правовой статус



Вывод:

- обмен данными основан на иерархии связей ID, а не одной связи ID
- без хранения в системе иерархии связей ID, каждый ресурс должен запрашивать Δ обновлений по всей цепочке, что существенно усложняет обмен
- хранение иерархии связей ID в системе также позволяет одновременно обновлять данные всех потребителей непосредственно из ресурсов-эталонов

Актуализация данных через единую систему



- 1 Обновление данных в ресурсе-этalone
Выгрузка Δ обновлений
- 2 Сравнение Δ обновлений с хэш-копиями ресурсов
Очистка данных
- 3 Определение соответствующих реестровых записей и видов данных
Формирование пакетов обновлений
- 4 Выгрузка обновлений потребителю

Ресурс-этalone ЕГРЮЛ

| ID | A | B | C |
|-----------------|----------------|----------------|----------------|
| ID ₁ | A ₁ | B ₁ | C ₁ |
| ID ₂ | A ₂ | B ₂ | C ₂ |
| ID ₃ | A ₃ | B ₃ | C ₃ |

1
 Δ обновлений

| | | | |
|-----------------|----------------|----------------|----------------|
| ID ₁ | A ₁ | B ₁ | C ₁ |
| ID ₂ | A ₂ | B ₂ | C ₂ |

2

Хэш-копии ресурсов

| ID | A | B | C |
|-----------------|-----------------|-----------------|-----------------|
| ID ₁ | #A ₁ | #B ₁ | #C ₁ |
| ID ₂ | #A ₂ | #B ₂ | #C ₂ |
| ID ₃ | #A ₃ | #B ₃ | #C ₃ |

СИСТЕМА

НСИ

4

| | | |
|-----------------|-----------------|----------------|
| ID ₇ | ID ₁ | A ₁ |
| ID ₇ | ID ₄ | Z ₄ |

РФИ (ресурс-потребитель)

| ID | ID | ID | A | C | Z |
|-----------------|-----------------|-----------------|----------------|----------------|----------------|
| ID ₇ | ID ₁ | ID ₄ | A ₁ | C | Z ₄ |
| ID ₈ | ID ₂ | ID ₆ | A ₂ | C ₂ | Z ₆ |

4

| | | | |
|-----------------|-----------------|----------------|----------------|
| ID ₈ | ID ₂ | A ₂ | C ₂ |
| ID ₈ | ID ₆ | Z ₆ | |

4

Иной ресурс-потребитель

| ID | ID | ID | B | Y |
|-----------------|-----------------|-----------------|----------------|---|
| ID ₉ | ID ₂ | ID ₅ | B ₂ | Y |

Ресурс-этalone ЕГРП

| ID | X | Y | Z |
|-----------------|----------------|----------------|----------------|
| ID ₄ | X ₄ | Y ₄ | Z ₄ |
| ID ₅ | X ₅ | Y ₅ | Z ₅ |
| ID ₆ | X ₆ | Y ₆ | Z ₆ |

1
 Δ обновлений

| | | | |
|-----------------|----------------|----------------|----------------|
| ID ₄ | X ₄ | Y ₄ | Z ₄ |
| ID ₆ | X ₆ | Y ₆ | Z ₆ |

2

Хэш-копии ресурсов

| ID ² | X | Y | Z |
|-----------------|-----------------|-----------------|-----------------|
| ID ₄ | #X ₄ | #Y ₄ | #Z ₄ |
| ID ₅ | #X ₅ | #Y ₅ | #Z ₅ |
| ID ₆ | #X ₆ | #Y ₆ | #Z ₆ |

Связи ID

| | | |
|-----------------|-----------------|-----------------|
| ID ₁ | ID ₇ | - |
| ID ₂ | ID ₈ | ID ₉ |
| ID ₃ | - | - |
| ID ₄ | ID ₇ | - |
| ID ₅ | - | ID ₉ |
| ID ₆ | ID ₈ | - |

РВД

| | | |
|---|---|---|
| A | A | - |
| B | - | B |
| C | C | - |
| X | - | - |
| Y | - | Y |
| Z | Z | - |

3

| | |
|-----------------|----------------|
| ID ₄ | Z ₄ |
| ID ₆ | Z ₆ |



I вариант – от единых форматов к самостоятельному обмену данными

Разработать единые форматы данных и внедрить их во все ресурсы.

Обмен самими данными между ресурсами сохранить по принципу «каждый с каждым», не связывая данные ресурсов.

Результат: общие неизменные форматы и структура данных

II вариант – от организованного обмена данными к унификации востребованных форматов

Связать данные по ID и осуществлять обмен через единую систему. Затем перейти к унификации форматов там, где есть обмен данными.

Результат: единая система обмена актуальными данными



Спасибо за внимание!