

Информационная справка о международной практике сдачи статистических данных в формате SDMX и взаимосвязи форматов SDMX и XBRL

Спонсорами и разработчиками стандарта SDMX (ISO 17369) являются международные организации, деятельность которых связана с составлением, анализом и предоставлением статистических данных:

1. Банк международных расчетов (БМР);
2. Европейский центральный банк (ЕЦБ);
3. Статистическое бюро Европейского Союза (Евростат);
4. Международный валютный фонд (МВФ);
5. Организация экономического сотрудничества и развития (ОЭСР);
6. Организация объединенных наций (ООН);
7. Всемирный банк.

SDMX применяется для обмена официальной статистикой между международными организациями, а также между международными и государственными организациями (например, центральными банками и статистическими агентствами в странах-членах международных организаций). В случае с Европейской статистической системой (ESS) это обмен между Статистическим бюро Европейского Союза (Евростат) и национальными статистическими управлениями государств-членов ЕС.

Тем не менее, в настоящее время SDMX также используется организациями за пределами мира официальной статистики. Формат SDMX может представлять интерес для любой организации, которая собирает, обрабатывает, анализирует и распространяет статистические данные и метаданные.

В международной практике наиболее распространенными являются два варианта применения SDMX:

1. SDMX в качестве формата представления и сбора отчетности.

Особенно распространено в сообществе центральных банков (в результате более раннего внедрения формата GESMES/TS, известного сейчас как SDMX-EDI), а также среди статистических агентств в Европе (исторически они также были пользователями GESMES).

2. Распространение статистических данных с помощью веб-сайтов.

Наблюдается в широком спектре учреждений, включая центральные банки (ЕЦБ и Европейская система центральных банков, БМР, Федеральная резервная система США и Федеральный резервный банк Нью-Йорка), другие спонсорские организации (МВФ, Всемирный Банк, ОЭСР) и национальные статистические агентства (INEGI в Мексике, Статистическое управление Новой Зеландии, Австралийское бюро статистики, статистические управления в Европейской статистической системе и т.д.).

В меньшей степени распространено такое применение, при котором SDMX является основой для хранилища данных и других форм управления данными.

Одним из лучших примеров является Европейский центральный банк, в котором все внутренние хранилища данных построены на базе информационной модели SDMX.

В международной практике в формате SDMX собираются статистические данные по следующим категориям:

- перепись населения и демография;
- образование;
- финансовые показатели;
- экономические показатели;
- национальные счета;
- труд;
- продовольствие и сельское хозяйство, включая рыболовство;
- эпидемиология;
- транспорт;
- качество данных;
- индикаторы развития.

Для целей обмена данными в формате SDMX могут использоваться другие форматы, например, CSV или MS-Excel. Как правило, международные организации публикуют файл MS-Excel, описывающий структуру требуемой информации. Файл MS-Excel содержит всю информацию, необходимую для понимания, как информация организована в «условиях SDMX».

Кроме того, спонсорскими организациями разработан ряд программных продуктов с открытым исходным кодом, которые помогают упростить внедрение SDMX. Также увеличивается количество инструментов, предлагаемых коммерческими компаниями-разработчиками.

Пример бесплатного программного решения: «SDMX Converter», инструмент, который конвертирует статистические наборы данных между различными форматами. Это Java-приложение, разработанное Евростатом, публикуется как программное обеспечение с открытым исходным кодом. Наиболее распространенные варианты использования конвертера:

- преобразование файлов статистических данных, представленных в CSV или в виде специальных шаблонов Excel, в один из форматов SDMX-ML, принятых системами сбора, проверки и распространения данных, основанных на SDMX;
- преобразование статистических данных SDMX-ML файлов в CSV или другие текстовые форматы, принимаемые различными статистическими системами, которые не работают с SDMX.

Конкретным примером применения стандарта SDMX является обмен данными между семью международными спонсорскими организациями, которые являются членами Межведомственной группы по экономической и финансовой статистике (IAG).

В рамках инициативы «Основные глобальные показатели» предоставляются сопоставимые на международном уровне данные для стран Группы 20 (G-20) и стран с системно значимыми секторами экономики, которые не являются членами G-20. Инициатива облегчает мониторинг экономического и финансового развития этих стран. Веб-сайт инициативы, созданный в 2009 году, находится в ведении МВФ и является совместной деятельностью Межведомственной группы по экономической и финансовой статистике (IAG).

Другие примеры SDMX в международной практике:

- Федеральный резервный банк Нью-Йорка размещает на своем сайте данные о курсах валют, используя стандарт SDMX. Такие статистические ряды доступны для автоматизированного использования;

- Евростат совместно с компанией «Google» реализовал проект представления и визуализации данных. Например, при поиске в Google выражения «Unemployment rate Poland» на основе базы связанных данных выдается график, отображающий динамику этого индикатора, при этом есть возможность перейти к базе данных и определить страны, индикаторы, период времени и прочие параметры для получения данных;

- веб-сайт «Наша статистика» Европейского центрального банка и национальных центральных банков Евросистемы. Визуальное представление статистики облегчает понимание, позволяет использовать и сравнивать показатели еврозоны и национальной статистики. Информация доступна на всех официальных языках ЕС;

- приложение «Data Mapper» Международного валютного фонда используется для построения и сравнения данных из разных стран и регионов;

- мобильное приложение Европейского центрального банка «ECBstatsApp» предлагает предварительно определенные визуализации данных в нескольких разделах, например, ключевые индикаторы еврозоны, государственные финансы, курсы валют и цены.

Сравнение стандартов SDMX и XBRL

Существует историческая взаимосвязь сбора статистических данных в центральных банках Европы и представления отчетности в ЕЦБ с форматом Gesmes-TS (позже ставший известным как SDMX-EDI). К настоящему времени формат SDMX доработан до последней версии 2.1 и поддерживает другие форматы, такие как XML, JSON, CSV. Сбор статистических данных с использованием SDMX начался в 90-х годах до появления XBRL.

Для сбора и обмена статистическими данными можно использовать и XBRL. При этом следует учитывать, что SDMX был разработан в том числе для данных, представленных в виде временных рядов.

Особенности двух форматов:

✓ Данные в SDMX определяются комбинацией измерений аналогично формату XBRL, если используется спецификация «XBRL Dimensions» для описания данных, то есть каждая точка данных описывается как комбинация измерений.

✓ SDMX имеет преимущество в части работы с частотой статистических рядов. SDMX объединяет дату, частоту (ежедневно, еженедельно, ежемесячно, ежеквартально, ежегодно) и тип сбора (конец периода, сумма за период, среднее значение за период) для полного описания временного ряда, в то время как XBRL представляет данные на определенный момент времени (хотя «потoki» данных можно описывать с помощью startDate и endDate). Это особенно актуально, когда рассматриваются частоты, превышающие дневную (секундные, минутные или почасовые данные), а также для обработки еженедельных данных.

✓ Организация (entity) всегда является явным «измерением» (dimension) в XBRL, включенным в «контекст». Это может быть проблемой при моделировании статистических рядов, в которых нет реальной организации (entity) (например, национальные агрегаты или данные о курсах валют), хотя почти всегда можно найти то, что можно было бы сопоставить с понятием организации (entity), например, область данных, сектор и т.д.

✓ Оба стандарта XBRL и SDMX позволяют осуществлять валидацию данных. При этом в качестве сильной стороны стандарта XBRL следует отметить валидацию и представление данных, а у стандарта SDMX – распространение данных с использованием вэб-сервисов и обработку временных измерений.