

# IL CORPORATE DATA WAREHOUSE: UN PERCORSO DI PIANIFICAZIONE E SVILUPPO DI SERVIZI ALL'UTENZA

di Stefania Bergamasco (bergamas@istat.it)

■ Era l'ottobre del 2009 quando in Istat è nata l'idea di costituire un *data warehouse* unico per mettere a disposizione i dati prodotti dall'Istituto. Da più parti giungeva la richiesta di migliorare i livelli di servizio offerti nell'ambito della diffusione di informazione statistica, uscendo dalla logica di organizzazione della stessa per fonte di produzione, e di realizzare un ambiente unico, organizzato per temi, in cui cittadini, amministratori locali, utenti specializzati e politici potessero ritrovare agevolmente tutti i dati di interesse. Così, dopo una fase di studio, a dicembre 2009 si è deciso di iniziare questa nuova avventura adottando la tecnologia .Stat realizzata in ambito OECD, che consentiva di offrire importanti servizi: la possibilità di navigare i dati selezionando le variabili di diffusione e gli incroci di proprio interesse, posizionandoli in fiancata o in testata in base alle specifiche esigenze; di registrarsi e di effettuare il salvataggio delle proprie tavole costituendo, per ciascuno dei nostri utenti, una sorta di banca dati personalizzata sempre aggiornata; di visualizzare congiuntamente dati e metadati; di esportare i dati in diversi formati; di realizzare ricerche. Nel gennaio 2010 tutto l'Istituto ha avviato questa imponente trasformazione.

È stato subito chiaro però che l'Istat avrebbe dovuto ripensare completamente: i processi organizzativi per garantire la sostenibilità dei servizi messi a disposizione; le regole con le quali definire e governare le classificazioni in modo che esse fossero integrate tra più settori di produzione e che consentissero, lì dove possibile, la comparabilità



e integrabilità dei dati diffusi; le regole di governo dei comunicati stampa cosicché il sistema fosse a supporto integrato non solo del sistema di diffusione ma anche di quello di comunicazione. Sarebbe stato necessario migrare, all'interno del nuovo *data warehouse*, le tavole dati e le banche già presenti sul sito per realizzare una sorta di ricostruzione storica dei livelli di servizio. Inoltre è stata subito chiara anche la necessità di decidere secondo quale logica realizzare il processo di messa in produzione prima e poi di popolamento continuo. Nell'epoca del web 2.0, l'Istat ha scelto di intraprendere un cammino che fosse il risultato di uno sforzo comune. Quindi, sin da subito, sono state coinvolte tutte le strutture di produzione e si è deciso di affrontare la messa in esercizio del sistema attraverso step: a maggio 2010 il nuovo *data warehouse* I.Stat è stato pubblicato nella Intranet per una valutazione interna; a luglio I.Stat, mediante l'utilizzo di utenza e password, è stato reso accessibile a una lista di giornalisti, politici, esperti del settore e cittadini di fornire suggerimenti e commenti; a dicembre dello stesso anno I.Stat è stato rilasciato a tutta l'utenza (all'indirizzo <http://dati.istat.it>).

## CUBI DATI, DIMENSIONI, VARIABILI E METADATI

Ad oggi su I.Stat sono disponibili 256 "cubi dati" per un totale di 1.556 variabili di diffusione, delle quali 922

afferiscono all'ambito economico e 634 a quello sociale, accompagnate da circa 2.979 note esplicative; la lunghezza media delle serie storiche è di 10 anni e mediamente le variabili di diffusione vengono classificate con 6 dimensioni sino a picchi di 14 dimensioni; per ciascun cubo vi è una media di 6 "tavole salvate" dagli utenti registrati con picchi anche di 60. In media 17.000 utenti consultano mensilmente il sistema navigando su circa 145.739 pagine. Alcune banche dati, anche se ancora disponibili su web, sono state già migrate, come quelle dei dati congiunturali di Conlstat, quella dei prezzi al consumo, e quella delle retribuzioni contrattuali.

## SERVIZI E COLLABORAZIONI TRA GLI OBIETTIVI FUTURI

Ma il percorso non è finito. L'Istat sta creando nuovi servizi da mettere a disposizione dell'utenza, quali il *Single Exit Point*, che consente di esportare i dati contenuti nella banca dati di I.Stat, per un colloquio *machine-to-machine* secondo standard SDMX; sta costituendo collaborazioni in ambito Sistan tra differenti enti del Sistema per la realizzazione, con la medesima tecnologia, di sistemi Sistan integrati e tematici di diffusione dati. Il percorso inizialmente interno all'Istituto ora vuole arricchirsi dell'esperienza e della collaborazione di tutto il Sistan al fine di fornire servizi congiunti e trasversali agli utenti superando non solo la logica di diffusione in base alla fonte dati produttrice, ma anche di ente produttore del dato: un sistema che, pur nelle autonomie e responsabilità dei singoli, trovi unitarietà di fronte ai propri utenti e che seppur dotato di tante anime trovi canali congiunti di offerta di servizi.