



SUSTAINABLE DEVELOPMENT GOALS

Компиляция глобальных данных ЦУР с учетом стандарта SDMX Анализ

17 апреля 2018 г.

Эдуардо Белинчон

Статистический отдел Организации Объединенных Наций

Что такое стандарт SDMX

- «SDMX, что расшифровывается как «Обмен статистическими данными и метаданными» - это международная инициатива, направленная на стандартизацию и модернизацию («индустриализация») механизмов и процессов обмена статистическими данными и метаданными между международными организациями и их странами-участницами».

Источник: Википедия



Спонсоры SDMX

- [Банк международных расчетов](#) (БМР)
- [Европейский центральный банк](#) (ЕЦБ)
- [Евростат](#) (Статистическая служба Европейского союза)
- [Международный валютный фонд](#) (МВФ)
- [Организация экономического сотрудничества и развития](#) (ОЭСР),
- [Статистический отдел Организации Объединенных Наций](#) (СО ООН)
- [Всемирный банк](#)



Что такое стандарт SDMX

- Стандарт моделирования данных:
 - Как мы структурируем наши данные
- Технический стандарт
 - Как мы обмениваемся нашими данными
- Концепции
 - Величины, атрибуты..
- Перечни кодов
 - Ряды, значения величин, значения атрибутов...
- Потоки данных
- Валидация, трансформация и т.д.



Моделирование данных:

1.2.1 - Доля населения страны, живущего за официальной чертой бедности, в разбивке по полу и возрасту

США: Процент населения США, живущего за чертой бедности в США

<https://sdg.data.gov/>

год	tot_proportion_belownatpov_line	мужчины	женщины	младше 18 лет	18 – 64 лет	от 65 лет и старше	Примечание
2000	11,3	9,9	12,6	16,2	9,6	9,9	Реализация 2000...
2001	11,7	10,4	12,9	16,3	10,1	10,1	
2002	12,1	10,9	13,3	16,7	10,6	10,4	

1.2.1. - А) Доля населения, проживающего за национальной чертой бедности, в разбивке по регионам

Бангладеш

<http://www.sdg.gov.bd/>

Временной период	Национальный уровень/ Разбивка	Значение/ Описание	Замечания
2000	Всего	48,9	
2010	Всего	31,5	
2016 (базовый год)	Всего	24,3	
	Городской уровень	18,9	
	Сельский уровень	26,4	

- Представление данных
- Наименование
- Коды величин и рядов



Моделирование данных в SDMX:

REF_AREA	SERIES	URBANISATION	TIME_PERIOD	OBS_VALUE	TIME_DETAIL	NATURE	SOURCE	COMMENT_OBS
AF	SI_POV_NAHC	R	2007	38,2	2007	CA	Источник Всемирно го банка	Показатели бедности не сопоставимы из-за сезонности и изменений в используемых вопросниках Национальная черта бедности является абсолютной. Данные производятся Всемирным банком. Последние сопоставимые ряды: 2007, 2011. Полные ряды можно загрузить с банка данных, http://databank.worldbank.org/data/reports.aspx?source=poverty-and-equity-database .
AF	SI_POV_NAHC	U	2007	28,9	2007	CA	Источник Всемирно го банка	Показатели бедности не сопоставимы из-за сезонности и изменений в используемых вопросниках Национальная черта бедности является абсолютной. Данные производятся Всемирным банком. Последние сопоставимые ряды: 2007, 2011. Полные ряды можно загрузить с банка данных, http://databank.worldbank.org/data/reports.aspx?source=poverty-and-equity-database .
AF	SI_POV_NAHC	_T	2007	36,3	2007	CA	Источник Всемирно го банка	Показатели бедности не сопоставимы из-за сезонности и изменений в используемых вопросниках Национальная черта бедности является абсолютной. Данные производятся Всемирным банком. Последние сопоставимые ряды: 2007, 2011. Полные ряды можно загрузить с банка данных, http://databank.worldbank.org/data/reports.aspx?source=poverty-and-equity-database .



Глобальная компиляция ЦУР

Резолюция, принятая Генеральной Ассамблеей в отношении работы Статистической комиссии, касающейся Повестки дня в области устойчивого развития на период до 2030 года ([A/RES/71/313](#)):

«просит Генерального секретаря продолжать вести базу данных глобальных показателей достижения целей в области устойчивого развития и использовать ее при подготовке ежегодного доклада о ходе осуществления этих целей, а также обеспечивать транспарентность в отношении данных, статистических данных и метаданных, представляемых по странам и используемых для расчета региональных и глобальных агрегированных показателей»



Глобальная компиляция ЦУР

- Более 50 агентств
- 132 показателя с данными
- ~ 460 000 наблюдений



Первая волна

- Кратковременное решение: Использование ОСД ЦРТ
- Обмен на основе Excel
- Созданные вручную шаблоны Excel
- Инструменты SDMX для загрузки данных, но ручной проверки
- Быстрая настройка, но неудобная
- Поставщики данных не уверены в дополнительных величинах и других полях



Вторая волна

- Специальное решение
- Удобный интерфейс для загрузки Excel
- Интегрированная система управления версиями и выпуска для архивирования
- На основе ЦУР с дополнительными величинами
- Автоматическое создание шаблонов Excel из БД с учетом предыдущих данных
- Не полностью соответствует стандарту SDMX, но может быть преобразован без потери данных.



Извлеченные уроки

- ЦУР гораздо сложнее, чем ЦРТ
- Обновление показателя может создать трудности для реализации
- Инструменты эффективные, но специальное решение является идеальным вариантом (повышение производительности на 70%)
- Ориентация на использование условленных соглашений о наименовании: Перечни кодов
- Не всегда очевидно, какие значения следует использовать в некоторых величинах. (_T, NA, F) против отсутствия величины
- SDMX имеет ограничения по распространению данных (нижние и верхние границы, XML, Json,...)





**SUSTAINABLE
DEVELOPMENT GOALS**

Спасибо!