



Il progetto Sistan Hub per la diffusione delle statistiche ufficiali al servizio dei processi decisionali sul territorio

Lia Coniglio

Cremona, 22 ottobre 2015

Indice

1. Sistan Hub

- Definizione
- Caratteristiche
- Progetto

2. Annuario statistico Regionale della Lombardia

3. Prototipo

4. Conclusioni



Data Warehouse virtuale del Sistema statistico nazionale per la diffusione di macrodati statistici ma anche strumento strategico e operativo per la cura della qualità, della condivisione dei metadati e dell'integrazione dei dati *(Malizia, Convegno nazionale Usci 2015)*



Cos'è?

- Un network per la diffusione dei dati

Modello di riferimento?

- Census HUB project (Eurostat)

Obiettivo?

- Riprodurre dati statistici pluritematici con riferimenti territoriali
- Migliorare la condivisione e la diffusione di dati e metadati
- Garantire l'ottica territoriale all'utenza locale



Come è organizzato?

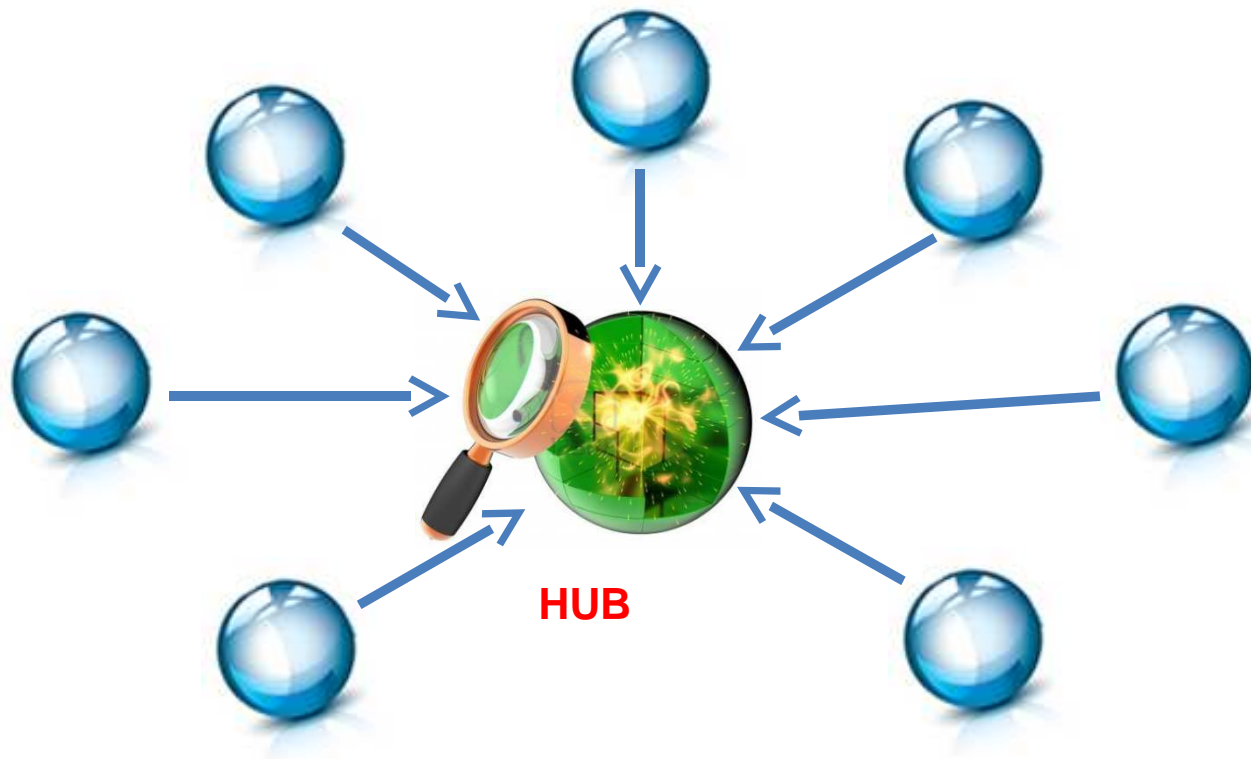
- Singoli enti producono, organizzano e aggiornano i propri dati
- Dati messi a disposizione in singoli nodi locali
- Dati armonizzati e corredati dai metadati
- Utilizzo strumento SDMX (Statistical Data and Metadata Xchange)



Come è organizzato?

- **Hub** porta di accesso via web (entrypointweb) ai dati disponibili sui nodi locali
- Tramite Hub sono interrogati i nodi locali
- I dati immessi nei nodi saranno rintracciati attraverso l'Hub
- Output dati (tabelle, grafici, mappe) e metadati

NODI LOCALI

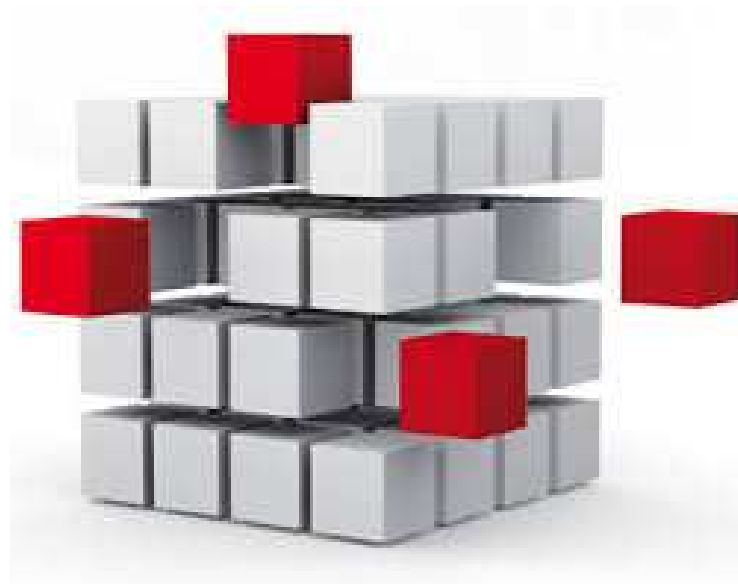


HUB

Caratteristiche

Punti di forza:

- Facilità di interrogazione
- Confronti
- Armonizzazione
- Dettaglio territoriale



Punti di debolezza:

- Disponibilità database locali
- Conversione dei dati degli enti in database relazionali

Attuazione del progetto

Sperimentazione con:

- Éupolis Lombardia
- Unioncamere Lombardia
- Regione Sicilia
- Comune di Vicenza



Attualmente:

- Definizione della piattaforma informatica
- Individuazione dei primi contenuti per il popolamento nodi
- Trattamento degli ipercubi
- Creazione dei nodi

Attuazione del progetto

Progetto innovativo che necessita di una forte governance da parte dell'Istat per:

- Coordinamento interno Istat e con i partner
- Aspetti tecnici e contenutistici
- Integrazione delle fonti
- Qualità dei dati
- Modalità di diffusione
- Tecnologia informatica
- Gestione e alimentazione dei nodi
- Armonizzazione metadati



Presentazione del prodotto in occasione della Conferenza Nazionale di Statistica del 2016

Esempi di diffusione dati in Lombardia

L'Annuario Statistico Regionale Lombardia

<http://www.asr-lombardia.it/ASR/>



Diffusione articolata su cinque prodotti

- Expandi tutto
- Espandi le categorie
- Datawarehouse

Tem

- Popolazione e famiglie
- Condizioni economiche delle famiglie e disuguaglianze
- Povertà
- Spesa media mensile familiare
- Salute e sanità
- Imprese
- Servizi
- Piezzi
- Lavoro
 - Lavoratori
 - Lavoratori autonomi
 - Lavoratori autonomi dati mensili - Commercianti
 - Lavoratori autonomi dati mensili - Artigiani
 - Lavoratori dipendenti
- Retribuzioni
- Pensioni
- Territorio
- Enti

SDMX Istat Framework v0.9

Questo è un sito di test!!!!

CHE COSA È L'SDMX ISTAT FRAMEWORK

È un insieme di tools sviluppati nell'ambito dei grant "Horizontal and vertical integration: Implementing technical and statistical standards in the European Statistical System" finanziato da Eurostat. Il Framework, compatibile con SDMX Reference Infrastructure sviluppata da Eurostat, ne estende le funzionalità permettendo di gestire l'intero processo di diffusione (creazione del database e/o interfacciamento a database esistenti (es. I.Stat), modellazione dei dataset, presentazione attraverso una Web GUI) in modalità "metadata and SDMX driven".

CHE COSA È L'SDMX WEB BROWSER

È un tool - fa parte dell'SDMX Istat Framework - che permette ad un utente di navigare dati provenienti da database distribuiti, e presentare i risultati attraverso tabelle e grafici. Può essere, quindi, utilizzato all'interno di una singola Organizzazione per diffondere dati memorizzati su uno o più database, oppure nel contesto di un progetto "multitente" in cui più organizzazioni rendono disponibili dati memorizzati nei propri database di diffusione. L'interconnessione tra l'SDMX Web Browser e i database di diffusione avviene utilizzando web service in SDMX, per cui anche dati diffusi da Organizzazioni Internazionali, quali Eurostat, OECD, IMF e ECB possono essere ricercati e presentati.

FUNZIONALITÀ DELL'SDMX WEB BROWSER

- Il software è composto da due moduli principali:
- gestione e amministrazione (utilizzato da un utente con le credenziali di "administratore")
 - ricerca (navigazione), presentazione e visualizzazione dei dati provenienti da database distribuiti (utilizzato dai "data users")

L'amministratore imposta le caratteristiche generali quallo sito, gli oggetti da visualizzare (es. l'albero dei temi, lo "switch" tra differenti web service, le dashboard, ecc.), i database distribuiti a cui cometersi. Inoltre può definire gli oggetti "informativi" utilizzati dai "data users" per compiere le ricerche e visualizzare i risultati. Infine l'amministratore può creare dashboard (questa pagina è stata creata attraverso una dashboard), componendole attraverso dei widget collegati a database comunque distribuiti: per esempio potrà creare una dashboard composta da un grafico i cui dati provengono da I.Stat, da una tabella con dati provenienti dal database di Eurostat, ecc.

Per ciascun dataset l'amministratore potrà facilmente creare un template preimpostato definendo dei filtri (query) e decidendo la disposizione degli assi della tabella, in maniera tale da assegnare un default di rappresentazione.

Il "data user" può:

- visualizzare i dashboard disponibili
- "switchare" tra differenti web service
- navigare uno o più alberi dei temi e selezionare i dataset d'interesse (è possibile avere in uno specifico tema dei dataset residenti in database distribuiti)
- inserire dei filtri sui dataset selezionato
- definire l'orientamento di presentazione del dataset selezionato
- calcolare on-the-fly la validazione congiunturale e tendenziale sul dataset selezionato
- creare grafici su dataset selezionato, aggiungendo anche la visualizzazione delle variazioni congiunturale e tendenziale, e personalizzando anche le label
- se autenticato, memorizzare le ricerche (query) in maniera tale da poterle riutilizzare in successive sessioni

- Expandi tutto
- Expandi le categorie
- Datawarehouse
- Temi
- Territorio
 - Lombardia
 - Imprese
 - Salute e sanità
 - Prezzi
 - Popolazione e famiglie
 - Indicatori demografici - Regione Lombardia
 - Veneto
 - Sicilia

Indicatori demografici - Regione Lombardia

Criteri | Assi | Valore | Selezione download | Table | Chart mode | Metadata

Periodicità	[A] annuale											
Area geografica di riferimento	Italia			Nord-ovest			Lombardia			Cremona		
Tipo di indicatore demografico	tasso di natalità (per mille abitanti)	tasso di mortalità (per mille abitanti)	tasso di nuzialità (per mille abitanti)	tasso di natalità (per mille abitanti)	tasso di mortalità (per mille abitanti)	tasso di nuzialità (per mille abitanti)	tasso di natalità (per mille abitanti)	tasso di mortalità (per mille abitanti)	tasso di nuzialità (per mille abitanti)	tasso di natalità (per mille abitanti)	tasso di mortalità (per mille abitanti)	tasso di nuzialità (per mille abitanti)
Periodo di tempo												
2002	9.4	9.8	4.7	9	10.5	4.3	9.6	9.4	4.4	8.4	11.3	4
2003	9.5	10.2	4.6	9.1	11	4.1	9.6	9.9	4.1	8.2	11.8	4
2004	9.8	9.5	4.3	9.5	9.9	3.8	10.1	9	3.8	8.6	10.5	3.7
2005	9.6	9.8	4.3	9.3	10.2	3.8	9.9	9.2	3.7	8.9	10.8	3.7
2006	9.6	9.6	4.2	9.5	10.1	3.8	10.2	9.1	3.7	8.8	10.6	3.7
2007	9.7	9.8	4.3	9.6	10.1	3.8	10.2	9.1	3.8	9.2	10.7	3.9
2008	9.8	9.9	4.2	9.8	10.4	3.7	10.4	9.4	3.6	9.7	10.8	3.6
2009	9.6	10	3.9	9.7	10.4	3.4	10.3	9.4	3.3	9.5	11.3	3.3
2010	9.5	9.9	3.7	9.5	10.3	3.2	10.2	9.4	3.1	9.3	11.2	3.2
2011	9.2	10	3.4	9.2	10.2	3	9.7	9.2	2.8	8.7	10.8	2.8
2012	9	10.3	3.5	9	10.6	3	9.4	9.6	2.9	8.7	11.5	2.9
2013	8.5	10	3.2	8.5	10.3	2.8	8.9	9.2	2.7	8	10.8	2.7
2014	8.3	9.8	3.1	8.2	10	2.7	8.6	9.1	2.6	8.1	10.7	-

Expandi tutto Espandi le categorie

Datawarehouse

- Temi
 - Popolazione e famiglie
 - Condizioni economiche delle famiglie e disuguaglianze
 - Poverta
 - Incidenza di povertà relativa familiare per regione
 - Incidenza di povertà relativa familiare
 - Spesa media mensile familiare
 - Salute e sanità
 - Imprese
 - Servizi
 - Prezzi
 - Lavoro
- Territorio
- Enti

Incidenza di povertà relativa familiare

▼ Criteri ▲ Assi Valore ▼ Selezione download ▼ Table Chart mode Metadata

Criteri - 2806 / 10000

FREQ (1/1)	REF_AREA (0/26)	IND_TYPE	MEASURE	NUM_COMPONENTI	TIPOLOGIA_FAM	FAM_MINORI
FAM_ANZIANI	CLASSE_ETA_PERS_RIF	TITOLO_STUDIO_PERS_RIF	COND_PROF_PERS_RIF	TIME_PERIOD		

Territorio

[IT] Italia

- [ITCD] Nord
- [ITE] Centro
 - [ITE1] Toscana
 - [ITE2] Umbria
 - [ITE3] Marche
 - [ITE4] Lazio
 - [ITFG] Mezzogiorno
- [ITC1] Piemonte
- [ITC2] Valle d'Aosta / Vallée d'Aoste
- [ITC3] Liguria
- [ITC4] Lombardia

Ok Annulla



- Expandi tutto | Espandi le categorie
- Datawarehouse
- Tem
 - Popolazione e famiglie
 - Condizioni economiche delle famiglie e disuguaglianze
 - Povertà
 - Spesa media mensile familiare
 - Salute e sanità
 - Imprese
 - Servizi
 - Prezzi
 - Lavoro
 - Lavoratori
 - Lavoratori autonomi
 - Lavoratori autonomi dati mensili - Commercianti
 - Lavoratori autonomi dati mensili - Artigiani
 - Lavoratori dipendenti
 - Retribuzioni
 - Pensioni
- Territorio
- Enti

Lavoratori dipendenti nel mese

Table | Chart mode | **Metadata**

QUAL - Qualifica professionale

TIME_PERIOD - Time Dimension Start and End periods

Option

Valori da visualizzare

- Valore
- Tendenziale
- Congiuntura

Tipo grafico

Spline

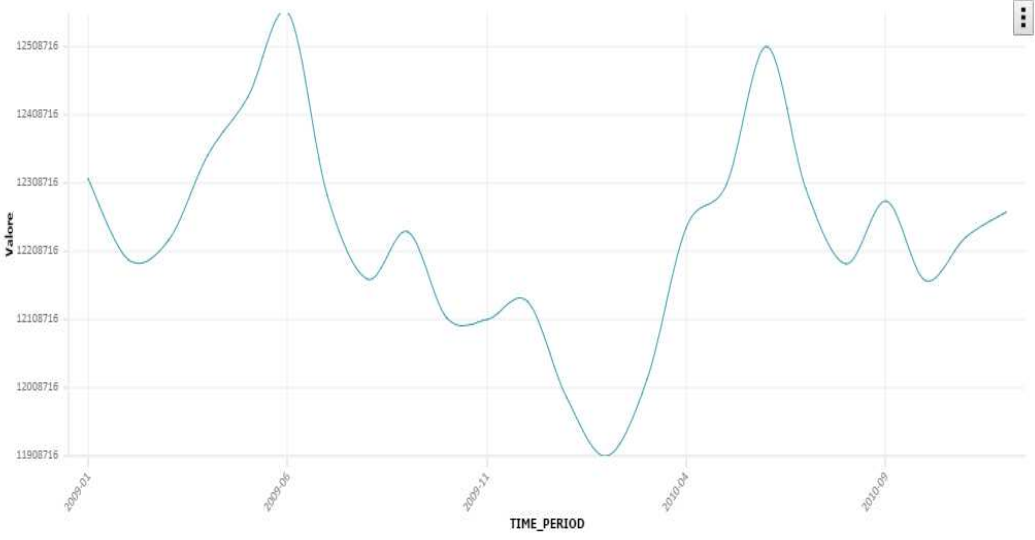
Dimensione asse X

Periodo

Personalizza etichette

Applica

Totale, Italia, mensile, Numero lavoratori



Conclusioni

Sistan Hub

8mila Census

UrBes

Bes- Province

Nel futuro non solo progetti riguardanti l'uso e l'acquisizione di microdati ma anche sviluppo di sistemi di diffusione del dato per realizzare infrastrutture di riferimento per tutti i soggetti locali a tutti i livelli

