

Sistan hub: il Data warehouse virtuale del Sistema statistico nazionale è ormai una realtà

Dopo tre anni di lavoro, l'Istat e un nucleo di soggetti del Sistan a livello territoriale hanno messo a punto il primo prototipo di "data warehouse virtuale" della statistica ufficiale

14. 04. 2016



L'idea progettuale di realizzare un sistema per la condivisione, l'integrazione e la diffusione dei dati statistici e dei relativi metadati prodotti dagli enti del Sistema statistico nazionale, presentata tre anni fa all'XI Conferenza nazionale di statistica, oggi è realtà.

Questa importante realizzazione, che si chiama *Sistan Hub*, annovera tra i suoi protagonisti non solo l'Istat ma anche alcuni soggetti territoriali del Sistan di particolare rilevanza. Si tratta di Regione Lombardia, Regione Siciliana, Unioncamere Lombardia e Comune di Vicenza, che hanno lavorato insieme all'Istituto nazionale di statistica sin dalla nascita del progetto, sviluppandolo poi secondo le linee programmatiche definite da una convenzione siglata nel 2014.

Sin dall'inizio di questo lungo cammino progettuale, l'ambiziosa idea sottesa a *Sistan Hub* è stata quella di costruire un "data warehouse virtuale" della statistica ufficiale, cioè un unico punto di accesso per la navigazione delle informazioni prodotte da tutti gli enti Sistan, così da garantire la progressiva armonizzazione dei dati, la loro comparabilità e integrazione, nonché il massimo di tempestività e facilità di reperimento, secondo i principi del Code of practice europeo e del Codice italiano delle statistiche ufficiali. Tra gli obiettivi del progetto, inoltre, è altrettanto importante il rafforzamento della cooperazione istituzionale tra i diversi enti produttori.

La logica di funzionamento di *Sistan Hub* prevede che le statistiche continuino a risiedere nei data base degli enti del Sistema, venendo però governate e armonizzate attraverso lo standard internazionale SDMX, sviluppato e adottato da tutti i più importanti organismi internazionali per lo scambio di dati e metadati. Il sistema è compiutamente *metadata driven* e si colloca nella fascia alta della filosofia *Open Data* nella linea dell'interoperabilità semantica di cui all'Agenda digitale. Le informazioni residenti nei singoli nodi di *Sistan Hub*, corrispondenti ai data base degli enti coinvolti, sono interconnesse fra loro, con possibilità di interrogazione in tempo reale da parte dell'utente. Si tratta, in definitiva, di una rete di basi di dati armonizzati, composta da nodi raggiungibili automaticamente dagli utenti in pochi click. Tale rete, inoltre, è scalabile, senza alcun impatto sulle performance del sistema. In questa prospettiva, può essere composta da due, cento o migliaia di nodi situati nel Paese, in Europa, nel mondo.

Il prototipo di *Sistan Hub* messo a punto dopo tre anni di lavoro è pienamente funzionante, assicurando elevatissime performance di navigazione e accesso alle informazioni. Parallelamente a questa realizzazione, prosegue l'attività di modellazione e mappatura in SDMX delle statistiche detenute nei data base degli enti partecipanti al *network*. Appena la partecipazione al sistema avrà raggiunto un livello sufficientemente ampio, esso verrà ufficialmente esposto su web per il libero accesso da parte dell'utenza.

L'Istat ha svolto e svolgerà in futuro la funzione di coordinamento del progetto, sovrintendendo alla cura del sistema di diffusione dei dati e metadati; verificando la rispondenza dei dati e dei metadati ai requisiti di qualità propri della statistica ufficiale, tali quindi da poter essere esposti nel sistema; assistendo e formando i partner per la realizzazione del singolo nodo.

In questa fase di sviluppo del progetto, prosegue anche l'attività di selezione dei dataset, a partire dai piani di spoglio presenti nel datawarehouse I.stat, che sono stati "ottimizzati" mediante le operazioni di modellazione e mappatura con lo standard SDMX-RI. Quanto ai nodi locali, la cui gestione e sviluppo è affidata alla responsabilità dei singoli enti partecipanti al progetto, con il supporto tecnico-organizzativo dell'Istat, sono stati predisposti gli ambienti per la successiva configurazione. Inoltre, nell'ambito dei nuclei locali – gruppi di lavoro costituiti dall'Ufficio territoriale Istat competente e dai partner coinvolti nella sperimentazione – sono stati definiti i primi dataset da pubblicare. Più di 200 ipercubi residenti in diversi nodi sono già oggi interrogabili. In definitiva, l'architettura di *Sistan Hub* è tale da consentire in prospettiva a ogni ente Sistan di pubblicare sul proprio nodo i dati di interesse, agganciandoli a un tema/sottotema standard. A questo fine, sono state realizzate diverse giornate di formazione su SDMX-RI, rivolte sia al personale dell'Istat sia ai rappresentanti degli enti partecipanti al progetto.

Nel corso del 2015 il progetto è stato presentato in occasione del ForumPA, svoltosi a Roma nel mese di maggio. Inoltre, è stato oggetto di incontri bilaterali e seminari con altri importanti soggetti produttori di informazioni statistiche di rilievo nazionale, fra cui molte Autorità statistiche nazionali, che sono componenti del Sistema statistico europeo. Il 26 gennaio di quest'anno, infine, è stato presentato in Comstat, che ne ha raccomandato una rapida estensione agli enti più importanti. Sono numerose le istituzioni coinvolte già oggi nelle attività di formazione, impianto e primo popolamento di *Sistan Hub*, secondo un approccio di graduale implementazione che, per alcuni enti, ha anche natura sperimentale. Precisamente, si tratta di: Inps, Ministero dell'Economia e delle finanze, Ministero della Salute, Ministero delle Infrastrutture e dei trasporti, Ministero dell'Interno, Ministero della Giustizia, Presidenza del Consiglio dei ministri, Dipartimento della Funzione pubblica, Dipartimento della Programmazione economica, Agenzia per la coesione territoriale e Ispra. Partecipa anche la Banca d'Italia, in virtù della specificità e dell'alta qualità delle informazioni prodotte, che ne rendono possibile l'inclusione in un concetto allargato di statistica ufficiale.

Con la realizzazione di *Sistan Hub*, il Sistan ha dimostrato concretamente di saper essere all'altezza del proprio ruolo al servizio della comunità nazionale e internazionale, offrendo un prodotto che collocherà il nostro Paese e la statistica ufficiale italiana alla frontiera dell'innovazione a livello mondiale.

Alla base di questi esiti, che potranno essere oggetto di un'ampia condivisione durante [la XII Conferenza nazionale di statistica](#), ci sono vari fattori, come le competenze tecniche e le capacità di innovazione tecnologica, metodologica e organizzativa. L'elemento decisivo, tuttavia, è rappresentato dalla reale cooperazione interistituzionale che gli enti coinvolti hanno saputo mettere in campo per conseguire un obiettivo strategico comune. Questa logica, che ha caratterizzato ogni azione progettuale attivata in questi anni nel Sistema, ha consentito ancora una volta alla rete Sistan di potere e saper essere viva e vitale, anche a normativa data, generando uno straordinario valore aggiunto per gli enti che compongono il Sistema e per l'intero Paese.

Raffaele Malizia