

# ESPERTI A CONFRONTO PER ORIENTARE LE SCELTE DELL'ISTAT SUI BIG DATA

di Monica Scannapieco (scannapi@istat.it)

■ L'HLG (High-level group for the modernisation of statistical production and services) - gruppo internazionale costituito da presidenti di organizzazioni statistiche con il compito di guidare le strategie di modernizzazione della produzione statistica - ha rilevato l'importanza del fenomeno dei Big Data e la necessità per le organizzazioni statistiche di rivedere le loro funzioni alla luce del fenomeno stesso.

Le modalità e i tempi di introduzione dei Big Data nelle attività di produzione della statistica ufficiale sono, tuttavia, ancora da definire.

A tal fine è stata costituita una Commissione di studio sui Big Data, presieduta dall'Istat, con l'obiettivo generale di "orientare le scelte dell'Istituto sul tema dei Big Data". La Commissione, costituita lo scorso 14 febbraio 2013, ha durata biennale.

## GLI OBIETTIVI DELLA COMMISSIONE

La Commissione ha due obiettivi principali. Da una parte, si propone di definire "cosa" utilizzare, in termini di dati, metodi e tecnologie afferenti al mondo Big Data, nel processo di produzione dell'informazione statistica attuale. Rientra in tale obiettivo, la valutazione dell'utilizzo delle sorgenti di Big Data come sorgenti ausiliarie che integrano le sorgenti tradizionali. Dall'altra parte, vuole definire "come", eventualmente, cambiare il processo di produzione dell'informazione statistica. A tal fine, le sorgenti di Big Data saranno valutate sia come sorgenti aggiuntive a quelle tradizionali per studiare fenomeni allo stato non considerati nell'ambito della produzione statistica ufficiale, sia come sorgenti sostitutive a quelle tradizionali, lad-

dove ne sia provata una maggiore efficacia dal punto di vista dei costi senza una sostanziale perdita di qualità dell'informazione prodotta.

## I COMPONENTI

I membri della Commissione sono sia interni all'Istat sia esterni. La componente interna della Commissione prevede direttori di dipartimento dell'Istituto e alcuni direttori centrali di strutture particolarmente coinvolte sul tema.

La componente esterna della Commissione prevede tre principali "anime": accademia, settore privato, comunicazione. I membri accademici, con competenze sia statistiche che informatiche, provengono dalle maggiori università italiane; sono, ad esempio membri della commissione Ugo Trivellato, professore della Facoltà di Scienze Statistiche dell'Università di Padova e Dino Pedreschi, professore dell'Università di Pisa, esperto di tecniche di Big Data mining. I componenti del settore privato includono grandi aziende quali Google e Cisco. Il settore della comunicazione è rappresentato da Luca De Biase, giornalista di importanti testate tra cui "Il Sole 24 Ore".

## LO STATO DEI LAVORI

La Commissione ha organizzato i lavori individuando tre sottogruppi:

- metodologia statistica;
- metodologia e tecnologia informatica;
- esperienze e sperimentazioni.

I tre sottogruppi sono stati definiti identificando da un lato la componente tecnica (sottogruppi Metodologia statistica e Metodologia e tecnologia informatica), dall'altro dando rilievo alle esperienze concrete (sottogruppo

Esperienze e sperimentazioni), al fine di avere un quadro completo delle problematiche dell'area Big Data. Si sottolinea come la componente tecnica della commissione con competenze relative ad aspetti di comunicazione e sociologici è stata coinvolta in tutti i sottogruppi proprio in virtù del carattere trasversale di tali aspetti.

Ciascun sottogruppo ha come primo obiettivo concreto la produzione di un *issue paper* in cui saranno presentate le problematiche che riguardano l'introduzione dei Big Data nella Statistica Ufficiale secondo quanto stabilito negli obiettivi della commissione.

Gli *issue papers* sono in corso di redazione e si prevede di averne una prima versione entro la fine di giugno 2013. Riporteranno una descrizione delle aree dei Big data di interesse per la statistica ufficiale, sia di natura statistica che di natura informatica, con un'ottica di rassegna delle principali problematiche per ciascuna area. I lavori della Commissione procederanno con l'individuazione di temi di particolare interesse da approfondire e con la proposta di un insieme di raccomandazioni per orientare concretamente il percorso di valutazione e progressiva introduzione dei Big Data nelle realtà produttive dell'Istat.

## I MEMBRI DELLA COMMISSIONE

### Esterni

Francesco Tortorelli, Agenzia per l'Italia Digitale (AgID) | Flavia Marzano, Presidente Stati Generali dell'Innovazione | Agostino Santoni, CISCO | Michele Festuccia, CISCO | Luca De Biase, giornalista, Direttore scientifico Digital Academia | Ugo Trivellato, Università di Padova | Bruno Scarpa, Università di Padova | Roberta Siciliano, Università di Napoli | Dino Pedreschi, Università di Pisa | Alessandro Panconesi, Università di Roma "La Sapienza" | Luigi Cannari, Banca d'Italia | Maurizio Lenzerini, Università di Roma "La Sapienza" | Diego Ciulli, Google | Fausto Giunchiglia, Università di Trento

### Interni

Emanuele Baldacci, Roberto Monducci, Andrea Mancini, Saverio Gazzelloni, Domenico Donvito, Patrizia Cacioli, Tommaso Di Fonzo

### Segreteria tecnica

Monica Scannapieco, Marina Rezzesi