

Un tesoro nascosto. Lettera aperta di una ricercatrice che non trova

scritto da Rina Camporese | 19 Maggio 2015



Evviva i dati aperti. L'Istat ne offre tanti. Ma accedere ai più preziosi – i microdati – è un' Odissea, e degli eroi che intraprendono l'avventura pochi arrivano a destinazione. Lavoro nella sede Istat di Venezia e continuo a ricevere da università ed enti di ricerca richieste di accesso ai microdati. Quasi sempre le richieste non si riescono a soddisfare perché le politiche Istat per l'accesso sono complesse. Non parlo dei microdati già accessibili ad oggi, e ce ne sono, ma di quelli veramente utili a chi è professionista della ricerca e non si accontenta.

Penso che in statistica ufficiale si commettano ancora troppi peccati di omissione: non si promuovono abbastanza tutti i modi possibili per valorizzare l'immenso patrimonio di dati, chiuso a doppia mandata nelle nostre segrete. Quante informazioni, quanta conoscenza si potrebbero far emergere se solo sfruttassimo, o lasciassimo sfruttare, tutti i dati che abbiamo? A volte mi sembra che ci comportiamo come quei compositori che tengono gelosamente i loro spartiti e fanno uscire solamente i brani eseguiti da loro stessi. Quella stessa musica, però, interpretata da altri, avrebbe colori e sfumature diverse e sarebbe un piacere e un arricchimento per molte altre orecchie.

In Istat, tra l'altro, non c'è nemmeno il problema del diritto d'autore: è sufficiente citare la fonte.

Qualche esempio

Un esempio? Microdati arricchiti dell'indicazione puntuale del luogo a cui si riferiscono. Se conosco le coordinate geografiche a cui si riferiscono i dati posso ricavare più facilmente variabili di contesto che aiutino a interpretarli, e posso studiare i fenomeni nello spazio senza i vincoli delle suddivisioni amministrative del territorio. Certo, ciò aumenta il rischio di identificazione dell'unità a cui i dati si riferiscono, ma qui entra in gioco l'etica professionale.

Un altro esempio: microdati provenienti da fonti diverse agganciati per chiave identificativa. Ad oggi non è così difficile ottenere i microdati provenienti da una singola fonte, ma l'integrazione di fonti diverse è dura da ottenere, anche se i dati potrebbero essere facilmente resi anonimi immediatamente dopo aver usato i codici identificativi per effettuare l'aggancio.

Un altro esempio ancora: i microdati più succosi al momento si possono elaborare solamente in

ambiente protetto, all'interno dell'Istat e secondo un piano di output rigorosamente definito e approvato a priori. A chi fa ricerca, tuttavia, pur nel rispetto rigoroso delle ipotesi di lavoro e delle metodologie, accade spesso di delineare gli output in modo maturo soltanto grazie al percorso di ricerca, un percorso che si svolge spesso in modo dialettico e in gruppo. Tutto ciò è improponibile con le attuali regole di accesso ai laboratori Istat di microdati.

Obiezioni e repliche

Si potrebbe controbattere che l'operazione di rilascio ha un costo, ma le tecnologie attuali non la rendono particolarmente complessa. Inoltre, i mancati guadagni in conoscenza per lo scarso sfruttamento del patrimonio di dati sono certamente superiori ai costi materiali del predisporre file di microdati più ricchi.

Non dimentico certo la tutela della riservatezza. Chi ha generosamente condiviso informazioni personali per costruire il bene collettivo della conoscenza statistica - oltre che essere ringraziato - ha diritto ad essere protetto grazie al segreto statistico.

Inoltre, microdati di questo tipo richiedono una professionalità statistica matura, padrona delle tecniche e consapevole delle responsabilità etiche nel maneggiare informazioni individuali delicate. Per fortuna c'è chi ce l'ha, dentro e fuori dall'Istat.

L'uso dei microdati da parte di professionisti della ricerca, che si impegnano, attraverso un progetto di ricerca, ad un uso etico del dato per fini di conoscenza, porta con sé lo stesso rischio di omicidio all'arma bianca dell'uso di un coltello da parte di un cuoco stellato, che nel menù propone tagliata di manzo con dadolata di verdure e aceto balsamico.

Non nego, quindi, che la diffusione di microdati ponga problemi organizzativi, professionali ed etici. Li pone, eccome. Sostengo, però, che sia doveroso affrontarli in vista della potenziale conoscenza che tali dati, patrimonio pubblico, possono esprimere.

A, le argomentazioni sacrosante di chi frena la loro diffusione per motivi di tutela della riservatezza e timori legati all'uso professionale del dato dovrebbero trovare un contraltare nelle ragioni di chi vorrebbe usare i dati per creare un bene comune, all'interno di un dialogo costruttivo che affronti le difficoltà in vista di un guadagno di conoscenza condivisa.

Possiamo fare qualche passo in più in questo senso?