

# IL PROGETTO IN PARTNERSHIP DI DATA ARCHIVE PER I MICRODATI ITALIANI

di Vittoria Buratta (buratta@istat.it)

■ Per venire incontro all'esigenza degli utenti di disporre di una maggiore offerta di microdati e contestualmente di un insieme di servizi che ne facilitino il reperimento, l'Istat sta avviando un progetto in partnership con la Banca d'Italia, la Fondazione Bruno Kessler dell'università di Trento e il Ministero dell'istruzione, dell'università e della ricerca, per la creazione di un *data archive* italiano per i microdati.

Si tratta di un portale unico di servizio grazie al quale ciascuna istituzione della statistica pubblica e della ricerca scientifica potrà mettere a disposizione la propria raccolta di dati elementari.

## IL CONTESTO DI RIFERIMENTO

Storicamente la circolazione dei dati elementari è stata assicurata dai *data archive*. Il *data archive* nasce alla fine degli anni '70, inizialmente come struttura interna all'università che raccoglie dati elementari provenienti da indagini statistiche di tipo sociale e che opera poi per renderli fruibili alla comunità, grazie ad un'attività coerente e sistematica di archiviazione dei file, a un processo di documentazione basato su sistemi standard di metadati e alla cura della loro anonimizzazione. Nel tempo si sono costituite vere e proprie istituzioni che sono diventate un polo riconosciuto per la raccolta di microdati prodotti dalle università di un Paese e che promuovono l'utilizzo di standard nei vari campi di interesse, propongono corsi per avviare studenti e ricercatori all'analisi di dati elementari, offrono fondi a studiosi per l'analisi dati degli stessi.

Oggi, i *data archive* costituiscono il punto di riferimento delle politiche europee di distribuzione e accesso ai dati elementari. Le iniziative in tema di accesso ai microdati stanno vivendo un'accelerazione grazie alla spinta proveniente dallo sviluppo tecnologico – che permette interrogazioni via web a cataloghi di metadati molto estesi – e, soprattutto, grazie alla volontà dei paesi europei di creare delle istituzioni dedicate a favorire la circolazione dei dati elementari (nella maggior parte dei casi raccolti con fondi pubblici) all'interno della comunità scientifica e più in generale all'interno della società nel suo complesso.

## RACCOLTA E ACCESSO AI DATI ELEMENTARI

Nell'ambito della politica di diffusione dei microdati, l'Istat ha sviluppato negli ultimi anni importanti progetti quali la realizzazione del repository di microdati validati dell'Istituto (ARMIDA), il rilascio di diversi tipi di file di dati elementari per utenti diversi (file standard, file per la ricerca e file per il Sistan), la creazione del Laboratorio per l'analisi dei dati elementari (ADELE), con il successivo sviluppo di una rete dei laboratori regionali e, da ultimo, la progettazione di file di dati di più ampia fruizione (mlcro.STAT). Peraltro, l'Istat ha promosso una riflessione volta anche alla diffusione di microdati raccolti da altri enti appartenenti al Sistema statistico nazionale, a partire dalle amministrazioni centrali fino ad arrivare agli enti locali, muovendosi quindi nell'ottica propria dei *data archive* che prevedono appunto la messa a disposizione di dati elementari non

solo propri, ma anche prodotti da altri enti o istituzioni.

## LA COSTITUZIONE DI UN SISTEMA INTEGRATO DI DATA ARCHIVE

Il progetto in partnership per la creazione di un *data archive* italiano consentirà di:

- offrire una visione d'insieme dei dati elementari prodotti dalla statistica ufficiale e dalle istituzioni della ricerca pubblica (università ed enti di ricerca) e delle corrispondenti modalità di accesso;
- svolgere interrogazioni flessibili (ricerche tematiche, per parole chiave o per variabili presenti nei singoli file);
- indirizzare l'utente all'istituzione che offre il servizio prescelto.

In questo modo il reperimento dei dati sarà facilitato, la documentazione e le procedure tecniche armonizzate con una ricaduta positiva sulle attività di analisi e ricerca scientifica. Più in generale, un *data archive* integrato a livello nazionale non solo renderà più visibili i dati elementari raccolti in Italia, ma potrà costituire l'occasione per ampliare il patrimonio informativo messo a disposizione da ciascuna istituzione, per standardizzare i metadati associati e per rivedere le modalità di accesso in un'ottica di armonizzazione del servizio e di una maggiore facilitazione all'utente.

Questa iniziativa, che potremmo inscrivere nell'ambito della più generale strategia dell'*open access* all'informazione statistica e del rilascio di dati aperti, sarà anche una straordinaria occasione per far maturare e crescere la cultura statistica nel nostro Paese.