



www.lavoce.info

[Informazione](#) / [Innovazione e Ricerca](#)

## L'ISTAT CHE VORREMMO

di [Andrea Ichino](#) 18.04.2005

*L'Istat è ancora molto lontano dagli standard di altri paesi per quel che riguarda la quantità, la qualità e la facilità di accesso ai dati per la ricerca. Non solo per i problemi connessi alla tutela della privacy, ma anche per inefficienze proprie. Molte importanti questioni vengono così discusse sulla base di pregiudizi ideologici, senza minimamente curarsi di misurare correttamente i fatti. E' una questione di democrazia e trasparenza: la comunità scientifica deve poter accedere ai microdati elementari per controllare e replicare i risultati.*

L'Istat che vorremmo di Andrea Ichino

Nonostante qualche recente passo avanti, l'**Istat** è ancora molto lontano dagli standard di altri paesi per quel che riguarda la quantità, la qualità e la facilità di **accesso ai dati** per la ricerca. Questo ritardo è dovuto in parte alla famigerata disciplina per la tutela della privacy (vedi i precedenti interventi su [www.lavoce.info](http://www.lavoce.info)) (1). **Tuttavia, grazie al "Codice deontologico" recentemente approvato, l'Istituto ha margini di azione che potrebbe sfruttare meglio.**

### Non solo tutela della privacy

**Ma agli effetti della tutela della privacy, si aggiunge anche un'inspiegabile inefficienza dell'Istat nel fare almeno quello che in altri paesi è considerato del tutto normale. Ad esempio, rendere i dati del censimento disponibili in meno di quattro anni dalla raccolta (2). Oppure, fare in modo che i file standard delle poche banche dati disponibili siano direttamente scaricabili dal sito dell'Istat a prezzi accessibili a chi fa ricerca (si veda ad esempio quello che offre lo UK Data Archive: <http://www.data-archive.ac.uk>). O, almeno, raccogliere quelle informazioni elementari la cui mancanza in Italia lascia letteralmente di stucco chi ci osserva dall'estero.**

**Per esempio, non è possibile ottenere attraverso l'Istat il salario individuale in un campione rappresentativo della popolazione. E tanto meno è possibile ottenere questo dato per un numero sufficiente di anni insieme ad altre informazioni sugli individui stessi, tra cui, in particolare, la loro collocazione geografica.**

### Dati e dibattito politico-economico

**In un paese in cui si discute all'infinito nei salotti televisivi, sulle pagine dei quotidiani e nelle famiglie di fenomeni statistici come la "perdita di potere d'acquisto dei salari", il "costo del lavoro", le "gabbie salariali", la "disuguaglianza salariale", le "insostenibili condizioni economiche dei lavoratori precari", l'Istat non è in grado di offrire a chi fa ricerca il dato**



[www.lavoce.info](http://www.lavoce.info)

statistico elementare con cui misurare e spiegare questi fenomeni. E il risultato è che tutti discutono sulla base di pregiudizi ideologici e di aneddoti privi di qualsiasi rappresentatività, senza minimamente curarsi di misurare correttamente i fatti.

L'Istat mi risponderà che, riguardo ai salari, le cose cambieranno presto con la nuova Indagine trimestrale sulle forze di lavoro. Ma allo stato attuale, l'informazione non è ancora disponibile perchè, dice il sito, ancora in fase sperimentale e soggetta al controllo dell'Eurostat. In ogni caso, con o senza i salari, gli altri dati dell'Indagine trimestrale non sono scaricabili direttamente dal sito e costano la ragguardevole cifra di 90 euro per quadrimestre (irraggiungibile, ad esempio, per i giovani ricercatori universitari notoriamente privi di fondi di ricerca). Un prezzo così elevato scoraggia chiunque dall'eseguire quelle elaborazioni preliminari necessarie per decidere se effettuare ricerche più approfondite. Anche se l'Istat fosse un monopolista che massimizza i profitti (cosa che chiaramente non è, e non dovrebbe essere), non sarebbe probabilmente conveniente fissare un prezzo così alto

### Le colpe di Eurostat

A giustificazione dell'Istat, bisogna ammettere che la Comunità europea e Eurostat non sono esenti da colpe gravi in fatto di dati per la ricerca, sia per quel che riguarda le direttive sulla privacy, sia per quel che riguarda la predisposizione di dati per la ricerca. Ad esempio, è totalmente inspiegabile la decisione di Eurostat di interrompere la raccolta dello European Community Household Panel che dal 1994 al 2001 ha fornito ai ricercatori europei una fonte inestimabile di micro-dati longitudinali comparabili tra paesi sulla situazione demografica, economica e lavorativa di individui rappresentativi delle rispettive popolazioni. (3) In altri paesi europei, dati simili venivano raccolti già prima e continuano a essere raccolti adesso. In Italia, per quel che ne so, l'indagine è stata interrotta. Il nostro Governo si scaglia contro l'Europa su questioni molto più controverse come la revisione del Patto di Stabilità, ma quando si tratta di fornire dati per la ricerca non esita ad allinearsi con le arretratissime posizioni di Eurostat.

Perché, nonostante le direttive europee, la situazione in altri paesi europei è migliore? Vorrei invitare i lettori a sfogliare gli indici delle riviste scientifiche internazionali, non solo in campo economico, per toccare con mano la quantità di questioni di enorme interesse per il dibattito politico che i nostri colleghi stranieri possono studiare grazie ai dati a loro disposizione. Un esempio per tutti: pensate a quanto si discute di immigrazione nel nostro paese senza uno straccio di banca dati che ci consenta di studiare il fenomeno. In Italia, se non fosse per l'Indagine sui bilanci delle famiglie italiane effettuata dalla Banca d'Italia, la ricerca microeconomica applicata sarebbe praticamente assente.

### La situazione in altri campi

Qualcuno potrebbe pensare che solo gli economisti abbiano queste esigenze. Non è così. Il cardiocirurgo Giulio Rizzoli scrive: "Tra gli ostacoli che il nostro paese pone al progresso scientifico ricordiamo la recente adozione di regole sulla privacy che impediscono all'Istat di permetterci la consultazione delle schede di mortalità, compilate dai colleghi. L'indagine



[www.lavoce.info](http://www.lavoce.info)

sulle modalità o le cause di morte è indispensabile alla comprensione dei rischi legati all'uso di nuove tecnologie biomediche, valvole incluse. Nel nostro caso deve rispondere al quesito se la morte è dovuta a complicanze della protesi o a cause cardiache o se non è a esse correlata". (4)

Questo è un esempio particolarmente significativo del fatto che la privacy non può essere considerata un bene assoluto. I dati che Rizzoli non può ottenere fornirebbero informazioni fondamentali per migliorare la terapie cardio-chirurgiche con ricadute positive per l'intera collettività.

L'Istat sognata dai ricercatori (e che i politici dovrebbero desiderare)

Qual è allora l'Istat che vorremmo? Vorremmo un Istat che facesse meno rapporti sintetici, meno comunicati stampa e meno tabelle aggregate su dati elementari che poi si tiene per sé. Vorremmo un Istat che invece usasse le sue risorse finanziarie e umane per raccogliere e rendere disponibili i micro-dati elementari di cui i ricercatori italiani hanno bisogno per poter fare un'attività di ricerca comparabile a quella dei loro colleghi stranieri. E si noti che gran parte di questi dati sono di fonte amministrativa e quindi raccolti con costi solo parzialmente a carico dell'Istat.

Vorremmo un Istat che catalizzasse e organizzasse la raccolta di dati in Italia seguendo le orme dello UK Data Archive, per evitare lo spreco attuale dei fondi spesi dai singoli ricercatori per raccogliere dati utilizzati una tantum e poi abbandonati. (5)

Vorremmo un Istat che sfruttasse in modo ampio e pro-attivo i margini di manovra lasciati aperti dal nuovo [Codice deontologico](#), almeno nella forma di permessi individuali su richiesta motivata nel caso di dati particolarmente sensibili e di ricercatori che diano prova di affidabilità. Vorremmo un Istat che si battesse a fianco dei ricercatori perché in Italia prevalga il principio enunciato nel [progetto di legge sulla tutela della privacy](#) presentato al Parlamento da Nicola Rossi, che mira a "consentire un accesso ampio e facile ai dati per la ricerca scientifica, anche in forma integrata tra archivi diversi, punendo però duramente un loro eventuale uso che danneggi i diritti della persona". Vorremmo un Istat che ci aiutasse a convincere il Garante per la privacy del fatto che i ricercatori non hanno alcun interesse a usare i dati individuali in contrasto con il rispetto del diritto alla tutela della riservatezza delle persone.

È anche una questione di democrazia e trasparenza: la comunità scientifica deve poter controllare e replicare i risultati che l'Istat trae dai dati che raccoglie. E per questo deve poter accedere ai micro-dati elementari.

Per saperne di più

Andrea Ichino "Le perplessità di un utilizzatore di dati di fronte al "Codice di deontologia e buona condotta per il trattamento di dati personali per scopi statistici e scientifici" in corso di pubblicazione su Rivista Statistica, 2005, scaricabile da <http://www.iue.it/Personal/Ichino/#downl.publ.papers>)

(1) Vedi in particolare: [Se la privacy non tutela la ricerca, Anche dati solo sui giornali](#),



[www.lavoce.info](http://www.lavoce.info)

Ancora sul codice deontologico.

(2) Attualmente sono disponibili solo i dati del censimento 1991. Negli Usa il sito <http://www.ipums.umn.edu> consente da tempo di scaricare i micro-dati di tutti i censimenti dal 1850 al 2000.

(3) Lo Echp sarà sostituito da una indagine non più longitudinale ma solo *cross section* (la Survey on Income and Living Conditions). Tuttavia dal sito [http://europa.eu.int/comm/eurostat/newcronos/reference/sdds/en/ilc/ilc\\_base.htm](http://europa.eu.int/comm/eurostat/newcronos/reference/sdds/en/ilc/ilc_base.htm) non si capisce quando i nuovi dati saranno disponibili, cosa conterranno in dettaglio e come vi si potrà accedere. Si sa però che il focus della nuova indagine sarà politically correct "a range of social statistics and indicators relating to the risks of income poverty and social exclusion".

(4) Dal saggio "Il follow-up del paziente portatore di protesi valvolari cardiache" che farà parte del volume "Guida ragionata alla cardiocirurgia" diretto ai medici di base.

(5) Vedi ad esempio <http://www.esds.ac.uk/aandp/create/depintro.asp>

Un commento all'articolo, di Giovanni Barbieri

Ogni contributo inteso a migliorare in quantità e qualità la restituzione dell'informazione statistica pubblica ai cittadini è prezioso per chi, come me, lavora ogni giorno con questo obiettivo. Anche noi abbiamo in mente "l'Istat che vorremmo" e cerchiamo di realizzarla. Mi sembra però ingeneroso attribuire all'Istituto "un'inspiegabile inefficienza" nel fare quello che negli altri paesi è considerato del tutto normale. A parte il fatto che ciò non è vero, poiché è facilmente dimostrabile e riconosciuto che in molti settori l'Istat è all'avanguardia e fornisce più informazioni statistiche degli altri paesi, non si può dimenticare che "l'Italia investe per la statistica ufficiale molto meno degli altri grandi paesi, sia in termini pro capite, sia rispetto al Pil", come sottolinea Enrico Giovannini sempre su LaVoce.info. Aggiungo soltanto che la mancanza di investimenti adeguati si fa sentire anche con riferimento alle risorse umane e materiali destinate alla diffusione: non c'è riunione internazionale in cui i colleghi degli altri istituti di statistica non siano sorpresi dal confronto tra ricchezza delle informazioni diffuse ed esiguità del personale dedicato a queste attività.

L'Istat ha dal 1996 un proprio sito destinato alla diffusione. Dal 2000 tutta l'informazione statistica è resa disponibile sul web a titolo assolutamente gratuito e senza altra richiesta che la citazione della fonte (anche tutte le pubblicazioni dell'Istituto sono "no copyright", da molto tempo prima che si aprisse il dibattito in materia). Tutta la produzione editoriale dell'Istat può essere scaricata senza vincoli, in Pdf (garantendo l'identità con la versione a stampa) e in formato rielaborabile. Un catalogo *on line* rende agevole le operazioni di ricerca. Sul sito sono presenti anche numerose banche dati, sistemi informativi, aree tematiche. Ogni giorno, gli 8mila utenti che ci visitano in media scaricano oltre due Gigabyte di dati; circa la metà, risulta dalle nostre rilevazioni sull'utenza, li rielabora per i propri scopi di documentazione e di ricerca. La scelta della diffusione gratuita sul web ha avuto l'Istat tra i precursori. Ancora oggi, non tutti gli istituti statistici dell'Unione europea mettono a disposizione gratuitamente l'intera produzione statistica diffusa. L'Eurostat, che attua questa



[www.lavoce.info](http://www.lavoce.info)

politica soltanto dal 1° ottobre 2004, non rende disponibile per il *download* tutta la produzione editoriale (ad esempio, non è possibile scaricare i Cd-Rom allegati ai volumi). Anche le osservazioni sulla disponibilità dei dati censuari mi sembrano ingenerose, oltre che inesatte. I dati dell'8° Censimento generale dell'industria e dei servizi sono *on line* dal 16 marzo 2004 e consentono l'elaborazione di tabelle personalizzate e confronti storici con i Censimenti precedenti, a partire dal 1951; tutti i fascicoli territoriali sono stati pubblicati e possono anche essere scaricati gratuitamente dal sito. Per quanto riguarda il 14° Censimento generale della popolazione e delle abitazioni, è stato adottato un diverso piano di elaborazione dei risultati proprio per venire incontro alle richieste rappresentate dagli utilizzatori: invece di procedere alla pubblicazione dei dati provincia per provincia, si è scelto di operare per variabili, in modo da anticipare la disponibilità delle informazioni sui fenomeni principali. Ad esempio, i dati sulle caratteristiche demografiche sono stati resi accessibili *on line* a due anni dalla raccolta.

In merito alla disponibilità di dati elementari resi anonimi, è bene precisare anzitutto che i prezzi praticati dall'Istat (90 euro) sono più che ragionevoli se confrontati con quelli praticati dall'Eurostat e dagli altri istituti di statistica europei. In Spagna, soltanto per fare un esempio, il file standard dell'indagine sulle forze di lavoro costa 201 euro a trimestre; Statistics Canada (un termine di riferimento per gli addetti ai lavori) vende il "*Survey of Labour and Income Dynamics: public use microdata file*" a 2.140 dollari canadesi (circa 1.300 euro). Senza cadere nella demagogia, i "giovani ricercatori universitari senza fondi di ricerca" trovano all'Istat facilitazioni. Anche l'accesso al *Laboratorio per l'analisi dei dati elementari* (Adele) è gratuito. Il problema dei dati elementari non è soltanto questo, e convengo con Andrea Ichino che si possono fare progressi. In seno all'Eurostat, l'orientamento – che trova l'Istat d'accordo – è quello di distinguere tra dati elementari resi anonimi confidenziali e non-confidenziali. I primi (in linea di massima, i *dataset* relativi alle imprese rientrerebbero in questa categoria) continuerebbero a essere sottoposti alle attuali restrizioni, ai sensi del Regolamento 831/2002 e della normativa italiana. I secondi sarebbero resi disponibili al pubblico a titolo gratuito e costituirebbero l'oggetto di un'area dedicata sul sito web dell'Istituto. Ritengo che questi sviluppi vadano incontro alle esigenze della comunità scientifica.

In conclusione, rimango convinto che il dialogo con gli utilizzatori dell'informazione statistica, e in particolare con quelli più attenti, per motivi professionali e di ricerca, sia fondamentale per migliorare la qualità e l'accessibilità dei dati. In questo convincimento sono confortato dall'opinione espressa da quanti utilizzano abitualmente i nostri dati e ci contattano per critiche e suggerimenti. Penso però che ci si debba attenere al confronto sui fatti, senza cedere alla tentazione del sensazionalismo. A mio parere, mondo della ricerca e Istat possono, anzi devono, lavorare insieme in modo costruttivo, perché migliorare i servizi e rafforzare le istituzioni della statistica pubblica sono due facce dello stesso obiettivo, quello di affermare il valore dei dati per tutti i cittadini.

Giovanni A. Barbieri

Direttore centrale per la diffusione della cultura e dell'informazione statistica  
Istat, Via Cesare Balbo 16, Roma

La controreplica dell'autore



[www.lavoce.info](http://www.lavoce.info)

Ogni istituzione pubblica Italiana si trincerava dietro la mancanza di risorse per giustificare la scarsità dei servizi offerti. Ed è indiscutibile che più fondi migliorerebbero le cose. Tuttavia, l'impressione (per quel che riguarda l'Istat ma anche il sistema universitario, sanitario, della giustizia etc.) è che, a parità di finanziamenti, le risorse esistenti (anche umane e materiali) possano essere utilizzate meglio o comunque in modo diverso. Nel caso specifico dell'Istat questo è quello che i ricercatori (almeno per quello che a me risulta) vorrebbero. Nel mio articolo ho cercato di dare suggerimenti concreti con intento costruttivo (in particolare il modello del <http://www.data-archive.ac.uk> che offre, gratuitamente ai ricercatori, una mole impressionante di dati elementari). Sono pronto a darne di ancora più concreti e costruttivi se l'Istat lo desidera.

Il mio articolo era in parte ispirato da un messaggio email ricevuto da un collega italiano che insegna in una università straniera (Daniele Paserman, Hebrew University, Jerusalem) e che mi ha scritto chiedendo: "Qual'è il metodo migliore per reperire dei dati sui salari in Italia a livello geografico abbastanza disaggregato (comune, provincia), possibilmente che vadano indietro di qualche anno. Sul sito dell'ISTAT non sono riuscito a trovare niente. Mi sapresti aiutare?" Come risponderebbe Giovanni Barbieri a questa richiesta?

Ai ricercatori interessano i dati elementari (downloadable e gratuiti) non la produzione editoriale. E non è una richiesta corporativa. Solo con i dati elementari i ricercatori possono essere di aiuto alla collettività affinché il dibattito politico economico sia basato su fatti e non su pregiudizi. Va benissimo che l'Istat faccia le sue analisi sui dati, se vuole e ha i fondi per farlo, ma anche i ricercatori devono poter aver accesso ai dati per darne la loro interpretazione in un dibattito serio e trasparente.

Vedo con piacere che il 4 aprile 2005 (dopo la stesura del mio articolo) l'Istat ha annunciato la diffusione dei dati del Censimento Generale della Popolazione del 2001. Ma andate a visitare il sito (<http://dawinci.istat.it/MD/>). Per quel che ne capisco mi sembra siano scaricabili solo tavole descrittive pre-elaborate. Probabilmente molte ore di lavoro e risorse sono state spese dall'Istat per produrre queste tavole e il relativo sito. Ma sono scaricabili anche (e più semplicemente) i dati elementari come ad esempio sul sito USA <http://www.ipums.umn.edu>? Per inciso, ecco un altro esempio che l'Istat potrebbe utilmente imitare.