

Uso e implementación de SDMX en América Latina: de la teoría a la práctica

June 10

SDMX Istat Toolkit

Metadata and data management, dissemination and reporting

Francesco Rizzo – ISTAT, Italy
rizzo@istat.it

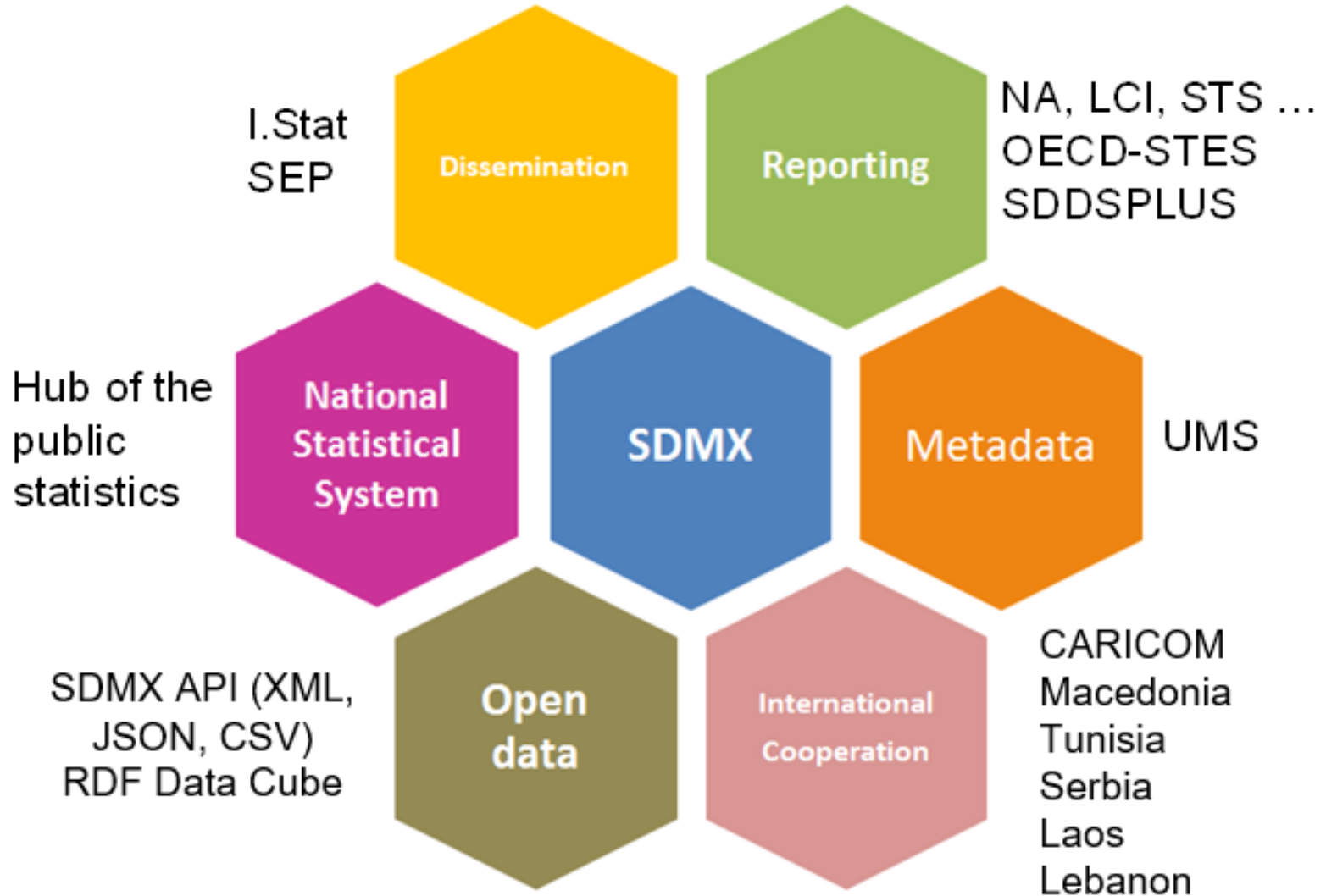
- ❑ SDMX use cases
- ❑ Herramientas SDMX disponible en sdmx.org
- ❑ El Toolkit como fundamento de la estrategia Istat SDMX
- ❑ SDMX Istat Toolkit (StatKit)
 - ❑ Excel2CSV
 - ❑ Meta and Data Manager
 - ❑ Data Browser

- ❑ Enviar datos a organizaciones internacionales (p.ej. Cuentas nacionales, SDDS Plus, etc.)
- ❑ Difundir datos y metadatos en modalidad de máquina a máquina (API)
- ❑ Descubrir y visualizar datos (GUI)
- ❑ Gestionar base de datos con un esquema conforme al modelo de información SDMX

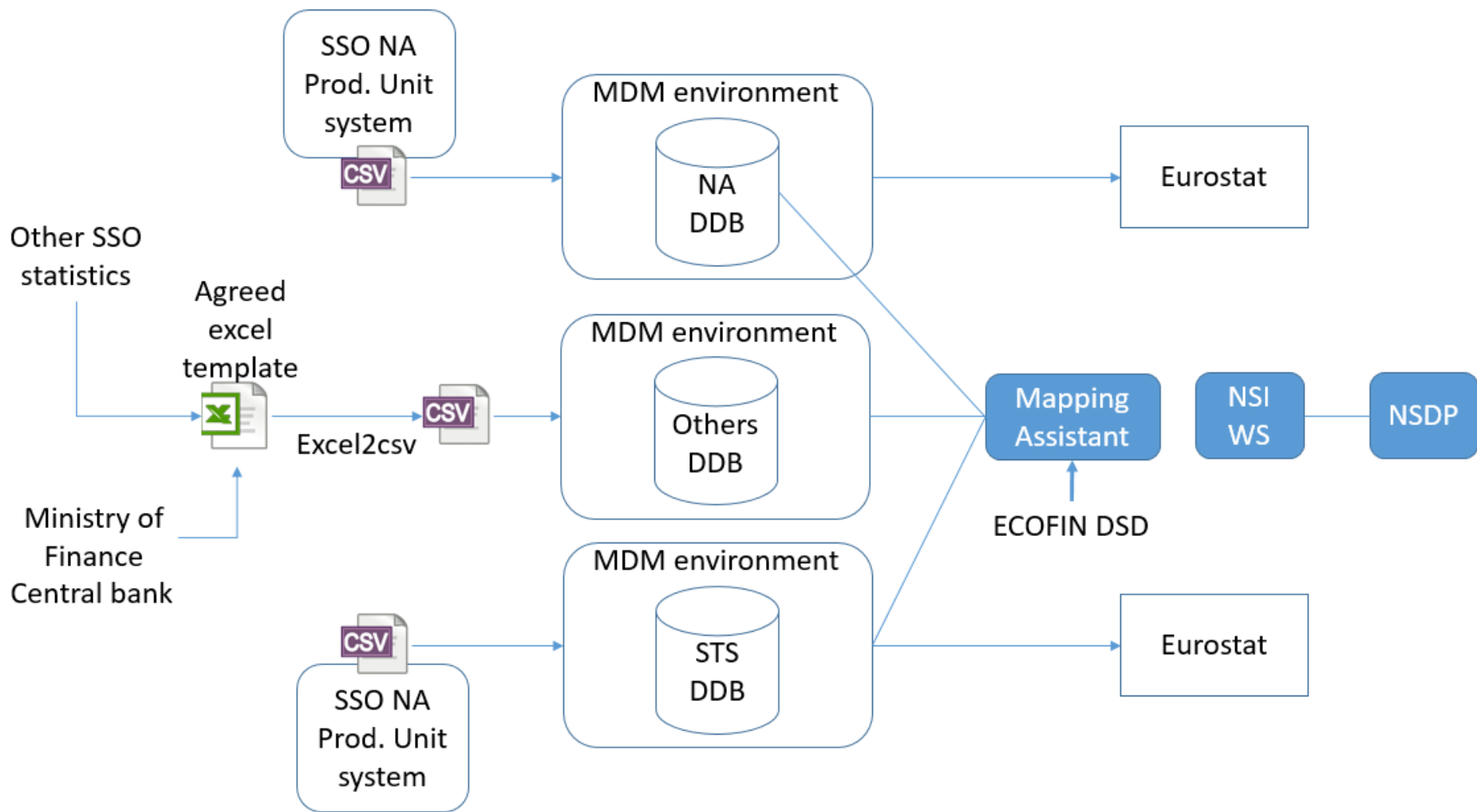
Herramientas SDMX disponible en sdmx.org

	Transformation/ Conversion	Data Modelling	Metadata management	Data management	Reporting	Dissemination GUI
Eurostat	SDMX Converter	Eurostat Registry				Web Client (GUI) NSI WS (API)
		Data Structure Wizard				
ILO	SMART	DSD Constructor				
ISTAT	Excel2CSV	Meta and Data Manager				Data Browser (GUI) NSI WS (API)
OECD		SDMX Matrix Generator		DLM		Data Explorer NSI WS (API)
Metadata Technology	Fusion Transformer	Fusion Registry				
	FusionXL					

El Toolkit como fundamento de la estrategia Istat SDMX



State Statistical Office (SSO) of the Republic of Macedonia use case



SDMX Istat Toolkit (StatKit)

<https://sdmxistattoolkit.github.io/index.html>

Principios utilizados para el diseño y la implementación

- Reingeniería basada en la experiencia
 - +500 comentarios recopilados de usuarios nacionales e internacionales
- Fácil de instalar (Plug and Play):
 - No más de 20 min
 - Habilidades técnicas fáciles de encontrar en las organizaciones estadísticas
- Fácil de configurar
 - Configuración predeterminada
 - Configuración extendida (a través de GUI)
- Fácil de cambiar de estilo y de localizar en diferentes idiomas
 - CSS files
 - Resource files
- Adecuada performance
 - Data Browser visualiza una tabla con 100.000 celdas en 10 segundos
 - Benchmark hecho con .Stat V.7 y .Stat V.8
- Fácil de usar (diseño de GUI impulsado por la experiencia de los usuarios)

Free and open source toolkit

Dissemination

Reporting

Collection

Data warehousing

Metadata handling

Data Source
transformation

Metadata
management

Data
management

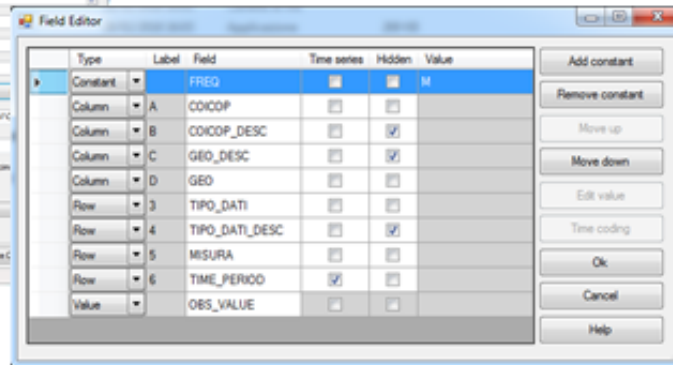
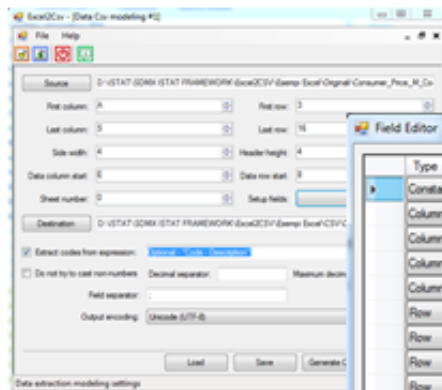
Data and Metadata
Dissemination

Excel2CVS

Meta & Data Manager

Data Browser
SDMX API
CKAN API

Excel2CSV



Index Type		Consumer price index for the whole nation (base 2015=100) - monthly data									
Measure		[4] Index Number					[6] Percentage changes on the previous				
Time period		Oct-2011	Nov-2011	Dec-2011	Jan-2012	Feb-2012	Oct-2011	Nov-2011	Dec-2011	Jan-2012	
COICOP Rev. Istat	Territory	ITC	ITD	ITE	ITF	ITG	ITC	ITD	ITE	ITF	
01 Food and non-alcoholic beverages	Nord-ovest	103,0	104,3	104,3	104,0	105,2	0,6	0,7	0	0,3	
	Nord-est	103,1	103,9	104	104,3	105,3	0,4	0,8	0,1	0,3	
	Centro	102,5	103,4	103,3	103,5	104,6	0,3	0,9	-0,1	0,2	
02 Alcoholic beverages and tobacco	Nord-ovest	107,2	107,3	107,4	107,5	107,7	3,6	0,1	0,1	0,1	
	Nord-est	107	107,1	107,1	107,2	107,3	3,5	0,1	0	0,1	
	Centro	107,4	107,4	107,4	107,6	107,6	3,7	0	0	0,2	
03 Clothing and footwear	Nord-ovest	102,6	102,7	102,6	102,9	102,7	0,6	0,1	0,1	0,1	
	Nord-est	102,4	102,5	102,6	102,8	102,7	0,6	0,1	0,1	0,2	
	Centro	103,1	103,2	103,2	103,3	103,4	1,4	0,1	0	0,1	



```

1 FREQ;COICOP;GEO;TIPO_DATI;MISURA;TIME_PERIOD;
2 M;01;ITC;9;4;Ott-2011;103,6
3 M;01;ITC;9;4;Nov-2011;104,3
4 M;01;ITC;9;4;Dic-2011;104,3
5 M;01;ITC;9;4;Gen-2012;104,6
6 M;01;ITC;9;4;Feb-2012;105,2
7 M;01;ITC;9;6;Ott-2011;0,6
8 M;01;ITC;9;6;Nov-2011;0,7
9 M;01;ITC;9;6;Dic-2011;0
10 M;01;ITC;9;6;Gen-2012;0,3
11 M;01;ITC;9;6;Feb-2012;0,6
12 M;01;ITC;9;7;Ott-2011;3,3
13 M;01;ITC;9;7;Nov-2011;3,7
14 M;01;ITC;9;7;Dic-2011;3,5
15 M;01;ITC;9;7;Gen-2012;3
16 M;01;ITC;9;7;Feb-2012;3,1
    
```

Excel2CSV - ejemplo de tabla Excel compleja

	A	B	C	D	E	F	G	H	I	J	K
2											
3					9						
4		Index Type			Consumer price index for the whole nation (base 2015=100) - monthly data						
5		Measure			[4] Index Number					[6] Percentage change	
6		Time period			Oct-2011	Nov-2011	Dec-2011	Jan-2012	Feb-2012	Oct-2011	Nov-2011
7		Istat	Territory								
8	01	Food and non-alcoholic beverages	Nord-ovest	ITC	103,6	104,3	104,3	104,6	105,2	0,6	0,0
9			Nord-est	ITD	103,1	103,9	104	104,3	105,3	0,4	0,0
10			Centro	ITE	102,5	103,4	103,3	103,5	104,6	0,3	0,0
11	02	Alcoholic beverages and tobacco	Nord-ovest	ITC	107,2	107,3	107,4	107,5	107,7	3,6	0,0
12			Nord-est	ITD	107	107,1	107,1	107,2	107,3	3,5	0,0
13			Centro	ITE	107,4	107,4	107,4	107,6	107,6	3,7	0,0
14	03	Clothing and footwear	Nord-ovest	ITC	102,6	102,7	102,8	102,9	102,7	0,6	0,0
15			Nord-est	ITD	102,4	102,5	102,6	102,8	102,7	0,8	0,0
16			Centro	ITE	103,1	103,2	103,2	103,3	103,4	1,4	0,0

Excel2CSV – ejemplo de tabla Excel compleja

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
17		Gross domestic product			Gross value added Total A10			Agriculture, forestry and fishing			Mining and quarrying; manufacturing; electricity, gas, steam and air conditioning supply; water supply; sewerage, waste management and remediation activities					
18											Total		of which: Manufacturing			
19	STO ▶	B1GQ	S	S	B1G	S	S	B1G	S	S	B1G	S	S	B1G	S	S
20	ACTIVITY ▶	_Z	A	A	_T	A	A	A	A	A	BTE	A	A	C	A	A
21	ACCOUNTING ENTRY ▶	B	T	T	B	T	T	B	T	T	B	T	T	B	T	T
22	TIME ▼	1=2+16+17	S	S	2=3+4+6+..+13	S	S	3	S	S	4	S	S	5	S	S
23	1995	1244538	A	F	1122615	A	F	26413	A	F	252752	A	F	223469	A	F
24	1996	1258660	A	F	1135998	A	F	26879	A	F	251989	A	F	221712	A	F
25	1997	1282146	A	F	1154423	A	F	27604	A	F	254242	A	F	224090	A	F
26	1998	1300714	A	F	1168621	A	F	28311	A	F	256061	A	F	226052	A	F
27	1999	1319588	A	F	1182242	A	F	30074	A	F	256249	A	F	225407	A	F
28	2000	1367801	A	F	1229008	A	F	29368	A	F	264569	A	F	233876	A	F
29	2001	1393278	A	F	1252220	A	F	28607	A	F	262305	A	F	232056	A	F
30	2002	1399568	A	F	1257988	A	F	27786	A	F	261154	A	F	230313	A	F
31	2003	1398916	A	F	1255411	A	F	26493	A	F	255273	A	F	224565	A	F
32	2004	1423126	A	F	1278452	A	F	29908	A	F	259508	A	F	227912	A	F
33	2005	1436379	A	F	1291692	A	F	28600	A	F	261909	A	F	229848	A	F
34	2006	1467964	A	F	1320418	A	F	28276	A	F	272010	A	F	239639	A	F
35	2007	1492671	A	F	1344313	A	F	28332	A	F	279679	A	F	247336	A	F
36	2008	1475412	A	F	1329002	A	F	28729	A	F	271375	A	F	238470	A	F
37	2009	1394347	A	F	1254718	A	F	28007	A	F	230422	A	F	198986	A	F
38	2010	1418376	A	F	1276477	A	F	27952	A	F	244266	A	F	214249	A	F
39	2011	1424752	A	F	1284355	A	F	28105	A	F	247946	A	F	217861	A	F
40	2012	1391018	A	F	1256553	A	F	26908	A	F	240500	A	F	210230	A	F
41	2013	1365227	A	F	1236836	A	F	26980	A	F	232792	A	F	203609	A	F

Meta and Data Manager (MDM)

Metadata manager

Data manager

Data modelling

Builder

Mapping

Loader

DF builder

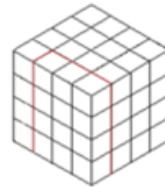
Structural metadata utilities

Reference Metadata Editor

Data catalog (DCAT) Editor

Nomenclature server

“layout” Annotations editor

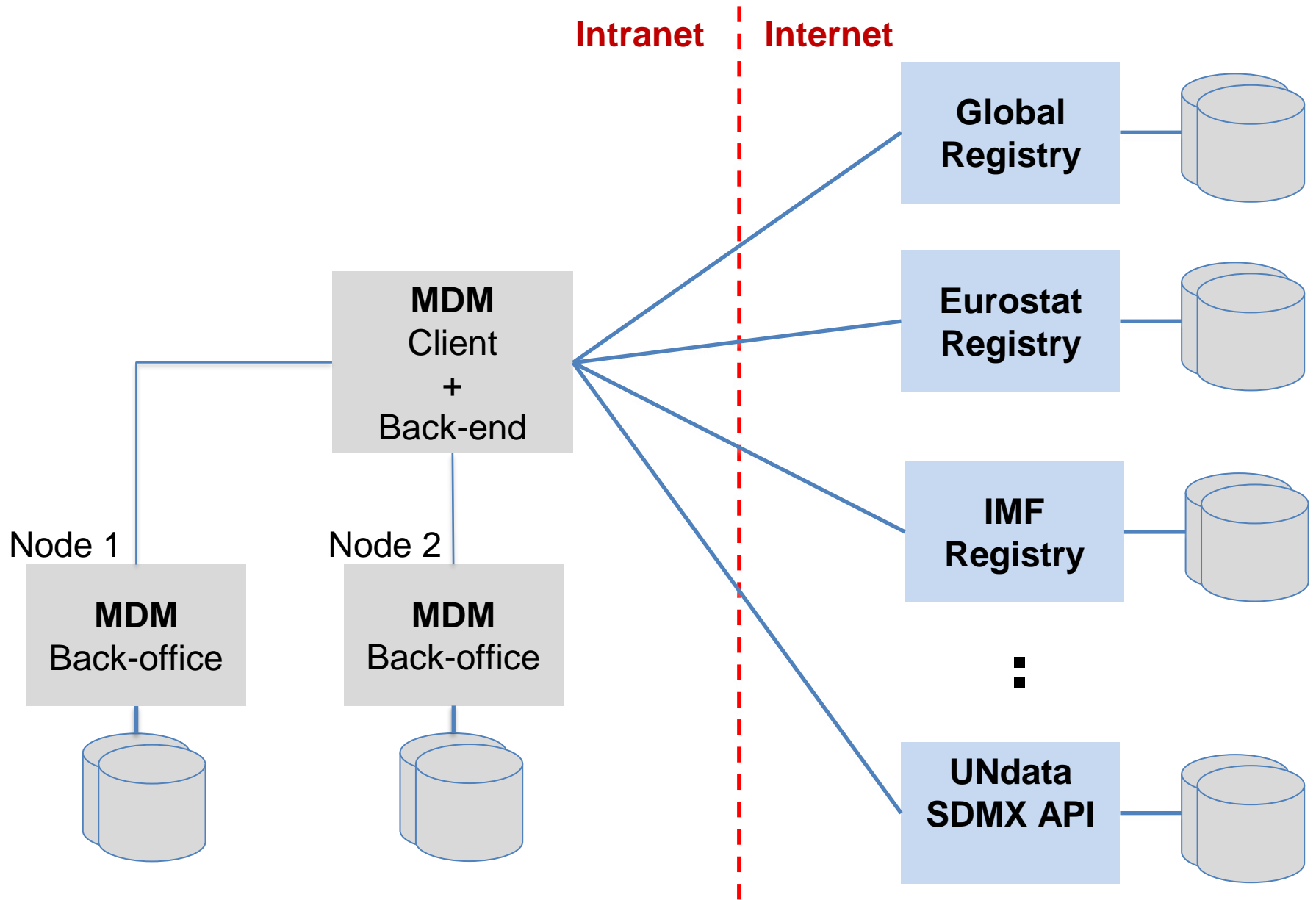


FREQ	↔	FREQ
TERRITORY	↔	TERRITORY
SEZNACE	↔	SEZNACE
LEGALFORM	↔	LEGALFORM
DEMOENT	↔	DEMOENT
TIME_PERIOD	↔	TIME_PERIOD
SOURCE	↔	SOURCE
NOTA	↔	NOTA



DBrowser cache invalidation

Meta and Data Manager – hub architecture




Meta and Data Manager – main Feature and Utilities

- ❑ **CRUD** Concept Schemes, Code Lists, Concept Schemes, Category Schemes, Agency Schemes, Data Structure Definitions, Data Flows, Categorisations
- ❑ **Usar** Hierarchical CL, Metadata Structure Definitions, Metadata Flows
- Utilities
 - ❑ **Derivar** item schemes (Code Lists, Concept Schemes, Category Schemes)
 - ❑ Incluir elementos preservando la jerarquía
 - ❑ Incluir “Parents”, “Children”, “Descendants”
 - ❑ **Combinar** item schemes (Listas de códigos, Esquemas de concepto, Esquemas de categoría)
 - ❑ **Comparar** item schemes
 - ❑ **Comparar** DSDs
 - ❑ **Actualizar** DSDs
 - ❑ **Superar** las deficiencias de SDMX 2.1: agregar elementos a los *item schemes* ya finalizados, modificar / agregar nombres, descripciones y anotaciones a los artefactos finales
 - ❑ **Administrar** gráficamente anotaciones para ordenar elementos (p. ej. listas de códigos) y para guiar el estilo de presentación de las tablas estadísticas


Meta and Data Manager


MM Meta Manager 


 Concept Schemes


 Codelists


 **Data Structure Definitions**

 Dataflows


 Category Schemes

 Agency Schemes

 Content Constraints

 Configuration 

 **Data Structure Definitions**

na main 

<input type="checkbox"/>	ID	Ag.	Name	Vers.	Final
<input type="checkbox"/>	NA_MAIN	ESTAT	NA Main Aggregates	1.8	✓
<input type="checkbox"/>	NA_MAIN	ESTAT	NA Main Aggregates	1.9	✓
<input type="checkbox"/>	NA_MAIN	ESTAT	NA Main Aggregates	1.9.1	✓
<input type="checkbox"/>	NA_MAIN	ESTAT	NA Main Aggregates	1.10	✓
<input type="checkbox"/>	NA_MAIN	ESTAT	NA Main Aggregates	1.11	✓
<input type="checkbox"/>	NA_MAIN	ESTAT	NA Main Aggregates	1.12	✓

Data Browser

Istat Istituto Nazionale di Statistica

1 NO POVERTY 	2 ZERO HUNGER 	3 GOOD HEALTH AND WELL-BEING 	4 QUALITY EDUCATION 	5 GENDER EQUALITY 	6 CLEAN WATER AND SANITATION
7 AFFORDABLE AND CLEAN ENERGY 	8 DECENT WORK AND ECONOMIC GROWTH 	9 INDUSTRY, INNOVATION AND INFRASTRUCTURE 	10 REDUCED INEQUALITIES 	11 SUSTAINABLE CITIES AND COMMUNITIES 	12 RESPONSIBLE CONSUMPTION AND PRODUCTION
13 CLIMATE ACTION 	14 LIFE BELOW WATER 	15 LIFE ON LAND 	16 PEACE, JUSTICE AND STRONG INSTITUTIONS 	17 PARTNERSHIPS FOR THE GOALS 	

Data Browser

Current activity status - Calabria

Frequency: Annual, Indicator: Resident population, Age: 15 years and over



Criteria



Pivoting



Table



Chart



Map

Territory

Calabria

Label format: Name

Year	2018			2019		
	Males	Females	Total	Males	Females	Total
Current activity status						
CURRENTLY ECONOMICALLY ACTIVE	458.008	325.315	783.323	452.055	317.377	769.432
Employed person	358.288	239.059	597.347	361.257	239.825	601.082
Unemployed person	99.720	86.256	185.976	90.798	77.552	168.350

Current activity status - Calabria

Frequency: Annual, Indicator: Resident population, Age: 15 years and over



Criteria



Pivoting



Table



Vertical Bar



Map

Primary dimension

Current activity status

8 values selected

Secondary dimension

Gender

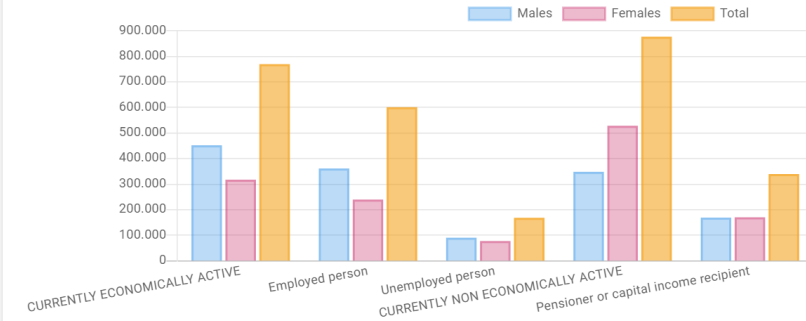
3 values selected

Territory

Calabria

Year

2019



Current activity status - Calabria

Frequency: Annual, Indicator: Resident population, Age: 15 years and over



Criteria



Pivoting



Table



Chart



Map

Gender

Total

Current activity status

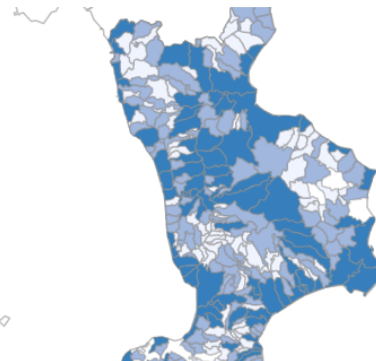
CURRENTLY ECONOMICALLY ACTIVE

Year

2019

Detail level

Municipality



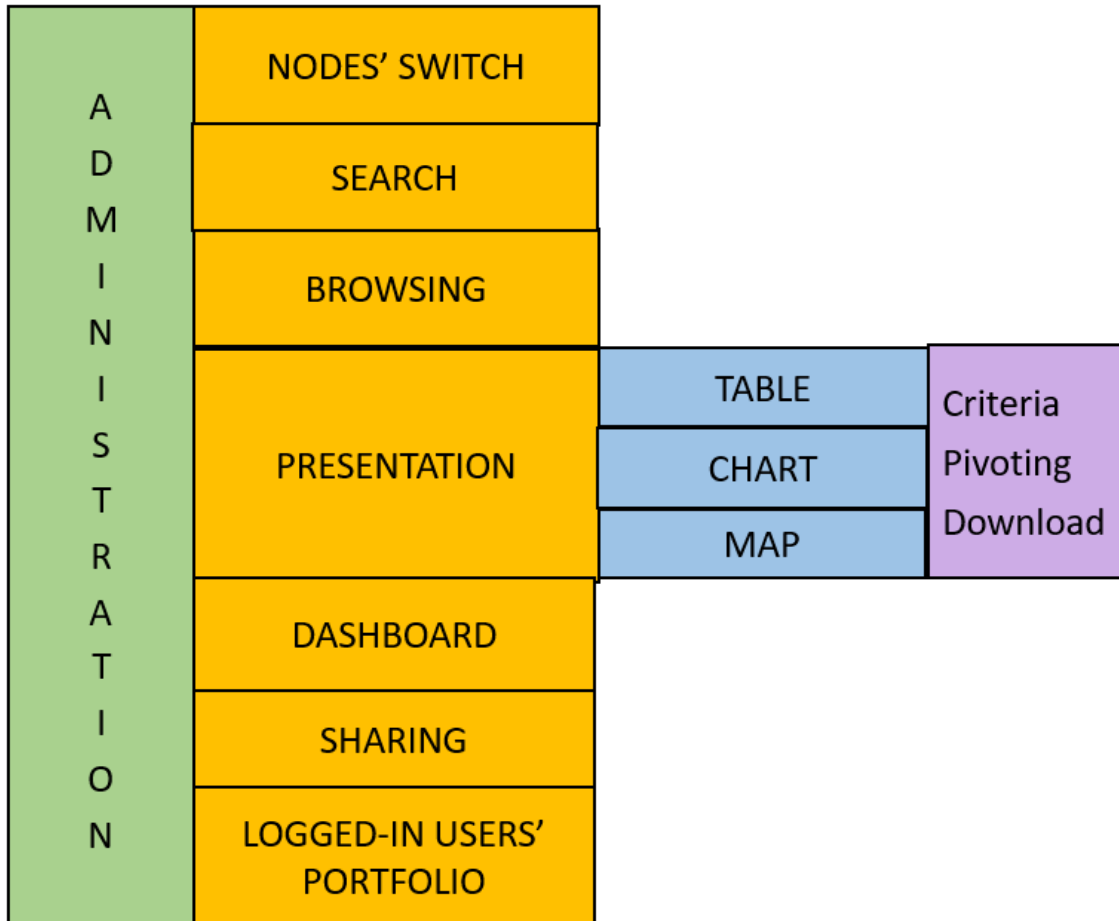
Legend

- from 63 to 506
- from 506 to 1.326
- from 1.326 to 73.563

Configure

Opacity: 90%

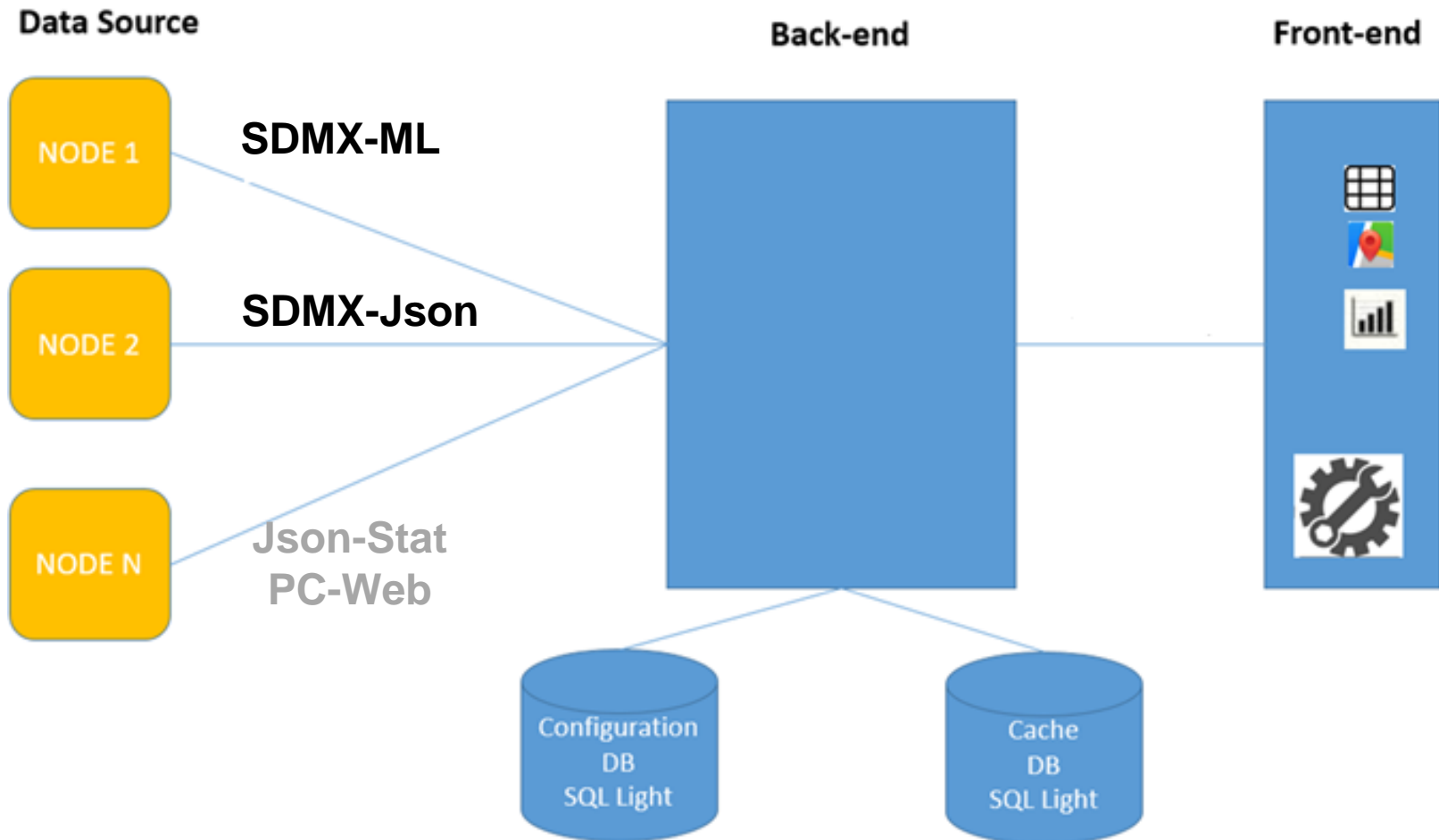
Data Browser



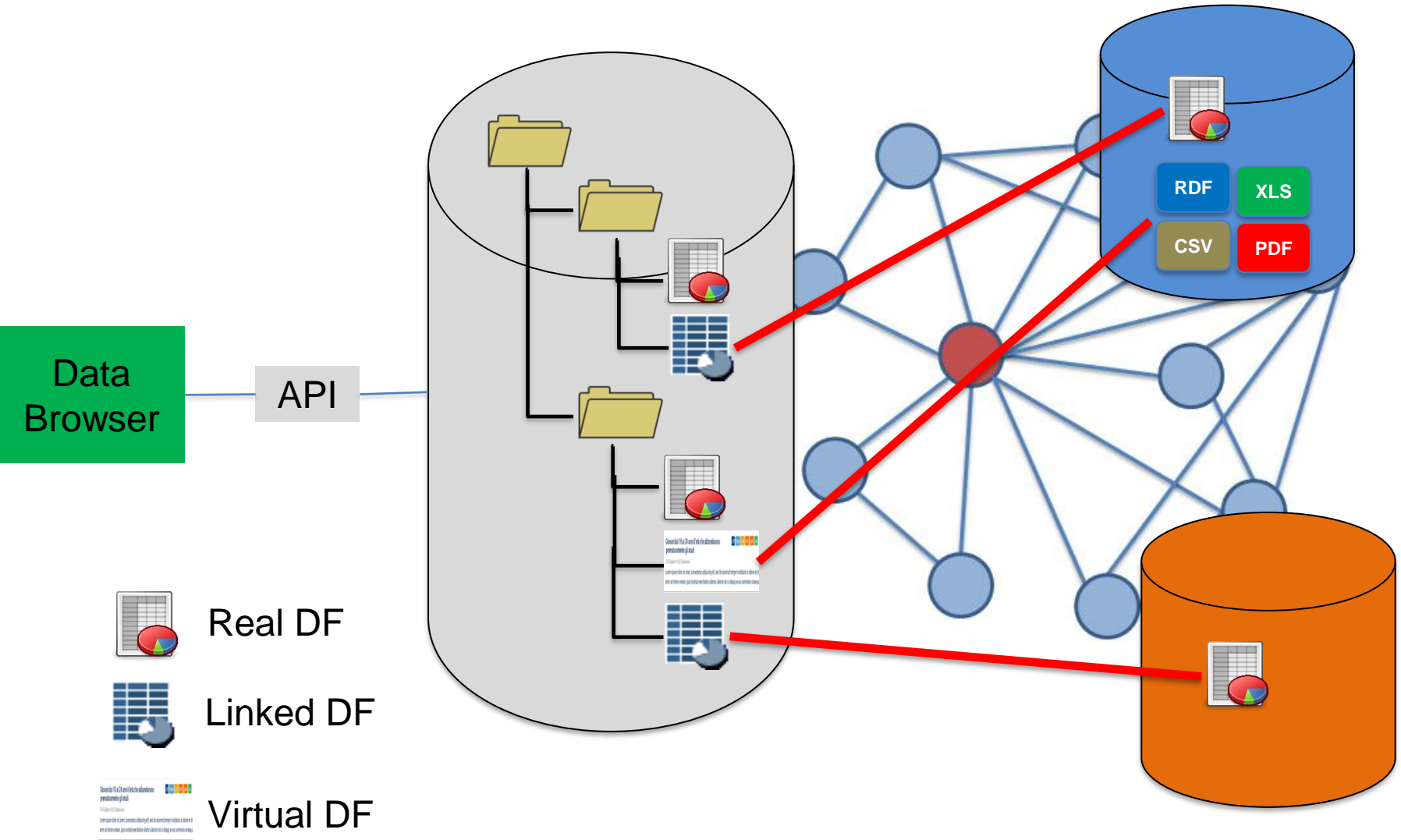
Download

SDMX – Generic 2.1
SDMX – Generic 2.0
SDMX – Compact 2.0
SDMX – Structure Specific 2.1
SDMX-CSV
SDMX-JSON
RDF/XML
Excel
Custom CSV
JSON-STAT
PC-AXIS

Data Browser – client and backend

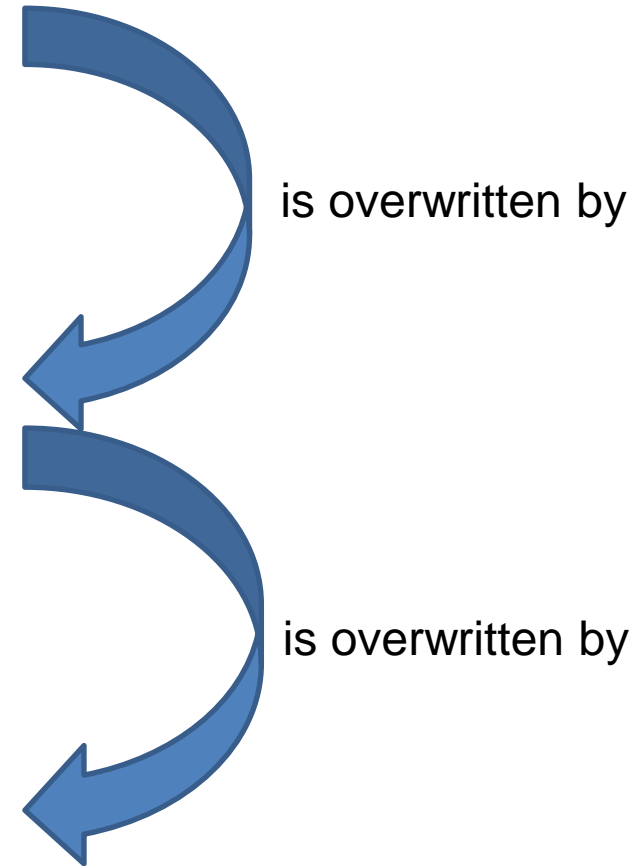


Real, Linked and Virtual DFs



Three configuration levels for presenting information

- ❑ Built-in Template (managed centrally through GUI)
- ❑ Dataflow-driven Template (annotations)
- ❑ Administrator Template (managed per Dataflows (managed centrally through GUI))



References

- ❑ SDMX Istat toolkit download:

<https://github.com/SDMXISTATTOOLKIT>

- ❑ Data Browser for the “Permanent census of population and housing”

<https://esploradati.censimentopopolazione.istat.it/databrowser/#/en>

- ❑ More information:

rizzo@istat.it (Francesco Rizzo)

alcardac@istat.it (Alessio Cardacino)

