



Misurazione della qualità dei dati

L'ISO/IEC 25024 estende l'ISO/IEC 25012 "Data quality model" del 2008 al campo delle misurazioni, definendo 63 misure di qualità applicabili alle caratteristiche di qualità dei dati.

La conformità all'impiego di tale standard è assicurata anche motivando il mancato uso di alcune delle misure, modificandole o aggiungendone altre purché seguano il metodo di applicazione proposto. Non sono previste soglie di accettabilità, ma gli utilizzatori dello standard sono invitati a definirne di proprie nel proprio contesto d'uso.

Tra le 15 caratteristiche di qualità del 25012 sono state enfatizzate in passato, dalla Determinazione Commissariale 68/2013 dell'AgID, l'accuratezza (sintattica e semantica), la coerenza, la completezza e l'attualità del dato, in particolare da applicare alle banche dati di interesse nazionale:

- repertorio nazionale dei dati territoriali;
- anagrafe nazionale della popolazione residente;
- banca dati nazionale dei contratti pubblici di cui all'articolo 62-bis;
- casellario giudiziale;
- registro delle imprese;
- gli archivi automatizzati in materia di immigrazione e di asilo.

Nel 2014 inoltre l'ISO/IEC 25012, tradotto in italiano con la sigla UNI ISO 25012, è stato raccomandato come standard di riferimento per assicurare il raggiungimento di determinati livelli di qualità dei dati gestiti ("Linee guida nazionali per la valorizzazione del Patrimonio informativo Pubblico", 2014, pag. 19).

Proprio per la migliore quantificazione dei livelli di qualità dei dati che l'ISO/IEC 25024 propone elementi di misura che prevedono differenti funzioni di calcolo.

Nello standard vengono direttamente toccati vari temi che comprendono, oltre quelli citati: credibilità, sicurezza (privacy), accessibilità dei dati e loro comprensibilità, disponibilità, portabilità e reperibilità, fattori importanti per l'uso di dati pubblici (come per esempio dell'Anagrafe Nazionale e del Fascicolo Sanitario Nazionale).

L'applicazione della norma potrà essere utile per controllare il livello di qualità dei dati e favorire l'interscambio, l'integrazione e l'interoperabilità, la condivisione e l'ottimizzazione di servizi. Gestire la qualità dei dati con gli standard citati è



importante anche per favorire:

- il governo della crescente disponibilità di dati;
- l'acquisizione di dati la cui qualità è sconosciuta;
- la gestione di informazioni spesso insoddisfacenti;
- la focalizzazione dei requisiti di qualità dei dati;
- la riduzione della dispersione di dati tra i vari owner ed utenti;
- l'incremento di dati riusabili (eliminando ambiguità semantica);
- la co-esistenza di legacy system con sistemi aperti;
- la riduzione delle duplicazioni e l'impegno di risorse;
- il miglioramento dei processi che causano dati errati;
- la stima dei costi della non qualità;
- l'eliminazione progressiva dei modelli cartacei di acquisizione dati.

Il miglioramento della qualità dei dati, e la maggiore diffusione delle tecniche di misurazione, dipende da vari fattori tra cui primeggia l'adesione a modelli di qualità condivisi. Il raggiungimento della qualità non è in ogni caso frutto di un impegno sporadico di singole istituzioni, ma il frutto di una sinergia concertata che, basata su un cambio culturale, si apra a collaborazioni orizzontali che, pur nel rispetto della privacy, consentano un maggior dialogo tra le banche dati e razionalizzazione delle informazioni. Tra i benefici generali possono includersi anche quelli relativi ad una sempre più ricca disponibilità di Open data che possa soddisfare i principi di credibilità e accuratezza.

In un modello cooperativo per il miglioramento dei dati, anche il cittadino, conoscitore della realtà che lo riguarda, può essere parte attiva. E' sempre più frequente infatti la richiesta di collaborazione, da parte di istituzioni, dell'apporto del cittadino per la verifica di informazioni complesse orientate al raggiungimento della qualità globale intersettoriale.

Una nuova cultura della qualità dei dati è uno stimolo allo sviluppo e una spinta al miglioramento dei servizi, di persone e tecnologia, accompagnandosi alle nuove tendenze che vanno verso un maggior riuso dei dati, una qualità cross-enterprise, una concezione di sistemi orizzontali più che a Silos verticali non comunicanti.

La qualità dei dati (e delle informazioni) sta divenendo sempre più una componente essenziale della qualità in uso, caratterizzata da servizi con caratteristiche di efficacia, efficienza, mitigazione dei rischi (economici, ambientali e della salute), flessibilità.



Lo standard, di circa 50 pagine, è composto dai seguenti capitoli:

- Ambito di applicazione;
- Definizione dei termini;
- Insieme delle misure basilari con indicazione di:
 - o Sigla identificativa della misura;
 - o Breve descrizione;
 - o Funzione o algoritmo di calcolo;
 - o Fase del ciclo di vita del dato alla quale si applica la misura;
 - o Prodotto specifico di interesse;
- Allegati informativi riguardanti esemplificazioni degli algoritmi più complessi e quadri sinottici di ausilio all'applicazione delle misure.