



FOCUS:

**PUBBLICA AMMINISTRAZIONE: DIGITALIZZAZIONE,
INFORMAZIONI E DATI**

Intervista a Giorgio Alleva – Presidente ISTAT

a cura di Daniela Carlà e Guido Melis

In tutti i Paesi della UE cresce la domanda, istituzionale e non, di informazione quantitativa utile alla comprensione di fenomeni complessi e sempre più interconnessi. In Italia, l'Istituto nazionale di Statistica (ISTAT) si prepara ad affrontare le nuove sfide puntando a realizzare l'integrazione delle informazioni provenienti dalle tradizionali indagini statistiche, degli archivi amministrativi della pubblica amministrazione e delle "nuove fonti" di dati, ancora in parte da esplorare.

Come s'inquadra la strategia di creazione di "registri statistici" nel quadro della statistica europea?

La creazione del sistema dei registri che stiamo portando avanti si colloca nella cornice della strategia Vision 2020 che costituisce il programma di modernizzazione degli Istituti statistici europei e si fonda su alcuni obiettivi strategici, tra cui (I) l'integrazione dei dati; (II) l'uso massivo del patrimonio informativo della PA, (III) la diminuzione del fastidio (o costo) statistico per rispondere alle indagini; (IV) l'uso di nuove fonti di dati tra cui i Big Data. Questo modello di produzione delle statistiche ufficiali è incorporato formalmente nei regolamenti europei, approvati dagli organi rappresentativi attraverso procedure complesse e rigorose, e verificato dagli organismi di controllo. L'Italia non è un caso a sé: tutti i Paesi europei si stanno muovendo in questa direzione.

Quali sono i vantaggi e le criticità di questo nuovo orientamento?

La costruzione del sistema di registri statistici è finalizzata a produrre nuova informazione aggregata per i cittadini, le imprese e i decisori pubblici, attraverso l'unificazione e l'integrazione di fonti di dati differenti (archivi amministrativi, rilevazioni e Big Data). L'integrazione è finalizzata anche



alla produzione di nuova informazione su fenomeni emergenti (es. l'economia digitale o la globalizzazione) non catturabili mediante gli approcci statistici tradizionali. I vantaggi di questo tipo di approccio sono molteplici. Lo sfruttamento sistematico delle fonti amministrative e di altre tipologie di bacini informativi semplifica e riduce gli adempimenti statistici di cittadini e imprese, attraverso un contenimento del ricorso alle indagini dirette e alla loro focalizzazione su fenomeni non coperti dai dati già disponibili. In questo modo, si generano anche risparmi per la Pubblica amministrazione per la realizzazione della funzione statistica, come nel caso del Censimento permanente della popolazione, che stiamo per avviare.

Si tratta di una tendenza che va affermandosi a livello internazionale e che ha visto, in Italia, rilevanti avanzamenti: tra il 2013 e il 2019, all'interno delle attività del Sistema statistico nazionale, la raccolta diretta di dati da famiglie, unità economiche e istituzioni diminuisce da 376 a 310 lavori (-17,6%); contestualmente aumentano da 206 a 275 gli utilizzi di fonti amministrative (+33,5%), spesso in combinazione con fonti statistiche dirette. Un ulteriore vantaggio è dato dall'aumento dell'accuratezza delle stime, che vengono ora generate in misura crescente sulla base di registri statistici esaustivi, con la conseguente riduzione o annullamento dell'errore campionario. Un terzo, sostanziale, vantaggio è dato dall'ampliamento della capacità informativa dell'output, con la possibilità di analisi di grande precisione e rilevanza delle eterogeneità presenti tra i soggetti (es. unità economiche, individui, ecc.), delle interrelazioni tra i diversi fenomeni (es. performance delle imprese e capitale umano; retribuzioni, caratteristiche dei lavoratori e strategie delle imprese, ecc.), degli aspetti dinamici. Un ultimo enorme vantaggio è costituito dalla possibilità di seguire i soggetti (siano essi individui, imprese o istituzioni) nel tempo e di coglierne le trasformazioni e i cambiamenti di stato in relazione agli eventi del ciclo di vita.

Alcune criticità di questo approccio riguardano la dipendenza significativa della produzione statistica dalla modulistica amministrativa e dalle scelte della Pubblica amministrazione in merito agli adempimenti informativi richiesti alle imprese, ai cittadini, alle istituzioni. Un ulteriore aspetto critico si può rilevare nella complessità organizzativa e gestionale del modello di produzione basato sui registri e sulla loro integrazione, generata sia dai notevoli flussi di dati necessari all'alimentazione del sistema sia dall'aumento delle interdipendenze tra i diversi settori statistici.



Nel progetto di arricchire la statistica ufficiale con nuove fonti, non tradizionali, quali tipologie di Big Data possono essere utilizzate e in che settori?

L'approccio della statistica ufficiale ai Big Data tende a considerarli non solo come una fonte di dati ma come una realtà informativa complessa, in grado di configurare nuovi paradigmi di misurazione. L'utilizzo dei Big Data a supporto della produzione di statistiche ufficiali è da diversi anni al centro dell'agenda europea del network di istituti nazionali di statistica coordinato da Eurostat. In particolare, nel 2013, con il cosiddetto "Scheveningen memorandum", si è aperta ufficialmente una prospettiva strategica di utilizzo di grandi masse di dati per finalità statistiche, sia in forme integrate con altre fonti sia come nuova fonte in sostituzione di quelle tradizionalmente utilizzate. Ciò ha richiesto la realizzazione di complesse attività di sperimentazione e innovazione per lo studio di metodologie adeguate a sfruttare le potenzialità dei Big Data e di definizione di complessi strumenti di valutazione della qualità dei dati utilizzati e degli indicatori prodotti. In Italia, le attività condotte dell'Istat di sperimentazione ed esplorazione dell'utilizzo dei Big data per migliorare la qualità delle statistiche correnti o per sviluppare nuovi indicatori sono orientate da un Comitato formato da esperti internazionali e nazionali, oltre che da membri Istat, con un programma di sperimentazioni su diversi temi.

Un avanzamento di rilievo è stato realizzato a partire da gennaio del 2018 con la messa a regime dell'utilizzo di scanner data come fonte che affianca le rilevazioni Istat per il calcolo mensile dell'inflazione, con riferimento ai prodotti alimentari e per la cura della persona e della casa. Entro breve tempo l'Istat diffonderà diverse statistiche sperimentali basati sui Big Data. Tra queste, oltre a statistiche dettagliate sui siti web esposti dalle imprese industriali e dei servizi e indicatori sulla propensione digitale delle unità produttive, è da citare un indice sperimentale, denominato "Social Mood on Economy Index", che fornisce misure ad alta frequenza - basate su una complessa metodologia di filtraggio, classificazione e aggregazione - sull'evoluzione del sentimento italiano sull'economia, derivate da campioni di tweet pubblici in lingua italiana, catturati in streaming. Si tratta di un indicatore che ha un valore informativo in sé e se messo in relazione ad altre serie economiche o sociali anche a più bassa frequenza. Inoltre, l'Istituto sta investendo sull'utilizzo dei dati di telefonia mobile per rilevare i flussi di mobilità e turismo, su quelli generati dagli Smart meters per i consumi energetici, sulle informazioni sul traffico ottenibili da webcam per il rischio d'incidentalità e per le statistiche sul traffico e sull'utilizzo di immagini per l'uso e la copertura del suolo, la stima delle aree



verdi e la stima delle produzioni agricole. Il mondo, anche dal punto delle tecnologie ICT e di comunicazione, è in continua evoluzione ed il piano dell'Istituto è di selezionare gli utilizzi più promettenti e sostenibili e incorporarli nella produzione ufficiale.

La normativa¹ italiana assegna all'Istat il compito di coordinare la modulistica amministrativa, intervenendo direttamente sul contenuto, la qualità e l'organizzazione degli archivi e dei sistemi informativi della P.A. al fine di renderli maggiormente adatti all'utilizzo statistico. L'Istituto sta svolgendo questo ruolo e con quali risultati?

In realtà, la normativa italiana assegna all'Istat il ruolo di promuovere l'utilizzo "a fini statistici degli archivi gestionali e delle raccolte di dati amministrativi" fin dal 1989 (d.lgs. n. 322/89), concetto ripetuto più volte e ribadito infine nel D.P.R. n. 166/2010 ("coordinare modificazioni, integrazioni e nuove impostazioni della modulistica e dei sistemi informativi utilizzati dalle pubbliche amministrazioni per raccogliere informazioni utilizzate o da utilizzare per fini statistici"). Si tratta di disposizioni che possono essere applicate solo in un clima di forte collaborazione tra amministrazioni, intervenendo in diverse fasi del processo legislativo ed esecutivo. I risultati sono quindi alterni: in alcuni casi si è riusciti a ottenere una solida convergenza tra interessi statistici ed esigenze amministrative; in altri, le finalità amministrative sono risultate prevalenti e non modificabili. La linea da seguire appare sempre più quella di garantire qualità agli archivi amministrativi e chiarezza delle definizioni, in modo da non creare ambiguità nel momento in cui si utilizzano i dati in essi contenuti.

Come valuta, ai fini statistici, la qualità degli archivi amministrativi che Istat sta utilizzando?

La qualità è ovviamente variabile tra i diversi archivi, ma i loro contenuti sono comunque di estremo interesse per l'Istat. Nel corso degli anni si sta registrando una crescita esponenziale del ricorso ai dati contenuti negli archivi. Nel 2015 si acquisivano 90 archivi da 40 Enti, con 280 forniture; nel 2018 sono 50 gli Enti che ci forniscono 190 archivi, per un totale di 428 forniture. Il lavoro "statistico" da fare sugli archivi è fondamentale e anche molto oneroso: in primo luogo si deve lavorare

¹ Art.2, c.2, l. c) del D.P.R. n. 166/2010



sui metadati, capire cioè se e in che misura le informazioni contenute negli archivi sono effettivamente quelle che servono alla produzione statistica; se, in altri termini, vengono usate le stesse classificazioni e le stesse definizioni che interessano il fenomeno che si sta studiando. In seconda battuta, va svolto un enorme lavoro sulla corretta identificazione delle unità che si ritrovano nell'archivio per essere sicuri che siano le stesse unità che si ritrovano anche in altri archivi.

Come step successivo, è necessario procedere alla messa in sicurezza delle informazioni rendendole non direttamente identificabili. Soltanto a questo punto inizia il lavoro di analisi dei contenuti, in termini di coerenza, pertinenza e valore aggiunto che le informazioni presenti negli archivi possono offrire come contributo alla produzione di statistica ufficiale. La qualità dell'archivio è quindi un concetto articolato, riguarda i dati amministrativi in entrata, ma riguarda anche tutto il complesso lavoro di trattamento che li trasforma in dati statistici. Il passo ulteriore, che è poi quello decisivo, consiste nel riuscire a considerare congiuntamente tali dati (statistici) con i dati che provengono dalle rilevazioni dirette che l'Istat effettua su famiglie, istituzioni e imprese, prevedendo la complementarietà tra le fonti e, in alcuni casi, anche la possibile sostituzione di una fonte diretta con l'elaborazione effettuata sulla fonte amministrativa. L'utilizzo dei dati amministrativi per la costruzione delle statistiche pubbliche non solo alleggerisce il disturbo sui cittadini, evitando la proliferazione di indagini statistiche, ma arricchisce la capacità informativa e dà profondità alle analisi, consentendo una lettura moderna ed efficace della complessità dei fenomeni economici e sociali. Così è stato per la salute, per portare un esempio, in cui si sono potute porre in luce le disuguaglianze sociali. Anche in tema di disabilità, solo utilizzando tutte le fonti disponibili è stato possibile fornire l'informazione necessaria per la tutela di diritti delle persone disabili.

Il punto cardine della nuova strategia dell'Istat è l'integrazione tra archivi e indagini. Proprio recentemente, però, il Garante della Privacy ha rilasciato un parere molto "preoccupato" per il modo in cui l'Istat sta costruendo il "sistema integrato dei registri": La raccolta dei dati dei cittadini fatta dall'Istat per stilare censimenti e statistiche attraverso il "codice unico" chiamato Sim, determina una vera e propria schedatura permanente di ogni individuo, nel tempo e nello spazio, con gravi rischi per i diritti e le libertà degli interessati». Come e in che misura questo parere impatterà sulle attività dell'Istat? Il sistema di produzione statistica basato sui registri e sulla loro integrazione richiede infrastrutture tecnologiche e di governance adeguate ad utilizzare per finalità statistiche grandi masse di dati individuali nell'assoluto rispetto della privacy. Il SIM è comunque un'infrastruttura interna all'Istituto, soggetta ai più stretti requisiti di sicurezza, trasparenza e tracciabilità relativamente



all'accesso e al trattamento dei dati, in modo che i dati stessi non siano usati per scopi diversi da quelli della statistica ufficiale. Da questo punto di vista il SIM risponde alla necessità di creare un ambiente di memorizzazione di elevata sicurezza e protezione per i dati amministrativi; attribuire un codice numerico "non parlante" (cioè che non consente di identificare il soggetto) alle unità statistiche; procedere a integrazioni di dati allo scopo di mettere a disposizione degli utilizzatori interni autorizzati gli archivi integrati, privi di dati identificativi diretti, necessari alla successiva produzione dei dati statistici nel pieno rispetto della normativa in materia di protezione dei dati personali. Si tratta quindi di infrastrutture che rappresentano una garanzia per la tutela della riservatezza: i dati individuali vengono infatti utilizzati a fini statistici sulla base di codici che non consentono l'identificazione dei soggetti e permettono di sfruttare adeguatamente il patrimonio informativo in possesso della pubblica amministrazione. Riteniamo che ulteriori approfondimenti conoscitivi consentiranno di chiarire il ruolo del SIM con il Garante della Privacy ed evidenziare il suo ruolo infrastrutturale di strumento proprio a tutela della riservatezza.

Più in generale, come ritiene si possa conciliare l'esigenza di tutelare il diritto alla protezione dei dati personali con la necessità di dare risposte alla crescente richiesta di informazioni dettagliate e puntuali che la società nel suo complesso, non solo le istituzioni, richiede?

La fiducia da parte dei cittadini sul nostro operato e il loro profondo convincimento sull'utilizzo esclusivamente a fini statistici e nel pieno rispetto della privacy dei dati raccolti rappresentano le condizioni essenziali per assolvere il mandato che la legge affida all'Istat. Agendo in piena collaborazione con il Garante per la Protezione dei dati personali, l'Istituto è impegnato sistematicamente per garantire la tutela assoluta della riservatezza dei dati, rafforzando misure e policy per assolvere a tale obbligo. Sono anni che l'Istat lavora sugli archivi amministrativi e mai si è verificato un solo caso di violazione della privacy. E fino ad oggi l'Autorità ha sempre espresso parere favorevole all'acquisizione e al trattamento di archivi amministrativi. All'interno di questa cornice dobbiamo considerare, come evidenzia anche nella sua domanda, che il dettaglio e l'integrazione dei dati a supporto di prodotti statistici sempre più rilevanti per la collettività rappresentano aspetti di grande competitività del modello di produzione statistica basato sui registri. La statistica ufficiale si è dotata di strumenti metodologici avanzati che consentono di controllare se la produzione di statistiche aggregate di notevole dettaglio e multidimensionalità rispetti le esigenze di tutela della privacy,



Anno 6, n. 11 – LUGLIO 2018

Nuova Etica Pubblica

Rivista dell'Associazione Etica PA

individuando i necessari correttivi. Da questo punto di vista c'è da considerare che la rete della statistica ufficiale europea riesce a mobilitare competenze di elevato profilo e condividere tra i diversi Paesi soluzioni metodologiche e tecniche di eccellenza. Sul piano operativo, la consapevolezza del ruolo decisivo della tecnologia per garantire un'acquisizione, memorizzazione e trattamento dei dati individuali in condizioni di massima sicurezza ha stimolato la realizzazione di ambienti tecnologicamente avanzati e con una governance professionale e organizzativa di elevato livello qualitativo.