ATTUALITÀ

L'industria dei dati pubblici, il motore della riforma della PA

di **Fondazione AIDR** · 3 marzo 2021, 09:20

di Alessandro Capezzuoli, funzionario ISTAT e responsabile osservatorio dati professioni e competenze Aidr

Aperti, aggiornati, strutturati, machine readable e corredati dai metadati: i dati prodotti dalle Pubbliche Amministrazioni, per essere realmente utilizzabili, dovrebbero avere almeno queste caratteristiche. Sono decenni, ormai, che si sente parlare delle numerose possibilità offerte dai dati e delle ricadute, in termini di conoscenza e di benessere collettivo, conseguenti alla loro condivisione. Eppure, nonostante nel settore privato sia evidente il valore attribuito ai dati, talmente elevato da essere "pagato" con un corrispettivo in servizi gratuiti di ogni tipo, il settore pubblico sembra ancora troppo inconsapevole delle potenzialità informative di cui dispone e impreparato rispetto alle politiche da attuare. In realtà, l'impreparazione è più che altro dovuta a una specie di ostruzionismo burocratico e formale che impedisce di definire degli accordi snelli e veloci tra le amministrazioni. Per questo, la condivisione dei dati, prima di arrivare alle questioni tecnologiche riguardanti la cooperazione applicativa, viene ostacolata da protocolli d'intesa manzoniani firmati e controfirmati da dirigenti, direttori e presidenti, che, nel migliore dei casi, richiedono mesi di tempo per essere formalizzati. Nel peggiore, le trattative terminano con un nulla di fatto. C'è stato un periodo, circa quindici anni fa, in cui parlare di condivisione e open data andava di moda: chiunque si lanciava in riflessioni fantasiose e proiezioni spericolate di ogni tipo, a volte veniva perfino interpellato chi ne sapeva realmente qualcosa e che, proprio per questo motivo, è stato escluso dai consessi importanti. Poi, la moda è passata e la questione open data è stata considerata più o meno risolta. Anche perché si è palesata una parola sicuramente più comunicativa, misteriosa e affascinante, il termine "big", che ha avuto il potere di arrestare il processo di diffusione e di condivisione dei dati: tutto si è fermato ad alcune esperienze virtuose e a qualche file di testo che ancora resiste, eroicamente appeso alle pagine di un sito dimenticato, come una vecchia canottiera a costine stesa sui fili arrugginiti di una casa abbandonata. Come

spesso accade, la normativa esiste ed è chiara: l'articolo I del CAD prevede che i dati aperti debbano essere:

- *disponibili con una licenza o una previsione normativa che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali, in formato disaggregato;*
- *accessibili attraverso le tecnologie digitali, comprese le reti telematiche pubbliche e private, in formati aperti e provvisti dei relativi metadati;*
- resi disponibili gratuitamente attraverso le tecnologie digitali, oppure resi disponibili ai costi marginali sostenuti per la loro riproduzione e divulgazione (salvo quanto previsto dall'articolo 7 del decreto legislativo 24 gennaio 2006, n. 36).

A dispetto delle norme, però, la situazione reale è ben diversa. In primo luogo perché all'interno delle PPAA non sembrano esserci molte persone che conoscano approfonditamente i dati e il loro ciclo di vita e siano in grado di attuare strategie di condivisione stabili e di lungo periodo. I dati prodotti e condivisi dalle istituzioni, almeno di quelle che fanno parte del Sistema Statistico Nazionale, dovrebbero garantire la qualità, la completezza dei metadati e il rispetto degli standard internazionali di diffusione. Per produrre dei dati con queste caratteristiche, occorre **industrializzare** il processo di produzione e fare in modo che la diffusione non sia il compito di qualche volenteroso che inserisca manualmente un file di testo su uno dei tanti portali, ma la conclusione di un flusso informativo che passi per la raccolta, la validazione, l'archiviazione, la pubblicazione e, possibilmente, la visualizzazione. Costruire **"l'industria dei dati pubblici"** è molto oneroso e impegnativo: la pandemia ha dimostrato ampiamente l'impreparazione del sistema Paese, soprattutto in una situazione di emergenza, nella costruzione di una metodologia di raccolta rigorosa e affidabile e di un sistema di validazione e di condivisione trasparente e strutturato. Questi limiti, in una condizione di normalità, devono spesso fare i conti anche con la duplice anima delle istituzioni, che producono contemporaneamente dati di flusso e dati di stock. I due processi produttivi, pur avendo degli elementi comuni, sono governati da logiche molto diverse e richiedono l'impiego di metodologie e di tecnologie differenti per quanto riguarda le fasi di validazione, di diffusione e di visualizzazione. I dati di stock sono trattati attraverso l'impiego di tecniche consolidate e vengono aggregati con lo scopo di descrivere un certo fenomeno nella sua interezza, i dati di flusso descrivono l'evoluzione temporale di un fenomeno e, oltre a essere numericamente più consistenti, hanno delle specificità che richiedono trattamenti e tecniche di validazione e di diffusione diverse dai dati di stock, anche in relazione al GDPR. La validazione dei dati di stock, generalmente riferiti a un intero anno, richiede molto tempo in quanto gli archivi si devono consolidare e il processo scientifico per garantirne la qualità è molto oneroso: questo

vincolo non consente di avere dati aggiornati in tempo reale, ma permette di descrivere i fenomeni con molta precisione. La validazione dei dati di flusso segue un iter molto diverso, attraverso il quale non è al momento possibile garantire la stessa qualità dei dati di stock, ma in compenso risponde al bisogno crescente di numerosi ambiti di ricerca.

C'è poi una questione delicata che riguarda la distinzione tra i dati di sintesi e i dati puntuali: i primi possono essere trattati e condivisi senza vincoli particolari, i secondi, nella maggior parte dei casi, sono soggetti alla regolamentazione sul trattamento dei dati e impongono numerosi limiti non solo alla diffusione ma anche al trattamento e all'analisi da parte dei ricercatori.

Superato lo scoglio organizzativo e metodologico, che già di per sé rappresenta un limite notevole, c'è da affrontare la questione politica. Nonostante i proclami e le linee guida (molto spesso ignorate) dell'AGID, le pubbliche amministrazioni sono ancora dei feudi nei quali regnano le regulae societatis dei gesuiti, ovvero l'obbedienza incondizionata alle volontà dei superiori gerarchici e la negazione dell'evidenza, attraverso l'omissione della diffusione della conoscenza, per indirizzare il pensiero per mezzo di ordini precisi dettati dalla Divina Provvidenza, che, chissà perché, ha sempre sembianze molto umane. Questo aspetto rende gli archivi delle istituzioni assimilabili a dei fortini inespugnabili, protetti da un recinto chiamato "privacy", che ne legittima di fatto l'isolamento. Se è vero che negli ultimi anni la collaborazione tra istituzioni è stata rafforzata, e alcuni archivi, soprattutto stock, sono stati condivisi, è altresì vero che le metodologie adottate per la condivisione dei dati sono assolutamente inadeguate rispetto ai mezzi disponibili e fanno ricorso ancora a vecchi e insicuri metodi di trasferimento manuali (upload o FTP). In altre parole, non esiste una governance nazionale che definisca strategie, metodi e infrastrutture di condivisione, esistono più che altro prassi sedimentate che non tengono conto delle evoluzioni del mondo e della tecnologia e, soprattutto, della necessità di creare un'industria dei dati pubblici. Eppure, le pubbliche amministrazioni dispongono di patrimoni informativi ricchissimi, che vanno dalle caratteristiche dei singoli individui ai dati economici, dai fabbisogni di personale ai bilanci, dalle competenze alle professioni svolte, attraverso i quali sarebbe possibile attuare consapevolmente tutte le riforme di cui il Paese ha bisogno. Il rinnovamento della PA passa attraverso un reclutamento del personale più efficace e consapevole, un'erogazione dei concorsi pubblici fluida e trasparente, una valorizzazione del merito, della conoscenza e dell'esperienza dei lavoratori, un'ottimizzazione delle spese e degli assetti organizzativi attraverso l'attuazione di politiche sul lavoro sostenibili in termini economici, produttivi e ambientali. È difficile, se non impossibile, immaginare una riforma che, ancora una volta, ignori il valore dei dati e faccia ricorso alla volontà della Divina Provvidenza. Se è proprio necessario arrendersi all'idea che la salvezza degli uomini non sia frutto del contributo di ciascun individuo al benessere della

collettività, ma una specie di miracolo compiuto da uno dei tanti salvatori della Patria, molto cari alle masse, tanto vale identificare il salvatore nei dati e non in un santone improvvisato che dispensi l'elisir delle riforme perfette.

Articolo pubblicato su Istituzioni24: https://www.istituzioni24.it/2021/03/03/lindustria-dei-dati-pubblici-il-motore-della-riforma-della-pa/

© 2026 Istituzioni24