

# 55.8 Interoperabilità cross-border e cross-domain: l'esperienza del progetto INTERSTAT

67.412

-4.306

63.186

70.841

**25**  
GENNAIO  
2024

Raffaella Maria Aracri, Adele Maria Bianco,  
Paolo Francescangeli, Giuseppina Ruocco (ISTAT)  
Francesca D'Agresti (ENGINEERING)

*#IstatWebinar*

# Indice della presentazione

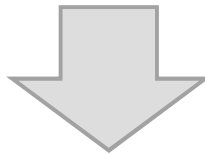
---

- Open data e interoperabilità
- Interoperabilità: framework di riferimento
- Servizi statistici cross-border e cross-domain
- Il progetto INTERSTAT: principali obiettivi
- Pilots implementati
- Focus on: SEP - Support for Environment Policies
- INTERSTAT e l'European data portal
- Conclusioni

# Open data e interoperabilità

---

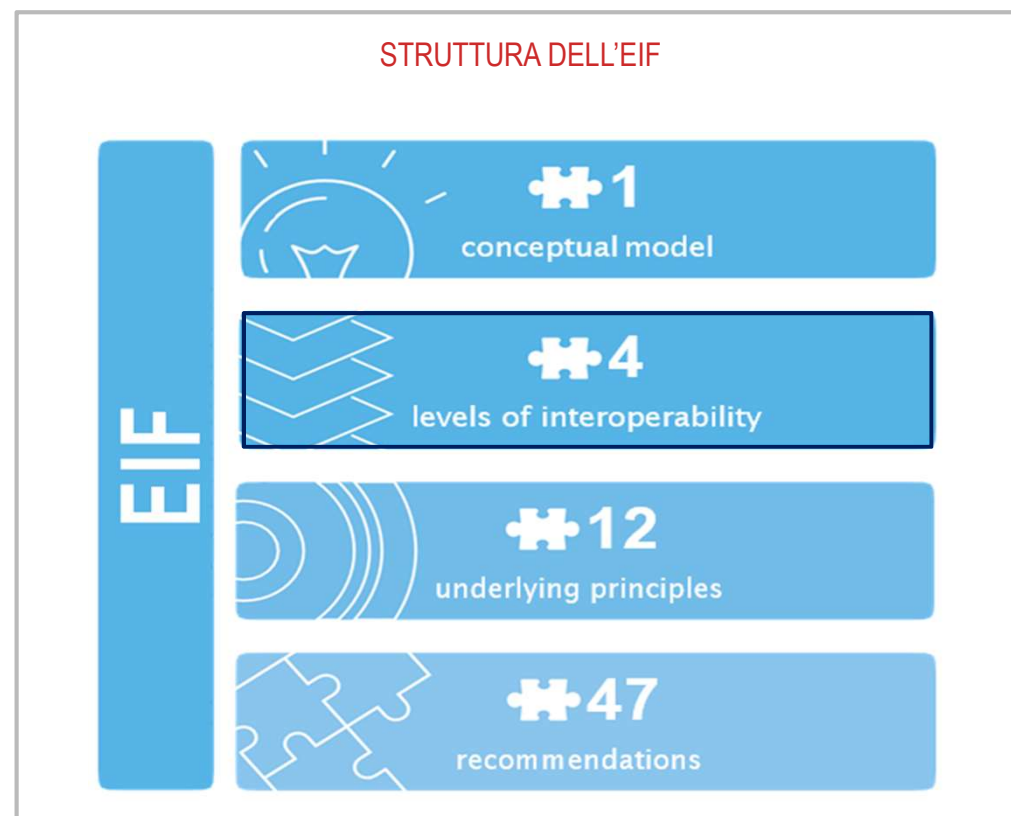
- Open Data: dati pubblicati in un formato 'machine-readable', accessibili via web e liberamente utilizzabili, nel rispetto delle indicazioni fornite nella licenza d'uso



- Implementazione di soluzioni e adozione di strumenti **interoperabili "by design"**
- Interoperabilità: Capacità di un sistema o di un prodotto informatico di cooperare e di **scambiare informazioni o servizi** con altri sistemi o prodotti in maniera più o meno completa e priva di errori, con affidabilità e con ottimizzazione delle risorse (*Fonte: Wikipedia*)

# Interoperabilità: framework di riferimento (1/3)

- In ambito europeo, l'**European Interoperability Framework**, revisionato nel 2017, ha definito le linee guida per l'interoperabilità sotto forma di principi, modelli e raccomandazioni comuni, finalizzate alla fornitura di servizi pubblici europei
- Creazione di **sistemi interoperabili** per realizzare un'efficace interazione tra dispositivi, reti e archivi di dati e incrementare l'efficienza dei collegamenti tra i diversi Stati Membri, e tra cittadini e autorità pubbliche



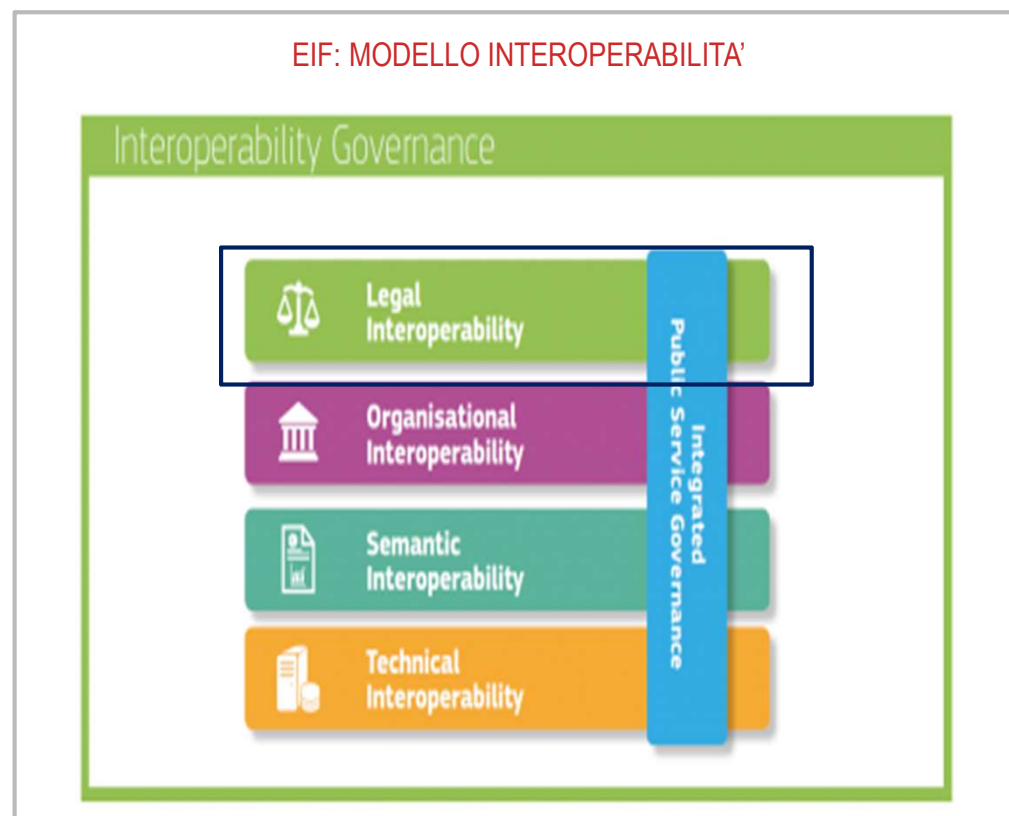
Fonte: EIF

## Interoperabilità: framework di riferimento (2/3)

Sviluppo di un framework giuridico per la produzione di dati open e il riutilizzo delle informazioni del settore pubblico mediante l'armonizzazione delle condizioni di base per tale riutilizzo e la rimozione dei principali ostacoli



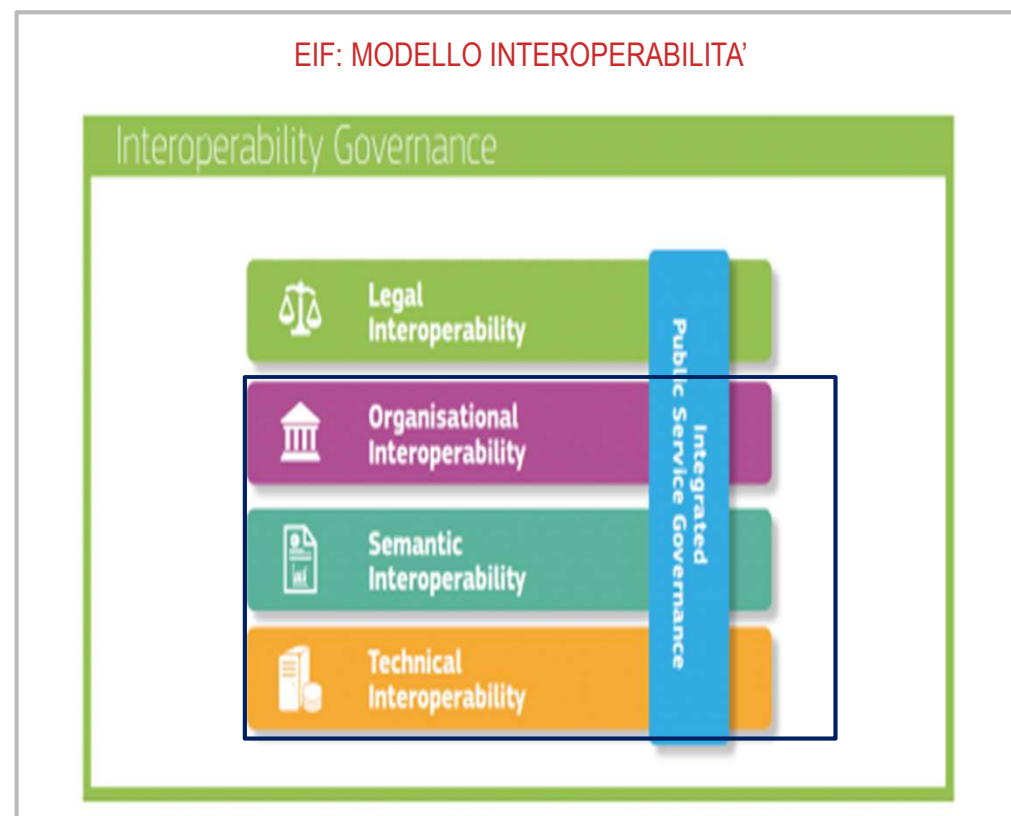
- **Direttiva (UE) 2019/1024**, relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico
- Definizione di una strategia europea e individuazione di set di dati particolarmente rilevanti (**HVD – High Value Dataset**) raggruppati in sei categorie: geospaziali, osservazione della terra e ambiente, meteorologia, statistica, imprese e relativa proprietà, mobilità



Fonte EIF

## Interoperabilità: framework di riferimento (3/3)

- Revisione e standardizzazione dei processi di creazione degli open data
- Metodi e strumenti per la pubblicazione di
  - ✓ **Linked Open Statistical Data**
  - ✓ Implementazione di **servizi statistici interoperabili**, cross-border e cross-domain



Fonte EIF

# Servizi statistici cross-border e cross-domain

---

## Interoperabilità semantica

- Modellazione del dominio di interesse mediante **ontologie** per la definizione formale dei concetti target e dei relativi metadati
- Ragionamento automatico, interoperabilità tra domini, **disaccoppiamento tra struttura e semantica dei dati**
- Utilizzo e riuso di concetti estratti da **meta-ontologie** di riferimento (e.g., Data Cube Vocabulary, SKOS, XKOS) o da ontologie di dominio già pubblicate
- Definizione di **Common Data Models** per facilitare l'armonizzazione di dati e metadati provenienti da fonti diverse

## Interoperabilità tecnica

- Sviluppo di piattaforme per facilitare l'interazione tra componenti applicative di terze parti, mediante **APIs pubbliche e open**
- Utilizzo di **componenti Open source** per promuovere l'adozione di tecnologie comuni
- Utilizzo di tools che facilitano l'internazionalizzazione, rimuovendo le barriere linguistiche
- Adozione di **standard tecnici internazionali**, ai fini della pubblicazione nell'European data portal
- Accesso all'informazione mediante diversi canali e devices, a seconda delle esigenze degli utenti finali

# Il progetto INTERSTAT: principali obiettivi

---

- Valorizzazione dei Linked Open Statistical Data (LOSD) mediante la creazione di un framework (Open Statistical Data Interoperability Framework) e lo sviluppo di soluzioni tecniche per incrementare l'interoperabilità tra i diversi portali statistici nazionali e lo European Data Portal
- Adozione di standard e ontologie comuni per l'automazione e l'armonizzazione dei metadati e condivisione dei tools utilizzati per la pubblicazione dei LOSD
- Creazione di strumenti per agevolare gli utenti finali (non tecnici) nella visualizzazione e analisi dei LOSD
- Validazione delle soluzioni tecniche sviluppate mediante l'implementazione e la sperimentazione di servizi cross-border e cross-domain e analisi dell'impatto per i potenziali utenti (tecnici e non)

- Partners:    



## Pilots implementati (1/2)

---

Test dell'interoperabilità cross-border e cross-domain



- **Servizi cross-domain:** valorizzazione del potenziale informativo derivante dall'integrazione di dati relativi a domini differenti ma correlati
- **Servizi cross-border:** analisi comparata di indicatori o variabili relativi a specifiche aree di Paesi diversi

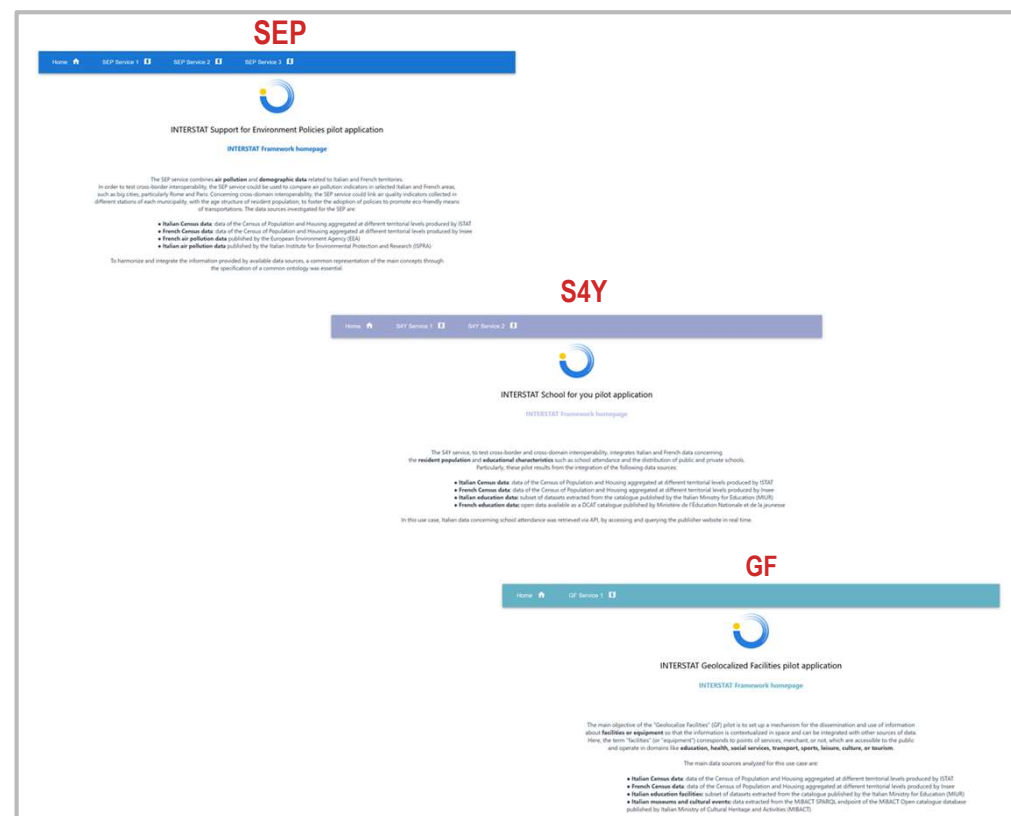
Implementazione di tre pilots al fine di validare le soluzioni tecniche per l'interoperabilità sviluppate mediante i tool del framework Interstat

Per ogni caso d'uso:

- ✓ Individuazione dei dataset da linkare
- ✓ Studio del modello dati
- ✓ Ontologia (integrazione **cross-domain**)
- ✓ Armonizzazione del modello dati italiano e francese (integrazione **cross-border**)
- ✓ Definizione della pipeline
- ✓ Potenziali utenti: cittadini e decisori politici

# Pilots implementati (2/2)

- **Support for Environment Policies (SEP):** integrazione tra dati francesi e italiani relativi alla qualità dell'aria e popolazione residente (derivante dal Censimento della Popolazione), al fine di supportare le autorità pubbliche locali responsabili delle politiche ambientali
- **Geolocalized facilities (GF):** integrazione tra informazioni relative alle strutture scolastiche pubbliche e private e popolazione residente in Italia e Francia
- **School for You (S4Y):** integrazione di dati provenienti da diverse fonti relative alla frequenza scolastica in Italia e Francia e confronto con la popolazione residente



# Focus on: SEP- Support for Environment Policies (1/3)

## Domande iniziali:

- Quali sono le aree maggiormente inquinate in Francia ed Italia?
- Quali caratteristiche ha la popolazione residente in tali aree?
- Quanto incidono le categorie di soggetti maggiormente a rischio?

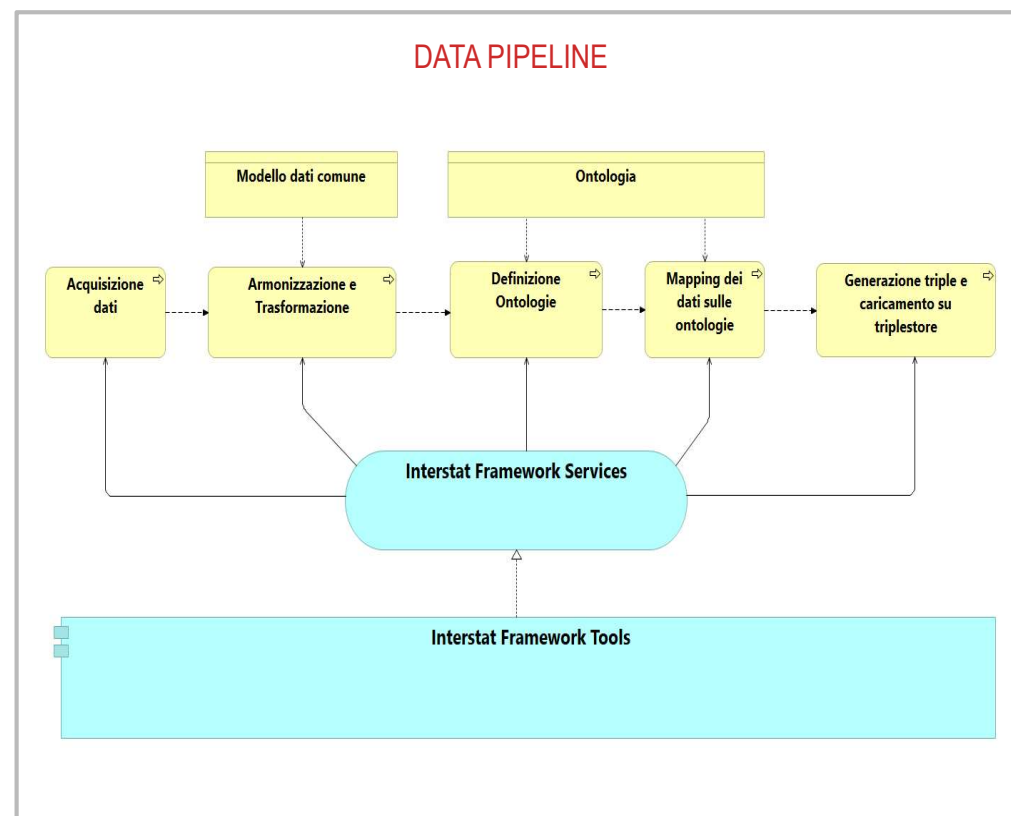
### FONTI STATISTICHE INTEGRATE

	Fonte	Territorio di riferimento
Dati relativi alla qualità dell'aria	ISPRA	Italia
	Agenzia Europea dell'Ambiente	Francia
Censimento della popolazione	ISTAT	Italia
	Insee	Francia

## Focus on: SEP- Support for Environment Policies (2/3)

### Data pipeline:

- Armonizzazione semantica mediante l'utilizzo di ontologie e modelli di dati
- Mapping concettuale tra ontologie e dati da integrare
- Generazione delle triple RDF
- Caricamento delle triple RDF sul triplestore



# Focus on: SEP- Support for Environment Policies (3/3)

## Armonizzazione semantica

- Definizione di una rappresentazione comune dei concetti principali mediante ontologie
- Analisi esplorativa delle ontologie di dominio pubblicate
- Ricerca di modelli dei dati relativi ai domini di interesse per l'armonizzazione delle strutture dati delle fonti selezionate

### ONTOLOGIE E MODELLI DI DATI DEL SEP

Dominio dei dati	Ontologie di Riferimento	Descrizione	Principali concetti
Air pollution	AQD data model	Fornisce alcune linee guida su come i paesi membri dell'European Environmental Agency (EEA) devono comunicare dati grezzi e statistiche aggregate all'EEA.	Pollutant, Aggregation Type, Media Value, Station, Reporting Year, Time Period, Sampling Point, Concentration
	GeoNames	Ontologia che descrive il dominio del territorio in termini di oggetto spaziale, geometria, paese, coordinate geografiche e così via	Location/Geometry, Longitude and latitude
	SOSA (Sensor, Observation, Sample, and Actuator) Semantic Sensor Network Ontology (SSN)	Ontologia che descrive i sensori e le loro osservazioni. È inclusa nell'ontologia SSN.	Sensor, Feature of Interest, madeObservation, hasFeatureofInterest
Census data integrated with Individual Register	Census data model	Struttura dei data cube diffusi	Total persons usually resident in the analysed area, Age class, Sex, Territory, Current activity status, Commuting

# SEP – Servizio 1

## Servizio Cross-domain

Un utente, selezionando un Comune italiano o francese può visualizzare:

- I Comuni con i valori più elevati di **PM10** (Particulate Matter < 10 µm) e la relativa popolazione per genere e classe di età (Municipalities with higher pollution level)
- I Comuni con i valori più elevati di **popolazione**, suddivisa per genere e classe di età e I rispettivi valori di PM10

Home SEP Service 1 SEP Service 2 SEP Service 3

### Resident population in the most polluted areas

The user can select a specific Country, between Italy and France, to be analysed. It is possible to view two different types of information: the first one allows to highlight the Municipalities with higher pollution level, based on the average value measured by all the stations in the selected municipality. This service linking pollution and census data, highlights the value of the resident population in the Municipality grouped by age and gender. In addition, specific information related to the PM10 pollutant considered in the analysis, is reported. The second table allows, instead, to obtain the Municipalities with higher resident population and also the age group and gender whose value is greater, in relation to the pollution data in the selected Municipality. In this service, which represents a specific cross-domain analysis for the selected Country, the areas with the highest or the lowest levels of air pollution, are highlighted. It is also possible to do an analysis on the resident population in those areas.

Select a Country to analyse  
IT

#### Municipalities with higher pollution

NUTS3	NUTS3_CODE	MUNICIPALITY	POPULATION	GENDER	AGE_CLASS
Frosinone	IT145	Ceccano	488	Female	10 to 14 years
Frosinone	IT145	Ceccano	523	Male	10 to 14 years
Frosinone	IT145	Ceccano	4	Female	100 years and over
Frosinone	IT145	Ceccano	1	Male	100 years and over
Frosinone	IT145	Ceccano	500	Female	15 to 19 years
Frosinone	IT145	Ceccano	490	Male	15 to 19 years
Frosinone	IT145	Ceccano	525	Female	20 to 24 years
Frosinone	IT145	Ceccano	572	Male	20 to 24 years
Frosinone	IT145	Ceccano	618	Female	25 to 29 years
Frosinone	IT145	Ceccano	680	Male	25 to 29 years

Rows per page: 10 1-10 of 1121

#### Municipalities with higher population

NUTS3	NUTS3_CODE	MUNICIPALITY	POPULATION	GENDER	AGE_CLASS
Roma	IT143	Roma	128164	Female	50 to 54 years
Roma	IT143	Roma	126275	Female	45 to 49 years
Roma	IT143	Roma	115414	Male	50 to 54 years
Roma	IT143	Roma	115046	Male	45 to 49 years
Roma	IT143	Roma	113004	Female	55 to 59 years
Roma	IT143	Roma	111452	Female	40 to 44 years
Roma	IT143	Roma	103528	Male	40 to 44 years
Roma	IT143	Roma	99562	Male	55 to 59 years
Roma	IT143	Roma	95843	Female	60 to 64 years
Roma	IT143	Roma	91339	Female	35 to 39 years

Rows per page: 10 1-10 of 1121

# SEP – Servizio 2

## Servizio Cross-domain e Cross-border

Un utente, selezionando un **Comune italiano**, un **Comune francese** ed una classe di età può visualizzare e confrontare:

- Il totale della **popolazione residente** relativo alla classe di età e al Comune selezionato
- I livelli di inquinamento atmosferico (**PM10**) nei Comuni selezionati e alcune informazioni relative alla rilevazione dell'inquinante (Station name and ID, Aggregation Type and description, Measure unit, Data Source and reference date)

The screenshot shows the SEP Servizio 2 web application. The interface includes a navigation bar with links to Home, SEP Service 1, SEP Service 2, and SEP Service 3. The main content area is titled "Evaluation of the effects of pollution (considering PM10) on specific municipalities in Italy and France". It features a search panel with dropdown menus for selecting a French NUTS3 Region (Paris), an Italian NUTS3 Region (Roma), a French Municipality (Paris), and an Italian Municipality (Roma). There is also a dropdown for selecting an Age Group (Y10-14). A map of Europe shows the selected regions. Below the map, two statistics are displayed: "Total population in the selected French territory for the Y10-14 Age Group" (388804) and "Total population in the selected Italian territory for the Y10-14 Age Group" (527280). Two data tables are shown: "Air pollution data (PM10) for the IT143 NUTS3 region" and "Air pollution data (PM10) for the FR101 NUTS3 region". Both tables have columns for NUTS\_CODE, MUNICIPALITY, STATION\_NAME, STATION\_ID, AIR\_QUALITY\_VALUE, and POLLUTANT. The first table lists various stations in the Roma area, and the second table lists stations in the Paris area.

# SEP – Servizio 3

## Servizio Cross-border

Un utente, selezionando un **Comune italiano** e un **Comune francese** può visualizzare e confrontare:

- I valori più elevati di due inquinanti atmosferici PM 2.5 (Particulate matter <2.5 µm) e NO2 (Nitrogen dioxide)
- Alcune informazioni relative alle stazioni di monitoraggio presenti nei Comuni selezionati (Station name and ID, Aggregation Type and description, Measure unit, Data Source and reference date)

The screenshot displays the SEP Servizio 3 web application interface. At the top, there are navigation tabs for Home, SEP Servizio 1, SEP Servizio 2, and SEP Servizio 3. The main content area is titled "Evaluation of the effects of pollution (considering PM2.5 and NO2) on specific municipalities in Italy and France". It includes a text box explaining that users can select a specific NUTS region and a specific Municipality within it, and that the application highlights different pollutant registration stations considering PM 2.5 (Particulate matter <2.5 µm) and NO2 (Nitrogen dioxide). It also notes that in the tabular view, it is possible to obtain and compare French and Italian air pollution data and information about the stations in the selected Municipalities, in relation with the two specific pollutants considered.

The interface features a "Geographic areas" section with dropdown menus for selecting a French NUTS Region (Paris), an Italian NUTS Region (Roma), a French Municipality (Paris), and an Italian Municipality (Roma). To the right is a map of Europe with markers for Paris and Roma.

Below the map, there are four data tables showing air pollution data for PM2.5 and NO2 for the IT43 NUTS region and the FR101 NUTS region.

NUTS_CODE	MUNICIPALITY	STATION_NAME	STATION_ID	AIR_QUALITY_VALUE	POLLUTANT
IT43	Roma	CINECITTA	120804	7.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	MALAGROTTA	120888	7.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	CINECITTA	120804	10.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	FERME	120842	11.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	CINECITTA	120804	14.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	MALAGROTTA	120888	14.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	CINECITTA	120804	16.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	FERME	120842	22.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	TENUTA DEL CAVALIERE	120817	16.0	Particulate matter < 2.5 µm (ar)
IT43	Roma	CINECITTA	120804	7.0	Particulate matter < 2.5 µm (ar)

NUTS_CODE	MUNICIPALITY	STATION_NAME	STATION_ID	AIR_QUALITY_VALUE	POLLUTANT
FR101	Paris	gma04a892c01664216823c0745	STA.FR04033	12.126609	Particulate matter < 2.5 µm (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04143	11.679009	Particulate matter < 2.5 µm (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04129	14.790907	Particulate matter < 2.5 µm (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04033	14.982202	Particulate matter < 2.5 µm (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04129	15.514122	Particulate matter < 2.5 µm (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04143	16.303828	Particulate matter < 2.5 µm (ar)

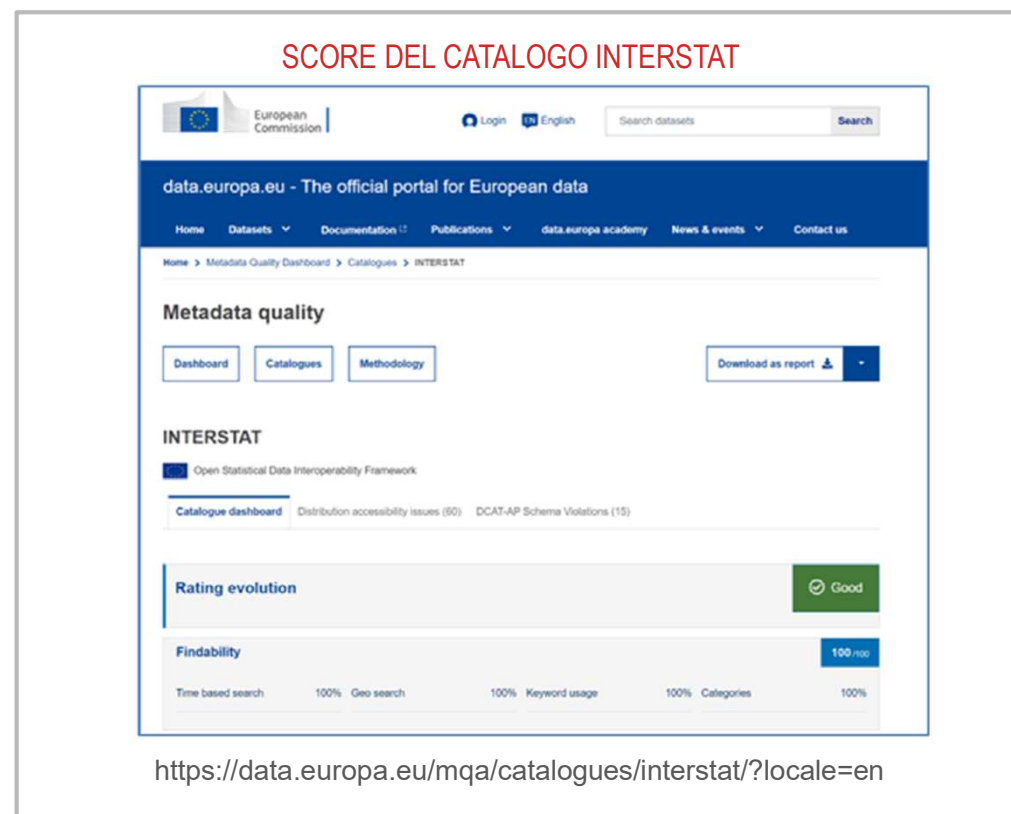
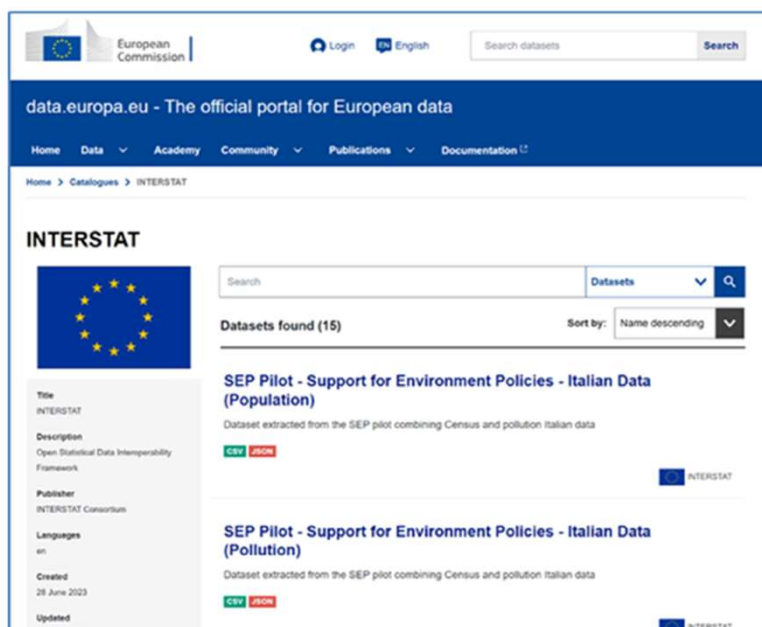
NUTS_CODE	MUNICIPALITY	STATION_NAME	STATION_ID	AIR_QUALITY_VALUE	POLLUTANT
IT43	Roma	MALAGROTTA	120888	2.0	Nitrogen dioxide (ar)
IT43	Roma	TENUTA DEL CAVALIERE	120817	4.0	Nitrogen dioxide (ar)
IT43	Roma	CIPRO	120843	9.0	Nitrogen dioxide (ar)
IT43	Roma	CIPRO	120843	11.0	Nitrogen dioxide (ar)
IT43	Roma	CASTEL DI GUIDO	120803	11.0	Nitrogen dioxide (ar)
IT43	Roma	CIPRO	120843	12.0	Nitrogen dioxide (ar)
IT43	Roma	CIPRO	120843	13.0	Nitrogen dioxide (ar)
IT43	Roma	ARENULA	120847	13.0	Nitrogen dioxide (ar)
IT43	Roma	MALAGROTTA	120888	13.0	Nitrogen dioxide (ar)
IT43	Roma	TENUTA DEL CAVALIERE	120817	14.0	Nitrogen dioxide (ar)

NUTS_CODE	MUNICIPALITY	STATION_NAME	STATION_ID	AIR_QUALITY_VALUE	POLLUTANT
FR101	Paris	gma04a892c01664216823c0745	STA.FR04129	11.942433	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04094	11.948424	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04131	14.848414	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04031	17.808964	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04033	18.117739	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04143	18.946563	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04118	20.789724	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04033	31.838264	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04143	32.382026	Nitrogen dioxide (ar)
FR101	Paris	gma04a892c01664216823c0745	STA.FR04094	34.109563	Nitrogen dioxide (ar)



# INTERSTAT e l'European data portal (1/2)

I dataset relativi ai pilot implementati sono stati pubblicati sul portale data.europa.eu e i rispettivi metadati verificati mediante la metodologia del Metadata Quality Assurance (MQA)



# INTERSTAT e l'European data portal (2/2)

## Compliance con MQA:

- Le dimensioni esaminate dall'MQA derivano dai FAIR principles (Findability, Accessibility, Interoperability, Reusability)
- Per ogni dimensione, un set di indicatori permette di valutare la qualità dei metadati dei Cataloghi pubblicati
- Verifica della conformità dei metadati alla specifica DCAT-AP tra le metriche usate per valutare l'interoperabilità di una distribuzione



# Conclusioni

---

- Tools e standards sono fondamentali per il riuso e l'integrazione dei dati open, ma devono essere integrati da descrizioni semantiche
- La disponibilità di ontologie di dominio, meta-ontologie e modelli di dati comuni ha facilitato lo sviluppo delle applicazioni pilota
- L'accesso a dati e metadati ben documentati ha influenzato la selezione delle fonti di dati di input durante la progettazione dei casi d'uso
- Le attività di sviluppo del framework INTERSTAT e dei servizi pilota hanno richiesto l'impiego di diverse competenze, soprattutto tecniche e di conoscenza dei domini

# Riferimenti

---

- *INTERSTAT Deliverable 1.1*: Report on the maturity of (Linked) Open Statistical Data in Europe (AA.VV., 2021)
- *INTERSTAT Deliverable 1.2*: Report on the overall technical scenarios in Europe in relation to the provisioning and reuse of (Linked) Open Statistical Data (AA.VV., 2021)
- *INTERSTAT Deliverable 2.1*: Ontologies and tools to enable cross-border semantic interoperability (AA.VV., 2021)
- *INTERSTAT Deliverable 3.1*: Report of the use cases to demonstrate the cross-border benefits of the proposed solution (AA.VV., 2020)
- *INTERSTAT Deliverable 3.2*: Pilot services deployed and working in real environments (AA.VV., 2022)
- *INTERSTAT Deliverable 3.3*: Assessing the impact of INTERSTAT project: methodology and KPIs (AA.VV., 2023)
- *INTERSTAT Deliverable 4.2: Exploitation & Sustainability plan* (AA.VV., 2023)
- *INTERSTAT*, Technical framework: <https://framework.cef-interstat.eu/> (2021)
- *INTERSTAT*, SEP Client Application: <https://framework.cef-interstat.eu/SEP-pilot-client/> (2022)
- *INTERSTAT*, GF Client Application: <https://framework.cef-interstat.eu/GF-pilot-client/> (2022)
- *INTERSTAT*, S4Y Client Application: <https://framework.cef-interstat.eu/S4Y-pilot-client/> (2022)

# grazie

GIUSEPPINA RUOCCO | [giuseppina.ruocco@istat.it](mailto:giuseppina.ruocco@istat.it)