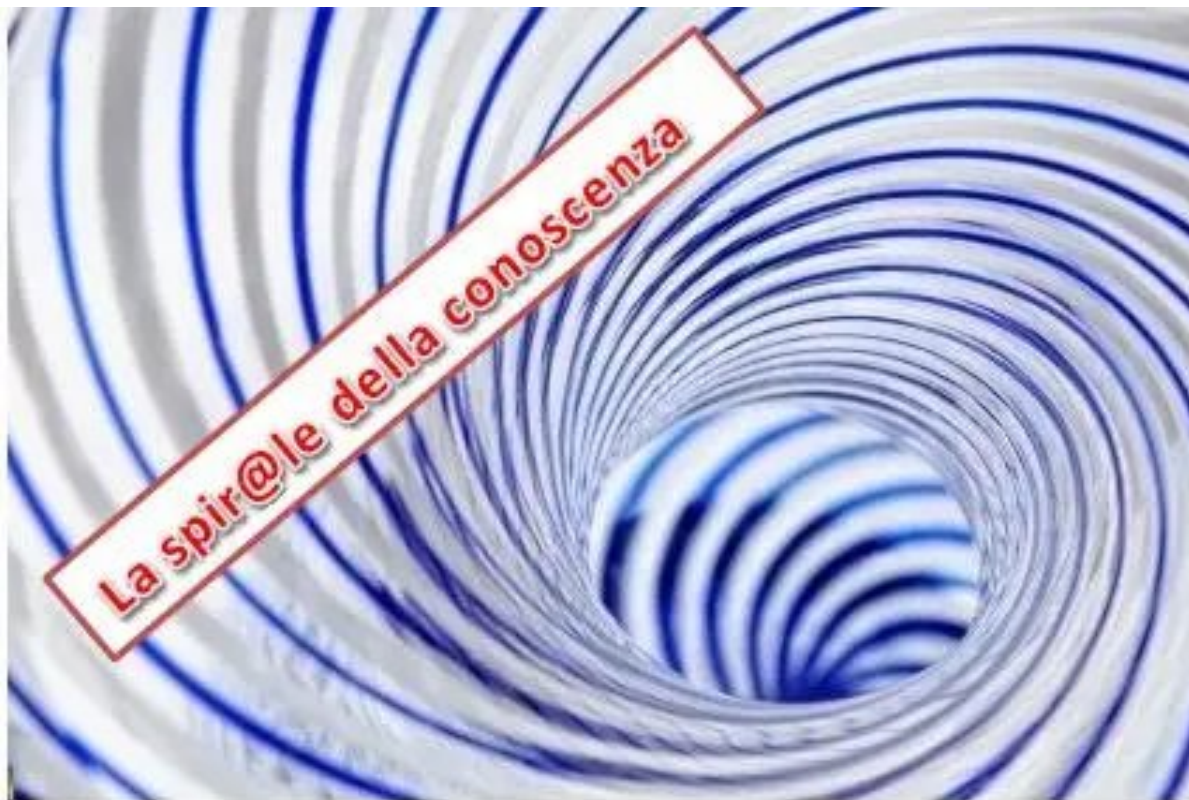


Big Data vs Open Data

Data: Invalid Date | Autore: Rosangela Muscetta



ROMA, 29 MAGGIO 2013 – Le due definizioni di Open Data e Big Data si accompagnano ormai sempre più spesso facendo riferimento alla necessità e possibilità di accedere a questi contenuti informativi per poter operare in termini di confronto, integrazione, analisi di grandi masse di informazioni riguardanti dati statistici, geografici, economici e demografici. [MORE]

Nello specifico, con l'espressione Open Data, o dati aperti, si fa riferimento alle informazioni raccolte dalle PA e messe a disposizione dei cittadini e delle imprese, attraverso lo strumento web, sotto forma di database interoperabili e con licenze libere. Nell'ambito della PA, infatti, da sempre vengono raccolte informazioni sulle tematiche più svariate e oggi, tramite queste collezioni di dati, si ha la possibilità di restituire il patrimonio informativo accumulato ai cittadini e agli operatori economici di settore. Le principali caratteristiche degli Open Data DataBase sono: trasparenza, partecipazione e collaborazione, che mirano a modificare i rapporti di condivisione di conoscenza tra PA e cittadini, imprese ed enti pubblici.

I Big Data (BD), invece, indicano un insieme di dati complessi ("big"), sia da un punto di vista quantitativo, sia dal punto di vista qualitativo, relativamente alla loro laboriosità: distribuzione su scala mondiale, eterogeneità delle fonti, difficile accesso alle informazioni intrinseche dei dati stessi. Da qui nasce l'esigenza di organizzare in maniera pianificata i dati, al fine di renderli fruibili senza vincoli tecnologici, spaziali e/o temporali. Si tratta, quindi, di riuscire a mettere in relazione diverse tipologie di dati: strutturate (in archivi e database), semi-strutturate e non strutturate (come blog, e-mail, testi e immagini).

Gli Open Data (OD), in quanto informazioni strutturate, sono quindi un sottoinsieme di dati che alimentano una cultura della partecipazione, contribuendo a creare un ecosistema collaborativo, composto da PA, cittadini e imprese, fondamentale per la crescita del territorio di riferimento. Di contro, le criticità possono essere ricondotte, da una prima analisi, all'impatto organizzativo dei dati: il patrimonio informativo pubblico è costituito, infatti, da una grande mole di dati "grezzi", detenuti da enti diversi. Riversare gli archivi pubblici su internet, potrebbe sembrare un'operazione di trasparenza, ma con poco significato da un punto di vista pratico. Occorre, quindi, processare i dati per ricavarne "dati raffinati", pronti per essere utilizzati; ma questa elaborazione presuppone un grande lavoro, anche e soprattutto riguardo la tempistica. Occorre, in secondo luogo, far percepire alla società civile il reale vantaggio degli OD, in modo che essi vengano poi opportunamente utilizzati.

Per quanto riguarda i BD, i vantaggi e le opportunità sono il monitoraggio di grosse quantità di dati, la visualizzazione e la recupero degli stessi. Questo significa processare grandi quantità di dati per ricavarne pattern adatti a individuare trend e opportunità nuove in campo politico, sociale ed economico. Questa è in realtà la sfida del Web 3.0. le criticità e le problematiche inerenti ai BD riguardano ovviamente il volume di dati da analizzare, ma anche la velocità di estrazione e l'eterogeneità di cui tener conto. Ciò implica costi, soprattutto iniziali, elevati, sia in termini di hardware che di software. Non esiste ancora, nella PA, un percorso predefinito, ma i benefici di tali investimenti potrebbero essere nel breve-medio periodo particolarmente significativi.

Rosangela Muscetta [<http://www.economia-conoscenza-itc-km.blogspot.it>]