



Organización
Internacional
del Trabajo



► Contexto SDMX: Intercambio de datos, gobernanza, artefactos

Edgardo Greising

Organización Internacional del Trabajo

Presidente del Grupo de Trabajo Técnico de SDMX



Organización
Internacional
del Trabajo



SDMX: Historia y protagonistas





Organización
Internacional
del Trabajo

¿Qué es SDMX?

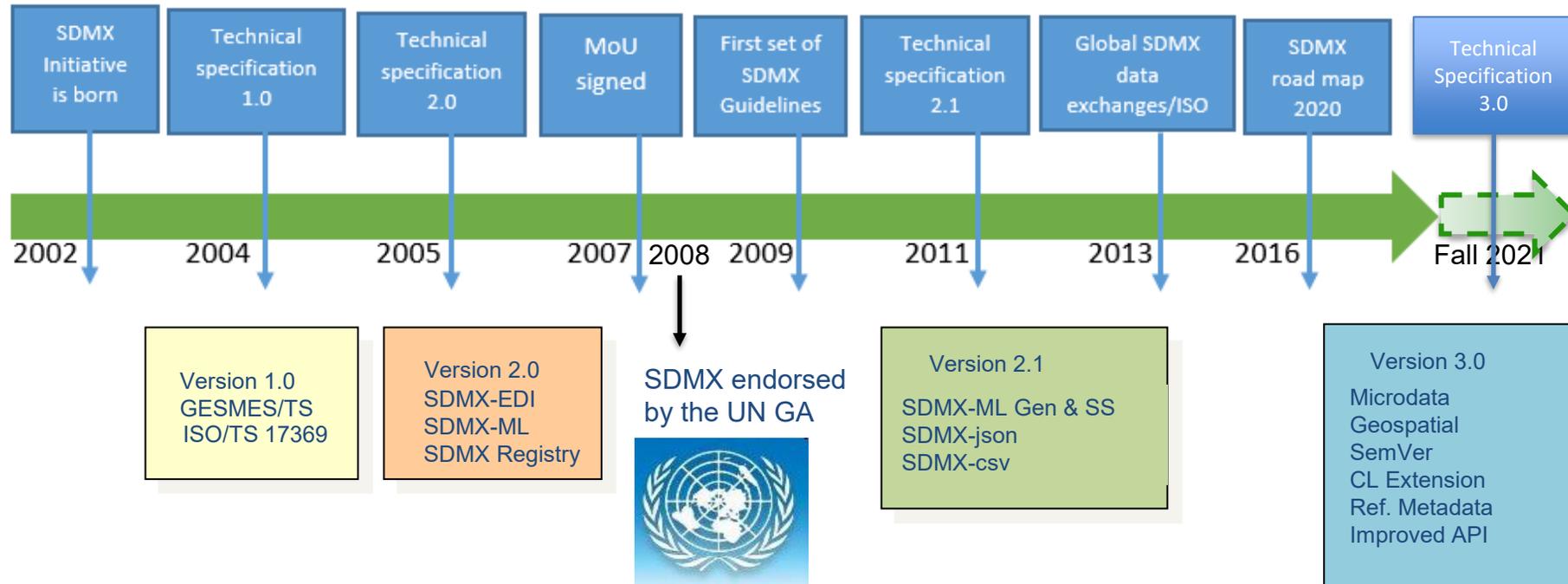


Reconocido internacionalmente y mantenido por





SDMX - Evolución





Estándar ISO

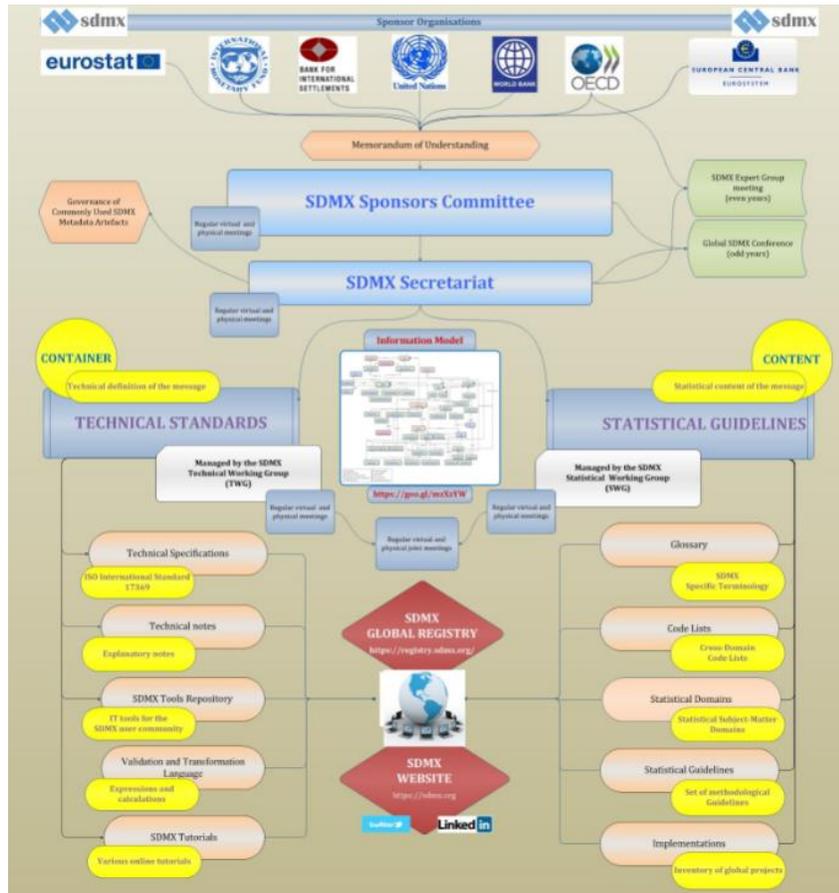


- SDMX es un estándar ISO (17369)

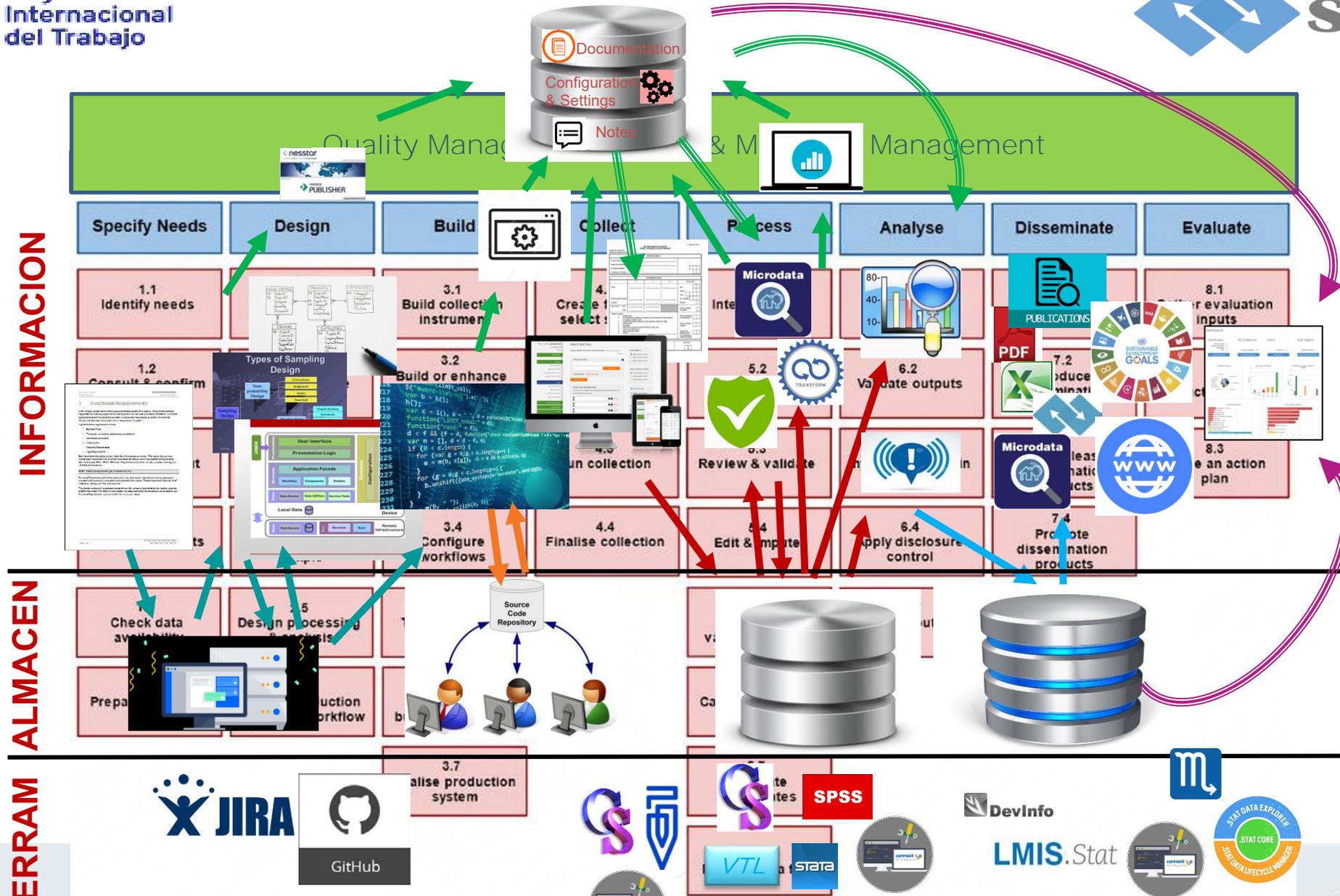




Gobernanza SDMX



- Comité de Patrocinadores
Máximo órgano de decisión que proporciona orientación general y decide sobre las cuestiones que la Secretaría de la SDMX no puede resolver..
- Secretaría de SDMX
Supervisa la aplicación y el funcionamiento del marco de gobernanza y presenta propuestas de mejora al Comité de Patrocinadores
- TWG & SWG
Mantener, mejorar y continuar desarrollando las normas técnicas y las guías estadísticas de SDMX





Organización
Internacional
del Trabajo



¿Qué encontramos en SDMX?



Principales componentes de SDMX

- Modelo de Información de SDMX (Ver. 2.1)
- Formato de mensajes SDMX

Formato	Datos	MD Estr.	MD Ref.
SDMX-EDI (GESMES/TS) (*)	✓	✓	✗
SDMX-ML Generic Data	✓	✓	✓
SDMX-ML StructureSpecific Data	✓	✓	✓
SDMX-JSON	✓	✓	D
SDMX-csv	✓	✗	✗



Principales componentes de SDMX

- API SDMX
 - RESTful web service – (Ver la «rest cheat sheet»)
 - SOAP (*)
- Especificación del Registro SDMX
 - SDMX Global Registry -- <https://registry.sdmx.org/>
- SDMX-TCK (Test compatibility kit)
 - Herramienta de prueba de "caja negra" que consulta todos los recursos y funciones especificados en la API Rest de SDMX
 - *Todavía no se ha publicado oficialmente*



Principales componentes de SDMX

- Guías SDMX
 - Listas de códigos inter-dominios de SDMX
 - Modelado de dominios estadísticos en SDMX
 - Guías para la definición de estructuras de datos en SDMX
 - Normalización de los reportes de metadatos de referencia en SDMX
 - Muchas más publicadas en https://sdmx.org/?page_id=4345
- Glosario SDMX
 - https://sdmx.org/wp-content/uploads/SDMX_Glossary_Version_2_0_October_2018.htm
- Artefactos Globales
 - Cuentas nacionales, BDP, índices de precios, **trabajo**, educación, etc.



Principales componentes de SDMX

- Herramientas desarrolladas por la comunidad
 - Registro SDMX: *Fusion Metadata Registry*
 - Servicio web SDMX: *SDMX-RI*
 - Desarrollo SDMX: *NSI WS, librerías SDMX (SDMXsource, rsdmx)*
 - Preparación de metadatos estructurales: *DSW, **DSD Constructor**, Converter, **ISTAT Meta&Data Manager***
 - Preparación de datos: *Converter, **SMART**, Fusion Transformer, **ISTAT Meta&Data Manager***
 - Difusión de datos: *.Stat Suite, **ISTAT Statkit***
 - Ver https://sdmx.org/?page_id=4500



SDMX es...



- ✓ Un **modelo de información** para describir datos y metadatos estadísticos
- ✓ Un **estándar** para la comunicación automatizada de máquina a máquina
- ✓ Una **tecnología** para desarrollar herramientas informáticas estandarizadas
- ✓ Una **iniciativa viva** respaldada por las Naciones Unidas y los productores de estadísticas oficiales, y mantenida por siete organizaciones internacionales



El modelo de información SDMX (MI)

- El modelo de información SDMX es un metamodelo que describe los objetos que intervienen en:
 - La recolección
 - La difusión
 - La publicaciónde datos estadísticos y metadatos relacionados
- Todo en SDMX se basa en el modelo de información:
 - Todos los mensajes e interfaces son implementaciones del modelo de información

¿Por qué son relevantes los metadatos?

Employment by sex, age and education (Thousands)

Source	Sex	Age	Education	2002	2003	2004	2005	2006	2007	2008	2009	2010	Trend
LFS	Total	Total	TOTAL...		1757.6	1768.8	1903.1	1946	2018.4	2059.6	2121.5		
LFS	Total	Total	X. No s...		51.8	10.5	48.1	47.5	47.4				
LFS	Total	Total	0. Pre-p...		1	55.1							
LFS	Total	Total	1. Prim...		728.1	730.6	789.6	795.2	795.6				
LFS	Total	Total	2. Lowe...		286.1	286.3	320	327.2	348.8				
LFS	Total	Total	3. Uppe...		309.6	320.7	329	342.5	374.4				
LFS	Total	Total	4. Post...		91.3	10.4	10.3	95.2	107.5				
LFS	Total	Total	5A. Firs...		129.5	355.2	136.2	148.1	149.5				
LFS	Total	Total	5B. Firs...		108		122.8	132.8	144.4	140.9	171		
LFS	Total	Total	6. Seco...		33		46.7	45.8	45	57.9	59.2		
LFS	Total	Total	? Level...		19.3		100.7	11.7	6	9.1	11.1		
LFS	Total	5-9	TOTAL...							0.1			NA
LFS	Total	5-9	X. No s...							0.1			NA
LFS	Total	5-9	0. Pre-p...										NA
LFS	Total	5-9	1. Prim...										NA
LFS	Total	5-9	2. Lowe...										NA
LFS	Total	5-9	3. Uppe...										NA
LFS	Total	5-9	4. Post...										NA
LFS	Total	5-9	5A. Firs...										NA
LFS	Total	5-9	5B. Firs...										NA
LFS	Total	5-9	6. Seco...										NA
LFS	Total	5-9	? Level...										NA
LFS	Total	10-14	TOTAL...		13.5	11.1	12	16.1	11.3	8.1	5.9		
LFS	Total	10-14	X. No s...		0.7			0.5	0.1	0.3			
LFS	Total	10-14	0. Pre-p...			0.4							NA
LFS	Total	10-14	1. Prim...		11.3	10	10.1	13.2	9.6	5.9	4.7		
LFS	Total	10-14	2. Lowe...		1.5	0.6	1.8	2.5	1.7	2.1	1.2		
LFS	Total	10-14	3. Uppe...										NA
LFS	Total	10-14	4. Post...										NA
LFS	Total	10-14	5A. Firs...										NA
LFS	Total	10-14	5B. Firs...										NA
LFS	Total	10-14	6. Seco...										NA
LFS	Total	10-14	? Level...										NA

Related metadata

Data reference period : July

Age coverage - minimum age : 12 years old

Age coverage - maximum age : No Maximum Age

Metadatos estructurales y de referencia

- *Metadatos estructurales*: Identificadores y descriptores
 - Definición de la estructura de datos
 - Esquema conceptual
 - Códigos
- *Metadatos descriptivos*: Proporciona información esencial para entender los datos
 - Desviaciones de las características previstas
 - Banderas y notas a pie de página
- *Metadatos de referencia*: Describe el contenido y la calidad de los datos
 - Definición del indicador
 - Comentarios y limitaciones

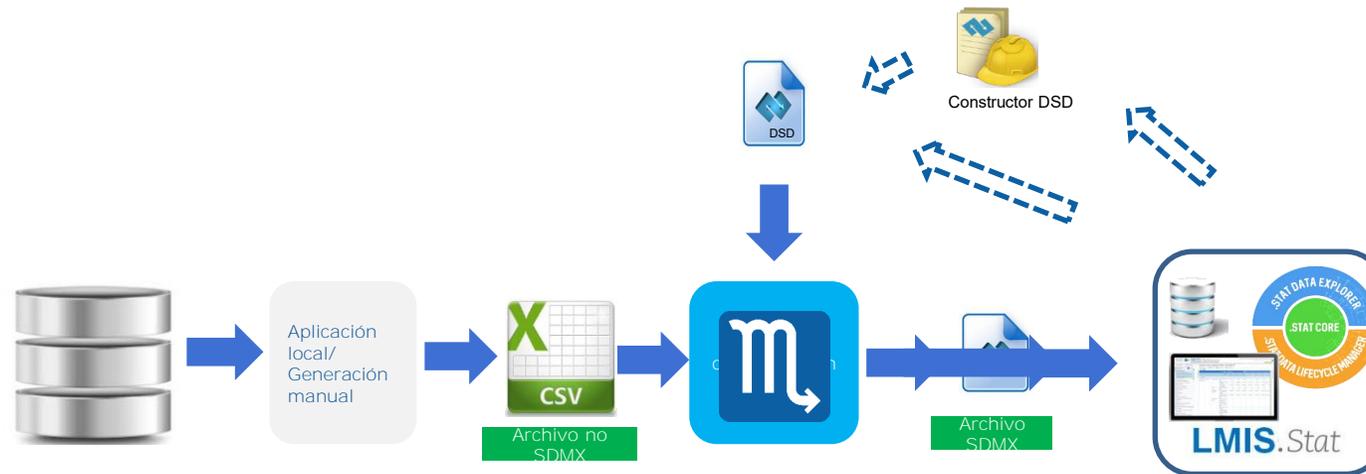
DSD
DF
Codelist
Concept
Category

Almacenado
en atributos
Viajan con
los datos

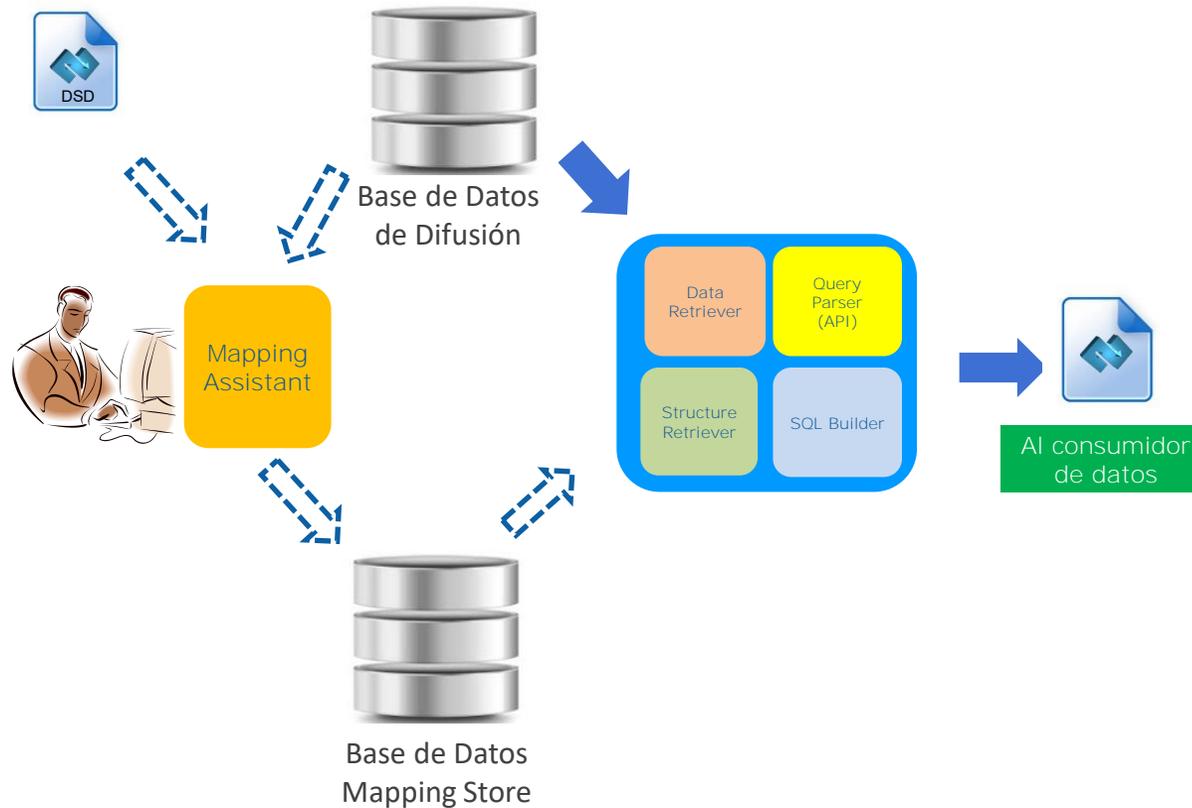
Se puede
almacenar en
atributos o en
artefactos de
metadatos de
referencia

Reporte de datos sin BD

- Sólo es necesaria una herramienta de conversión para convertir los archivos nativos en el formato SDMX deseado, utilizando el DSD proporcionado por el LMIS.



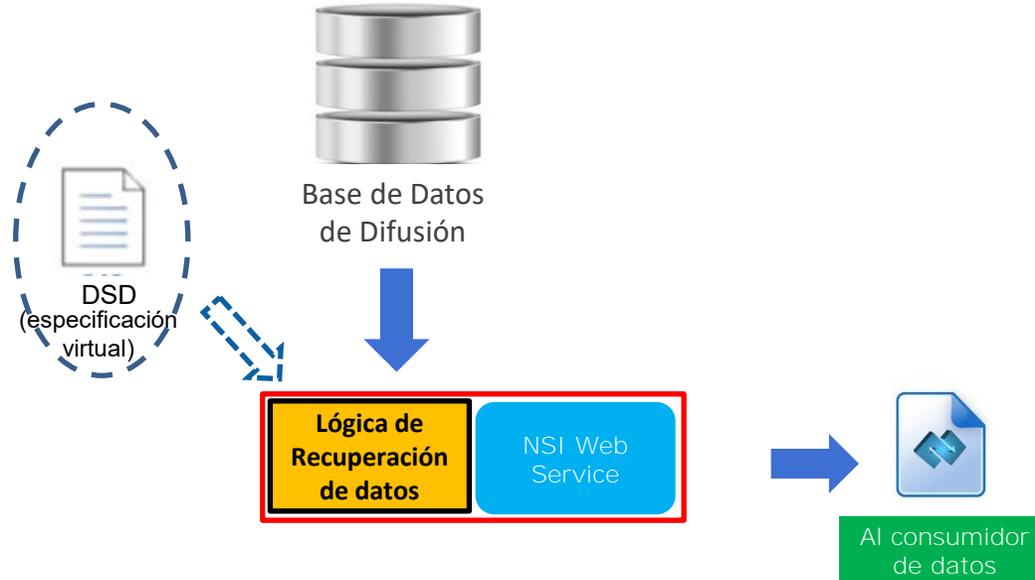
Incorporando SDMX-RI



- El Mapping Assistant mapea la estructura de la BD de Difusión al DSD (es decir, las columnas de la BD a los conceptos de SDMX)

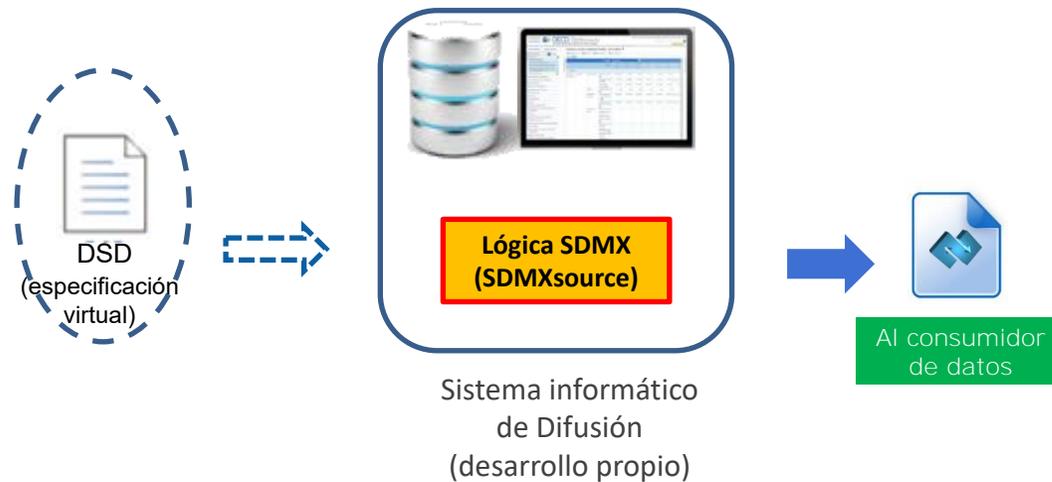
- El servicio web compatible con SDMX devuelve los datos SDMX recuperándolos de la base de datos heredada y aplicando los mapeos en el "Mapping store"

NSI WS + Lógica de recuperacion de datos



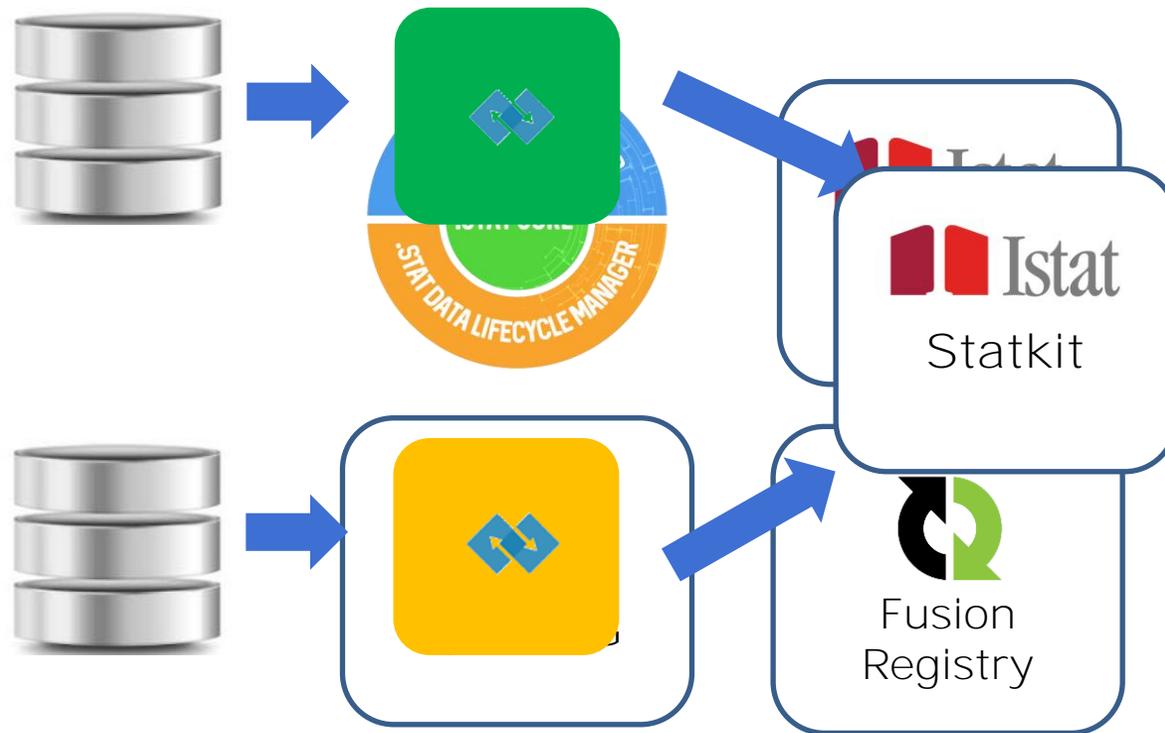
- El recuperador de datos contiene la lógica empresarial para construir los artefactos SDMX sobre la marcha.
- El servicio web compatible con SDMX devuelve los datos SDMX recuperándolos de la base de datos de difusión y aplicando la lógica del método Data Retriever

Desarrollo lógica SDMX (SDMXsource)



- El recuperador de datos contiene la lógica empresarial para construir los artefactos SDMX sobre la marcha.
- El servicio web desarrollado con la librería SDMXsource, decodifica los queries y devuelve los datos en SDMX

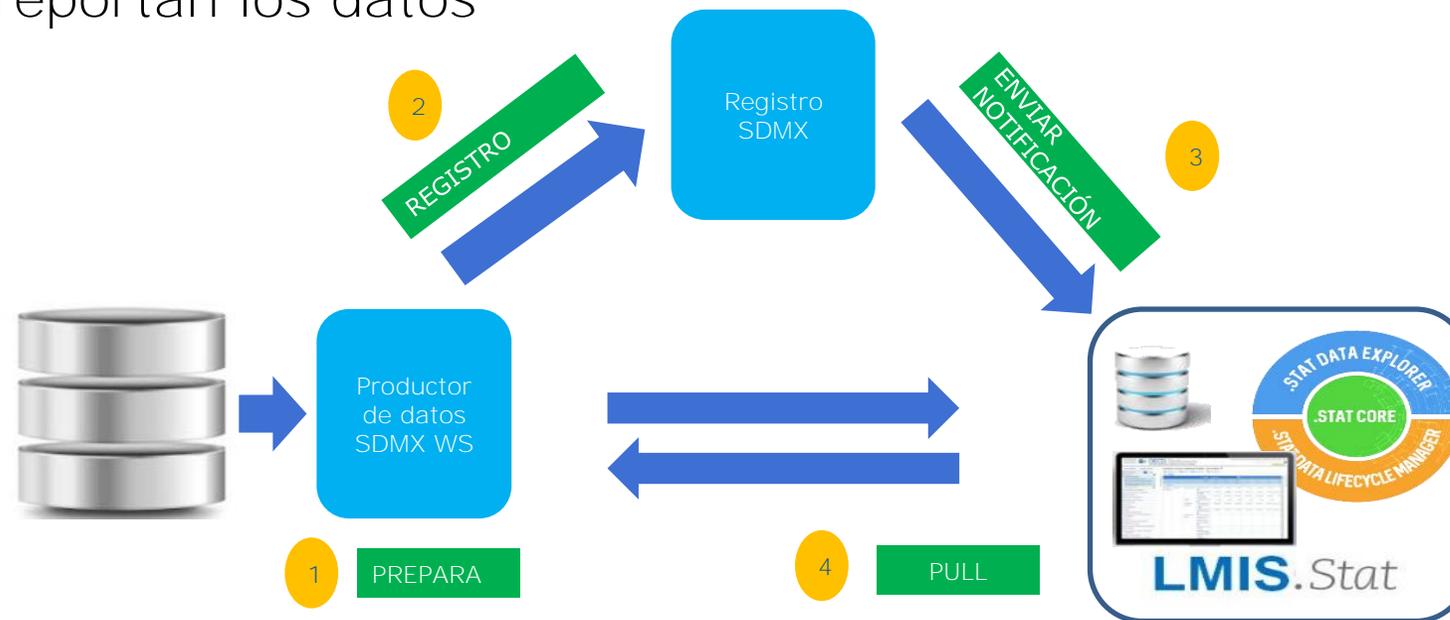
Sistema basado en SDMX



- Implementar una solución basada en una plataforma de difusión/reporte de datos basada en SDMX.
- Algunas soluciones ofrecen acceso federado a múltiples fuentes de datos (puntos de entrada SDMX) simultáneamente.

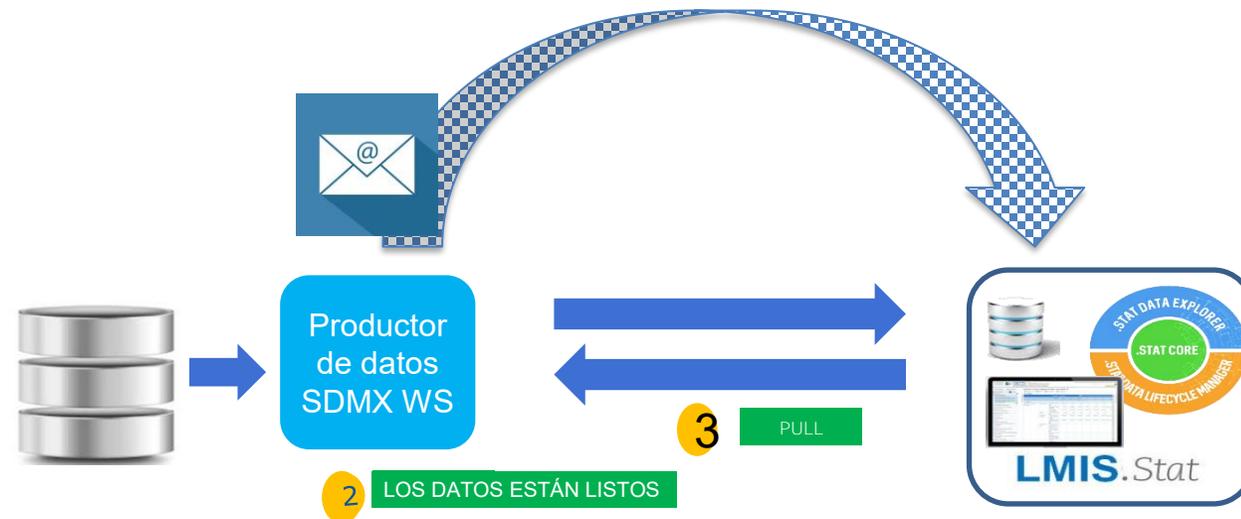
Interoperabilidad modo pull

- El modo "pull" requiere que los productores expongan sus datos a través de servicios web y que se registren en el registro SDMX utilizado por las organizaciones a las que se reportan los datos

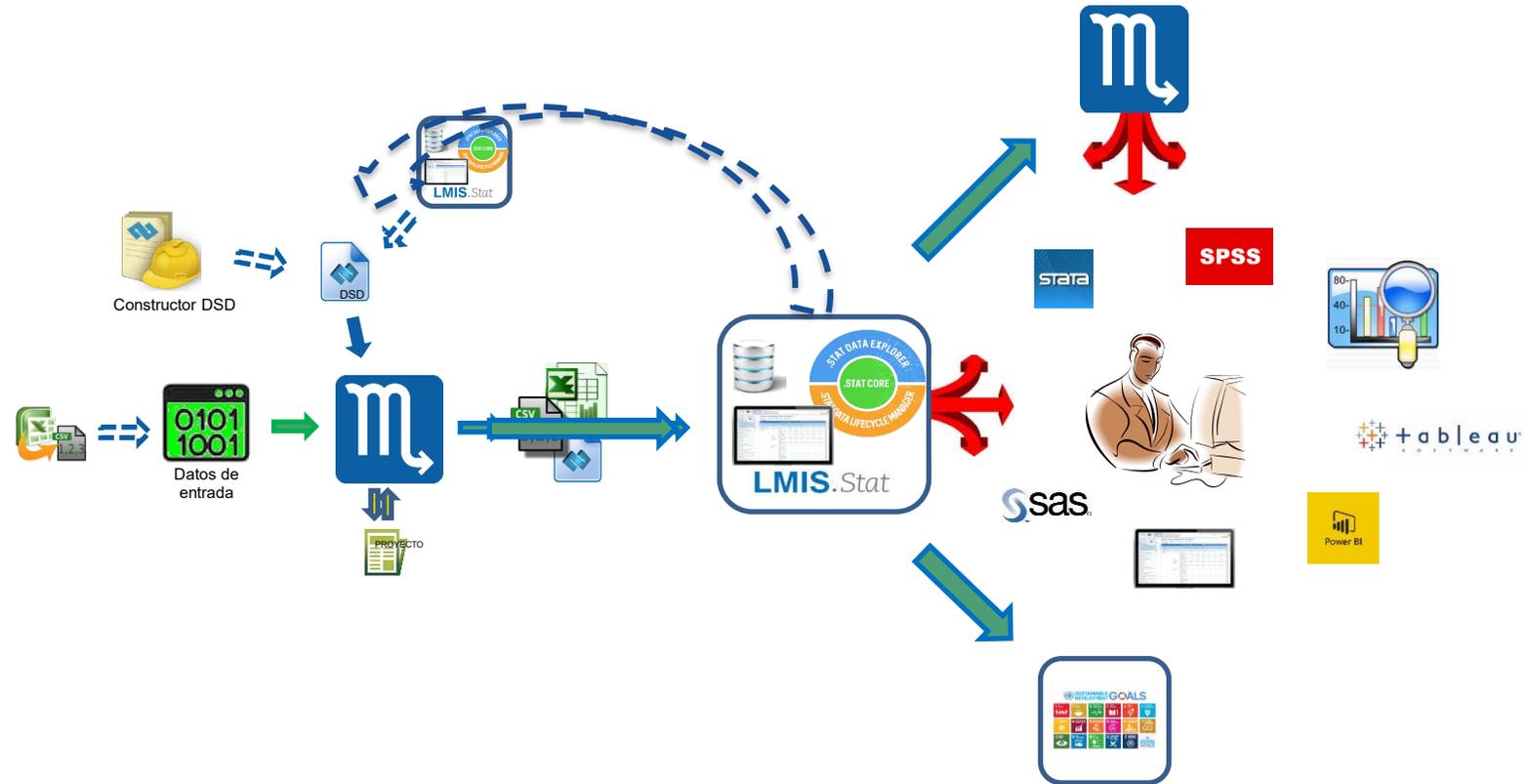


Interoperabilidad modo pull

- En la práctica, muchas veces existe un cronograma pre-acordado, o se envía la notificación via correo electrónico cuando los datos estén listos a la(s) organización(es) que los recoge.



Flujos de información SIMEL





Beneficios del estándar SDMX

- Estándar internacional para unificar todos los elementos que intervienen en el proceso estadístico con un modelo de información definido
- Sistema que maximiza la cantidad de información estadística que puede transmitirse a los usuarios, y tiene la capacidad de vincular estadísticas de diversas fuentes facilitando la armonización
- Automatización de procesos: SDMX permite el desarrollo de herramientas que pueden ser utilizadas por todas las organizaciones estadísticas para mejorar sus actividades



Beneficios del estándar SDMX

- Los estadísticos acuerdan utilizar una descripción común para los datos y metadatos
→ **ARMONIZACION** → **COMPARABILIDAD**
- El proceso de intercambio de datos se rige por definiciones comunes compartidas
→ **«METADATA DRIVEN»** → **AUTONOMIA**
- Las descripciones de los datos se ponen a disposición de todo aquel que quiera entenderlos y reutilizarlos
→ **REUSABILIDAD** → **INTEROPERABILIDAD**

Beneficios del estándar SDMX

SDMX es también un estándar avanzado para el acceso a datos mediante servicios web



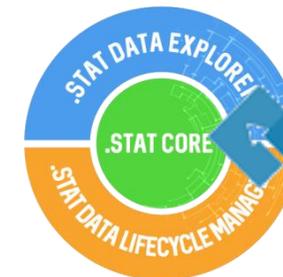
Los servicios web permiten la consulta, la visualización y la carga automática de datos y metadatos. Las herramientas de SDMX permiten consultar una base de datos, o un sistema de archivos, para la creación de tablas, cuadros y gráficos a partir de los resultados de la consulta.

Organización estadística

Infraestructura
SDMX



Sistema de información estadística





Producción/procesamiento de datos mejorado

- Estándar internacional:
 - Evita la duplicación de esfuerzos en el desarrollo y mantenimiento de herramientas informáticas
 - Aúna experiencia y recursos (comunidad de desarrolladores)
- Estandariza la estructura de los conjuntos de datos y el formato de transmisión
 - Los formatos heterogéneos se sustituyen por uno estándar
 - Las nuevas recopilaciones de datos se ponen en marcha con una configuración normalizada
- Validación de datos
 - SDMX permite detectar las inconsistencias de la estructura de datos en una fase anterior del procesamiento
- Aumento de la eficiencia
 - **Más tiempo para el análisis y la comprobación de la calidad de los datos**



Organización
Internacional
del Trabajo



Para finales de 2021: SDMX 3.0



SDMX 3.0





Érase una vez...

- "Hoja de ruta 2020" aprobada a principios de 2016
- Objetivos estratégicos:
 1. Fortalecer la implementación de SDMX;
 2. Facilitar el uso de datos a través de SDMX (especialmente para la formulación de políticas);
 3. Utilizar el SDMX para modernizar los procesos estadísticos, así como mejorar continuamente las normas y la infraestructura informática;
 4. Mejorar la comunicación en general, incluida una mejor interacción entre socios internacionales
- La aplicación de la hoja de ruta requería nuevas funciones no disponibles en SDMX 2.1



Nuevas características SDMX 3.0

- Mejorar el intercambio de metadatos de referencia
- Mejor soporte para el intercambio de microdatos
 - Múltiples medidas
 - Múltiples valores para los atributos
- Estandarizar el intercambio de datos geoespaciales
- Extensión/composición de listas de códigos
- Versionado semántico
- Mejoras de los artefactos:
 - Constraints
 - Structure set
 - HCL



Nuevas características SDMX 3.0

- Mejorar la API Restful
 - Soporte para metadatos de referencia
 - Consulta de artefactos estructurales
 - Consulta de datos mejorada
 - Mejor documentación
- Eliminación de los formatos EDI y ML obsoletos

Formato	Datos	Stru.MD	Ref.MD
SDMX-ML	✓	✓	✓
SDMX-JSON	✓	✓	✓
SDMX-csv	✓	TBD	X

- Eliminación de la API SOAP



Organización
Internacional
del Trabajo



Muchas gracias

contact@sdmx.org
twg@sdmx.org
greising@ilo.org