

## I CODICI DEONTOLOGICI PER IL TRATTAMENTO DI DATI PERSONALI A SCOPI STATISTICI E DI RICERCA SCIENTIFICA: QUALCHE RIFLESSIONE

Alberto Zuliani

### 1. PREMESSA

Il direttore della rivista mi ha chiesto un breve commento in occasione della firma del Codice di deontologia e buona condotta per i trattamenti di dati personali per scopi statistici e scientifici da parte della Conferenza dei rettori delle università italiane, di alcune importanti società scientifiche nazionali e dell'associazione tra istituti di ricerche di mercato, sondaggi di opinione, ricerca sociale. Ho aderito volentieri, poiché ritengo che il codice rappresenti un effettivo progresso, dovuto alla tenacia del collega Ugo Trivellato il quale è riuscito a organizzare la pressione delle società scientifiche nei confronti del Garante per la protezione dei dati personali e a coinvolgere in qualche misura l'opinione pubblica, tradizionalmente disinteressata a problemi di questo tipo, sopra tutto da noi, per la scarsa diffusione della cultura quantitativa (Trivellato, 2003).

In precedenza, era stato deliberato e pubblicato il Codice di deontologia e di buona condotta per i trattamenti di dati personali a scopi statistici e di ricerca scientifica effettuati nell'ambito del Sistema statistico nazionale.

Ispirai personalmente, in qualità di presidente dell'Istituto nazionale di statistica, la scelta di procedere alla redazione di due codici distinti, con riferimento rispettivamente ai soggetti del Sistema statistico nazionale e agli altri soggetti pubblici e privati che operano per fini scientifici. Tuttavia, insistetti perché ambedue si informassero agli stessi principi. Questo è avvenuto e numerose norme risultano identiche nei due codici.

Accenno soltanto ai principi, poiché molti altri autori, in questo stesso numero della rivista, si soffermano su di essi dettagliatamente. Si tratta dei criteri di pertinenza e non eccedenza, imparzialità e non discriminazione nei confronti degli utilizzatori, sicurezza. Spendo qualche parola in più su quest'ultimo. I canoni della sicurezza, rispetto a raccolta, trattamento, conservazione, disponibilità per altri soggetti, sono collegati con l'uso di "mezzi ragionevoli" da impiegare per l'eventuale violazione del diritto alla riservatezza e l'informativa ai soggetti è collegata con uno "sforzo proporzionato" all'esigenza. Sono previsioni importanti che fan-

no venire meno alcune preclusioni di principio al trattamento di dati personali, a mio parere non giustificate.

La distinzione tra i due codici trova fondamento nella specificità dei soggetti del Sistema statistico nazionale. La statistica pubblica ed in particolare l'Istituto, prima centrale ora nazionale, di statistica hanno da sempre adottato il rispetto del segreto statistico come norma di comportamento, rifiutando di esibire o trasferire a chiunque dati individuali riconoscibili. In quanto uffici pubblici, i soggetti del Sistema sono inoltre tenuti a rispettare il segreto d'ufficio. Questi vincoli sono stati definiti per tutti gli uffici statistici pubblici dal decreto legislativo n. 322 del 6 settembre 1989 il quale ha previsto che sul rispetto del segreto statistico vigilasse la Commissione di garanzia per l'informazione statistica. Alcune limitate competenze in materia sono state mantenute dalla Commissione, ma il soggetto principale in materia è ovviamente il Garante.

Il decreto n. 322/1989 stabiliva anche che gli uffici statistici “forniscono al Sistema statistico nazionale i dati informativi previsti dal programma statistico nazionale relativi all'amministrazione di appartenenza, anche in forma individuale, ma non nominativa, ai fini della successiva elaborazione statistica”; soltanto una successiva normativa comunitaria, il Regolamento CEE n. 2186 del 22 luglio 1993 ha consentito di utilizzare in modo esteso per un particolare fine statistico, la costruzione dell'archivio delle imprese, i dati amministrativi. Questo, peraltro, spiega il ritardo con il quale l'Istat ha iniziato ad approfondire la problematica del trattamento dei dati contenuti in archivi amministrativi.

Prima ancora dell'emanazione del codice, per i soggetti del Sistema statistico nazionale era stato previsto, in relazione al contenuto della legge n. 675 del 31 dicembre 1996, che dessero specifica indicazione del trattamento di dati sensibili all'interno del programma statistico nazionale e, inoltre, che quest'ultimo, prima dell'approvazione da parte del Consiglio dei ministri, passasse al parere del Garante per la protezione dei dati personali.

In definitiva, il codice disciplina la circolazione dei dati personali fra gli uffici di statistica del Sistema, nel rispetto dei principi già ricordati di pertinenza e non eccedenza, di tutela della riservatezza e inoltre di qualità dei dati (sulla circolazione di informazioni fra soggetti pubblici subordinata al canone della qualità, vedi Zuliani, 2002). Il codice regola inoltre l'accesso da parte di altri soggetti pubblici e di privati per fini di ricerca.

## 2. ACCESSO AI DATI STATISTICI PUBBLICI

Il codice relativo al Sistema statistico nazionale ha confermato le modalità di accesso ai dati individuali attivate dall'Istat dal 1994. In particolare, erano state allora previste le seguenti tre forme:

- acquisizione di collezioni campionarie di dati individuali, così detti file standard;
- accesso al laboratorio Adele, laboratorio per l'analisi dei dati elementari, per effettuare elaborazioni presso la sede centrale dell'Istat, da parte di soggetti aventi finalità di ricerca;

- svolgimento di progetti di ricerca comuni fra Istat e altre istituzioni che pure avessero finalità di ricerca.

Nel seguito, mi soffermerò essenzialmente sull'esperienza dell'Istat, in primo luogo perché la possibilità di accesso ai dati statistici pubblici, e in particolare a quelli prodotti dall'Istat, da parte dei soggetti privati continuerà ad esercitarsi secondo le stesse modalità; in secondo luogo perché, in qualità di presidente dell'ente, promossi e sostenni l'apertura delle raccolte di dati individuali all'uso da parte di soggetti privati aventi finalità di ricerca. Delle iniziative conseguenti desidero dare conto qui.<sup>1</sup>

### 3. COLLEZIONI CAMPIONARIE DI DATI INDIVIDUALI

La disponibilità di “collezioni campionarie di dati elementari, resi anonimi e privi di ogni riferimento che ne permetta il collegamento con singole persone fisiche e giuridiche” era prevista dall'articolo 10 del decreto legislativo n. 322/1989. Con la direttiva del presidente dell'Istat n. SP/250.94 del 23 marzo 1994, la previsione normativa venne resa operativa. Ovviamente, risultava necessario indicare le metodologie idonee a valutare il rischio di identificazione del soggetto al quale i dati si riferivano e, se tale rischio fosse risultato eccessivamente elevato, mettere in atto le opportune protezioni. Queste ultime, in generale, comportano una riduzione del contenuto informativo del file. La direttiva dava indicazioni relativamente alla stima del rischio di violazione per l'intera collezione campionaria da rilasciare e prendeva le mosse dalle analisi effettuate presso l'Istituto (Biggeri e Zannella, 1991, Coccia, 1993, Crescenzi, 1993). Essa stabiliva i criteri e le procedure per il rilascio dei file standard relativi al censimento della popolazione e delle abitazioni del 1991 e alle più importanti indagini campionarie in ambito sociale: forze di lavoro, consumi delle famiglie, multiscopo sulle famiglie, sbocchi professionali dei laureati e dei diplomati, panel europeo delle famiglie.

Una delle ipotesi sottostanti la metodologia allora adottata dall'Istat per la determinazione del rischio di violazione consisteva nell'indipendenza delle unità statistiche. Essa non è rispettata nei casi in cui la collezione campionaria è caratterizzata dalla presenza di informazioni relative all'appartenenza delle unità a sottogruppi di popolazione quali ad esempio le famiglie (così detti dati gerarchici). Il problema posto da quest'ultimo tipo di dati riguardo alla riservatezza è di particolare gravità poiché l'identificazione di un'unità appartenente ad un gruppo può comportare l'identificazione degli altri componenti, aumentando il numero di violazioni (Biggeri *et al.*, 1996). L'estensione della metodologia adottata al caso delle unità con struttura gerarchica rivelò un rischio troppo elevato per procedere al rilascio. Si dovette quindi definire un nuovo approccio nel quale il rischio rappresentasse un attributo stimato in base alle caratteristiche proprie di ciascuna unità presente nel file (Benedetti e Franconi, 1998). La procedura apprestata è stata uti-

---

<sup>1</sup> Ringrazio Luisa Franconi e Giovanni Serì per avermi fornito la documentazione e per i numerosi suggerimenti.

lizzata a partire dal 2000 per la produzione dei file standard dell'Istituto. L'introduzione di una misura del rischio individuale ha permesso di limitare gli interventi di protezione ai soli record per i quali esso risultasse superiore ad una soglia stabilita, definita in generale in funzione della natura dei dati e di altri fattori fra i quali la numerosità campionaria e la disponibilità e qualità di archivi nominativi esterni.<sup>2</sup> L'intera problematica è approfondita e la metodologia è descritta in modo dettagliato in Istat, 2004. Si può utilmente consultare anche Franconi e Poletti, 2004. Ulteriori sviluppi sono presentati in Poletti e Stander, 2004.

La tabella 1 riporta, per gli anni fra il 1994 e il 2003, il numero di file standard rilasciati dall'Istat per le diverse indagini. Via via, la maggiore disponibilità e qualità dei file ha contribuito a fare aumentare il numero di richieste e di rilasci. Le richieste provenienti dall'università risultano prevalenti (circa 55% nel periodo considerato) rispetto al restante settore pubblico (circa 25%) e al settore privato (circa 20%).

TABELLA 1  
*Numero di file standard rilasciati dall'Istat per indagine  
(anni 1994-2003)*

| INDAGINE   | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | TOTALE |
|--|------|------|------|------|------|------|------|------|------|------|--------|
| Censimento della popolazione 1991                                    |      |      | 1    |      |      | 1    |      | 3    |      | 2    | 7      |
| Forze di lavoro  | 3    | 1    | 8    | 4    | 7    | 8    | 12   | 18   | 17   | 26   | 104    |
| Consumi delle famiglie   | 6    | 15   | 11   | 26   | 18   | 16   | 22   | 33   | 38   | 27   | 212    |
| Indagini multiscopo sulle famiglie (87-91)                           | 28   | 5    | 14   | 4    | 7    | 5    | 1    | 1    | 3    | 2    | 70     |
| Aspetti della vita quotidiana  |      |      | 11   | 14   | 15   | 20   | 18   | 30   | 16   | 22   | 146    |
| Condizioni di salute e ricorso ai servizi sanitari                   |      |      |      |      | 11   | 5    | 5    | 5    | 17   | 14   | 57     |
| Sicurezza dei cittadini  |      |      |      |      |      | 2    | 3    | 1    | 2    | 1    | 9      |
| Tempo libero e cultura   |      |      |      |      |      |      | 1    | 4    | 1    | 8    | 14     |
| Viaggi in Italia e all'estero  |      |      |      |      |      | 3    | 2    | 4    | 3    | 2    | 14     |
| Famiglia e soggetti sociali  |      |      |      |      |      |      |      | 10   | 7    | 7    | 24     |
| Indagini sull'inserimento professionale dei laureati e dei diplomati |      | 1    |      | 6    |      | 7    | 6    | 10   | 6    | 16   | 52     |
| Panel europeo sulle condizioni delle famiglie                        |      |      |      |      |      | 2    | 2    | 1    | 2    |      | 7      |
| TOTALE   | 37   | 22   | 45   | 54   | 58   | 69   | 72   | 120  | 112  | 127  | 716    |

<sup>2</sup> Sugli aspetti della tutela della riservatezza sono state avviate importanti iniziative internazionali alle quali l'Istat ha partecipato e partecipa, come ad esempio i progetti europei Statistical disclosure control e Computational aspects of statistical confidentiality, nell'ambito rispettivamente del 4<sup>th</sup> e 5<sup>th</sup> Framework program della Commissione europea.

#### 4. DATI RELATIVI ALLE IMPRESE

Per i dati riferiti alle imprese il problema è ancora più delicato poiché, in questo caso, la minore variabilità dei caratteri osservati rende più facilmente identificabili le unità statistiche; inoltre, la violazione della riservatezza può avere conseguenze anche sulla concorrenza fra soggetti.

Le tecniche proposte in letteratura per la protezione dei dati di impresa sono di tipo perturbativo, ossia prevedono la produzione di file di unità statistiche fittizie i cui record non siano riconducibili ai dati originali, mantenendo al contempo, nella misura massima possibile, le caratteristiche dei dati originali dal punto di vista della capacità di interpretazione statistica (Cox, 1995).

Un possibile approccio è stato suggerito da Anwar, 1993 (si veda anche Defays e Anwar, 1998) e consiste nel produrre dati microaggregati, cioè pseudo-record ottenuti sintetizzando (normalmente mediante la media aritmetica) i valori di un gruppo di unità statistiche che presentino caratteristiche simili. Quanto più le imprese nel microgruppo esibiscano valori simili per le variabili di interesse, tanto minore risulterà la distorsione introdotta attraverso l'aggregazione. Di regola, si utilizza la stessa soglia adottata per il rilascio di dati in forma tabellare. In questo caso, come è noto, il numero minimo di unità statistiche in ciascun gruppo è pari almeno a tre. L'efficacia della tecnica descritta è notevole se la sintesi viene effettuata su tutte le variabili contemporaneamente. Il rischio di violazione è dunque contrastato prima dell'analisi statistica. Tuttavia, come si comprende, parte della variabilità viene perduta.

Nonostante le riserve degli utilizzatori, le tecniche di microaggregazione continuano a destare l'interesse degli istituti nazionali di statistica, poiché aprono una possibilità concreta di fornire dati trattabili sulle imprese. L'Istat ha condotto approfondimenti specifici su tali tecniche (Contini *et al.*, 1998, Corsini *et al.*, 1999) e ha prodotto sperimentalmente, mediante un'applicazione sviluppata internamente (Pagliuca e Seri, 2001), dati microaggregati per l'indagine sui conti economici delle imprese relativi agli anni 1995 e 1996 (Istat, 2001). La pubblicazione è stata corredata da informazioni utili per valutare la qualità statistica dei file. In particolare, sono state fornite le matrici di correlazione e di varianze e covarianze computate sui dati originali e sui dati microaggregati; inoltre, è stato fornito un indice di qualità, espresso come frazione di variabilità mantenuta dai dati microaggregati rispetto a quella riferita ai dati originali, per ciascuna variabile e per ciascuna sottopopolazione considerata nel processo di aggregazione.

L'accoglienza da parte degli utenti è stata tiepida. Compresi gli abbonamenti, sono state distribuite in tutto 323 copie della pubblicazione, di cui 60 gratuitamente.

Ulteriori sviluppi, dal punto di vista sia metodologico sia delle applicazioni software, sono stati prodotti nell'ambito del progetto europeo Computational aspects of statistical confidentiality, già richiamato alla nota 2. Recenti contributi di approfondimento della problematica relativa alla protezione dei dati relativi alle imprese sono in Franconi e Stander (2002 e 2003) e in Poletti (2003). Tuttavia, una metodologia consolidata ed accettata favorevolmente dagli utilizzatori riguardo alla produzione di file di dati elementari per le imprese non è ancora disponibile.

## 5. IL LABORATORIO ADELE

Il laboratorio è stato istituito nel 1998 e ha cominciato ad operare nel 1999. Inizialmente, vi hanno lavorato utenti “amichevoli” in modo da poter testare la funzionalità del servizio predisposto. Il 1° luglio 1999 il laboratorio è stato presentato ufficialmente. La registrazione delle richieste pervenute e delle elaborazioni svolte è cominciata a fine novembre.

Nella tabella 2 sono sintetizzati i dati relativi all’uso del laboratorio, dal 29 novembre 1999 al 30 aprile 2004.

Come si vede, non tutte le richieste si concludono con la realizzazione dell’elaborazione. Tuttavia, soltanto in un caso non è stato possibile procedere all’elaborazione perché i dati erano non validati o non disponibili. Se la risposta fornita all’utente non porta a sviluppi ulteriori, il contatto viene archiviato come “richiesta di informazioni”. In definitiva, le elaborazioni effettivamente eseguite sono state poco più della metà dei contatti.

Se si considera l’andamento temporale delle richieste pervenute (tabella 3), si conclude che l’attività del laboratorio non ha ancora avuto lo sviluppo augurabile. Ci si aspettava, infatti, anche in base all’esperienza di altri istituti nazionali di statistica e in particolare di quello olandese, che si registrasse un aumento consistente della domanda. La limitata frequenza è riconducibile alla scarsa conoscenza che si ha dell’iniziativa sia all’interno dell’Istat sia all’esterno. Lo si può dedurre dalle risposte fornite nel questionario di valutazione alla domanda “Come è venuto a conoscenza del laboratorio?”. Il 60% degli utenti ha risposto “Direttamente o indirettamente, tramite persone che lavorano all’Istat” e soltanto il 17% “Tramite il sito internet”. Hanno quindi funzionato essenzialmente i canali tradizionali e il “passa parola” fra gli utenti; poco è circolata, invece, l’informazione via web.

TABELLA 2

*Utenti del laboratorio Adele per esito della richiesta  
Novembre 1999 - aprile 2004*

| Esiti delle richieste                                      | Utenti |
|--|--------|
| Esecuzione dell’elaborazione                               | 50     |
| Fornitura di informazioni                                  | 37     |
| Richiesta non ben definita                                 | 4      |
| Richiesta avanzata per dati non validati o non disponibili | 1      |
| Rinuncia non motivata                                      | 1      |
| Soggetto non autorizzabile                                 | 5      |
| Totale   | 98     |

TABELLA 3

*Richieste di elaborazione soddisfatte presso il laboratorio Adele  
Novembre 1999 - aprile 2004*

| Anni                | Richieste di elaborazione | Elaborazioni effettuate |
|---------------------|---------------------------|-------------------------|
| 1999 (un mese)      | 2                         | 2                       |
| 2000                | 11                        | 9                       |
| 2001                | 17                        | 11                      |
| 2002                | 35                        | 11                      |
| 2003                | 28                        | 13                      |
| 2004 (quattro mesi) | 5                         | 4                       |
| Totale              | 98                        | 50                      |

Forse, le iniziative “promozionali” per l’uso del laboratorio (informazione in occasione di incontri scientifici e di corsi sulla tutela della riservatezza; notizia sul sito dell’Istat nella sezione dei servizi per i ricercatori) non sono state abbastanza convincenti, nel timore di non riuscire a sostenere un aumento troppo elevato e veloce della domanda.

La tabella 4 riporta i dati sullo stato delle elaborazioni richieste, mostrando un buon funzionamento del laboratorio.

Nella Tabella 5 gli utenti sono classificati secondo l’organizzazione scientifica di appartenenza. La maggior parte viene dalle università. D’altro canto, quasi tre quarti delle richieste sono pervenuti da soggetti che operano a Roma, a motivo della difficoltà di lavorare presso la sede centrale dell’Istat, eventualmente per tempi prolungati, da parte di ricercatori provenienti da altri luoghi del paese. Una soluzione convincente per ovviare a queste difficoltà è quella di sviluppare soluzioni on line sotto forma di laboratori virtuali, dove gli utenti possano effettuare elaborazioni in accesso remoto, con criteri di sicurezza equivalenti a quelli attuali. Il Dipartimento per le politiche di sviluppo del Ministero dell’economia e delle finanze potrebbe forse orientare nella direzione precedente una parte delle risorse rese disponibili dall’Unione europea, contribuendo a creare un’infrastruttura di ricerca di elevato interesse.

TABELLA 4

*Stato delle elaborazioni effettuate presso il laboratorio Adele  
Novembre 1999 - aprile 2004*

| Stato delle elaborazioni     | Numero |
|------------------------------|--------|
| Concluse positivamente       | 42     |
| In corso di esecuzione       | 2      |
| In attesa dei dati richiesti | 4      |
| In corso di valutazione      | 2      |
| Rinunce                      | 3      |
| Totale                       | 53     |

TABELLA 5

*Richieste di elaborazione pervenute al laboratorio Adele per tipo di organizzazione richiedente  
Novembre 1999 - aprile 2004*

| Organizzazioni richiedenti           | Numero |
|--------------------------------------|--------|
| Università pubbliche                 | 23     |
| Università private                   | 6      |
| Cnr                                  | 2      |
| Altri istituti di ricerca pubblici   | 5      |
| Altre amministrazioni pubbliche      | 9      |
| Banca d’Italia                       | 6      |
| Istituti e centri di ricerca privati | 2      |
| Totale                               | 53     |

Le richieste di elaborazione riguardano prevalentemente indagini su famiglie e individui rispetto a quelle sulle imprese; inoltre, per tre di esse (forze di lavoro, censimenti della popolazione e sbocchi professionali dei laureati), sono disponibili anche i file standard. Questa apparente contraddizione non deve stupire perché la possibilità di lavorare sull’intero file del censimento o dell’indagine sulle forze di

lavoro offre opportunità di analisi di gran lunga migliori rispetto a quelle possibili attraverso l'uso dei file standard i quali, come accennato in precedenza, subiscono una riduzione del contenuto informativo prima del rilascio. In particolare, il file relativo all'indagine sugli sbocchi professionali dei laureati presenta una significativa perdita di informazione rispetto ai dati originali. Non si capisce, invece, il motivo della limitata richiesta di elaborazioni su dati relativi alle imprese (soltanto il 20% del totale).

Il servizio ottenuto presso il laboratorio è valutato, in generale, molto soddisfacente dagli utenti.

## 6. PROGETTI DI RICERCA COMUNI

Le convenzioni fra l'Istat e soggetti esterni per lo sviluppo di progetti comuni di ricerca risultano al momento estremamente limitate in numero. Si tratta di una fattispecie che vede l'amministrazione dell'Istituto molto timida, poiché timorosa delle possibili conseguenze sul piano delle responsabilità personali. Non hanno aiutato in questa direzione l'iniziale incertezza sull'interpretazione della normativa relativa alla tutela della riservatezza e un atteggiamento del Garante generalmente abbastanza rigido.

I due codici deontologici che sono stati definiti potranno consentire un'estensione delle convenzioni di ricerca, se l'applicazione sarà intelligente e non invece riduttiva o, ancora peggio, burocratica.

Dei progetti potranno giovare sia i soggetti del Sistema statistico nazionale e tipicamente l'Istat, sia i ricercatori esterni. Potrà essere aumentata la capacità di analisi dei dati e per questa via migliorata la produzione.

## 7. CONCLUSIONE

In precedenza, mi sono soffermato essenzialmente su quanto è stato fatto dall'Istat per consentire l'accesso ai dati individuali da parte di terzi, prima della legge n. 675 del 31 dicembre 1996. Si tratta di aspetti che la riflessione promossa dalla rivista avrebbe potuto trascurare. Spero che la testimonianza portata, con il sostegno documentale dei colleghi dell'Istat, possa essere di qualche utilità.

Svolgo ora qualche riflessione conclusiva. I due codici riconoscono, il primo, il rilevante interesse pubblico del trattamento di dati personali da parte degli uffici del Sistema statistico nazionale, in quanto "fonte dell'informazione statistica ufficiale intesa quale patrimonio della collettività"; il secondo, la funzione sociale della ricerca scientifica e quindi possibilità di trattamento di dati personali più estese di quelle generalmente consentite. I due riferimenti si coniugano, nel primo caso, con le norme generali di comportamento degli operatori pubblici e con quelle particolari previste per gli operatori statistici pubblici, nel secondo caso, con la presunzione di buona fede dei soggetti privati aventi fini di ricerca, riconoscendo ad essi un credito di riservatezza. Su queste basi possono essere costruite prassi efficaci.

Si deve cogliere l'occasione per promuovere e sostenere forme di collaborazione più intensa fra la statistica pubblica e i soggetti esterni. Comunque, si deve rendere effettiva la possibilità di trattamento di dati personali da parte di ricercatori privati, attraverso investimenti in primo luogo dell'Istat (Zuliani, 2001) ed eventualmente, come ho già accennato, di altri soggetti pubblici.

Dipartimento di studi geoeconomici,  
linguistici, statistici, storici per l'analisi regionale  
Università di Roma "La Sapienza"

ALBERTO ZULIANI

#### BIBLIOGRAFIA

- M. N. ANWAR (1993), *Microaggregation: the small aggregates method*, Internal report, Eurostat, Luxembourg.
- R. BENEDETTI, L. FRANCONI (1998), *Statistical and technological solutions for controlled data dissemination*, "Pre-proceedings of new techniques and technologies for statistics", vol. 1, 225-232.
- L. BIGGERI, F. ZANNELLA (1991), *Release of microdata and statistical disclosure control in the new national system of Italy: main problems, some technical solutions, experiments*, "Bulletin of the International Statistical Institute, Proceedings", tomo LIV, vol. 1, 1-25.
- L. BIGGERI, M. ANGRISANI, L. BUZZIGOLI, A. GIOMMI, A. GIUSTI, L. GRASSINI, G. MARLIANI (1996), *Rapporto sull'attività svolta dal gruppo di lavoro su "Tutela della riservatezza e identificazione dei rispondenti alle rilevazioni statistiche svolte nell'ambito del Sistan"*, Commissione per la garanzia dell'informazione statistica, Roma.
- G. COCCIA (1993), *Disclosure risk in Italian current population surveys*, "International seminar on statistical confidentiality, Proceedings", Dublin, 1992. Office for the official publications of the European Communities, Luxembourg, 415-423.
- B. CONTINI, V. CORSINI, L. FRANCONI, D. PAGLIUCA, P. PAPA, F. PIERSIMONI, G. SERI, G. SIESTO, P. TACCINI (1998), *Metodi di microaggregazione per il rilascio di dati di impresa*, "Documenti Istat", n. 17.
- V. CORSINI, L. FRANCONI, D. PAGLIUCA, G. SERI (1999), *An application of microaggregation methods to Italian business survey*, "Statistical data protection, Proceedings", Lisbon, 1998. Eurostat, Luxembourg, 109-113.
- L. H. COX (1995), *Protecting confidentiality in business surveys*, in B. G. Cox, D. A. Binder, B. N. Chinnappa, A. Christianson, M. J. Colledge, P. S. Kott (Eds.), "Business survey methods", Wiley, New-York, 443-476.
- F. CRESCENZI (1993), *Estimating population uniques: methodological proposals and applications on Italian census data*, "International seminar on statistical confidentiality, Proceedings", Dublin, 1992. Office for the official publications of the European Communities, Luxembourg, 247-260.
- D. DEFAYS, M. N. ANWAR (1998), *Masking microdata using micro-aggregation*, "Journal of official statistics", 14, 449-461.
- L. FRANCONI, S. POLETTINI (2004), *Individual risk estimation in mu-Argus: a review*, in J. Domingo-Ferrer, V. Torra (Eds.), "Privacy in statistical databases", LNCS 3050, Springer-Verlag, Berlin, 262-272.
- L. FRANCONI, J. STANDER (2002), *A model based method for disclosure limitation of business microdata*, "Journal of the Royal Statistical Society", D, 51, 1-11.
- L. FRANCONI, J. STANDER (2003), *Spatial and non-spatial model-based protection procedures for the release of business microdata*, "Statistics and computing", 13, 295-305.

- ISTAT (2001), *Microaggregazione dei dati economici strutturali delle imprese industriali e dei servizi - Anni 1995-1996*, "Informazioni", n. 34.
- ISTAT (2004), *Metodologie e tecniche di tutela della riservatezza nel rilascio di informazione statistica*, "Metodi e norme", in corso di stampa.
- D. PAGLIUCA, G. SERI (2001), *Dissemination of business data: MASQ a software for single axis micro-aggregation of quantitative variables*, "Research in official statistics", 2, 73-82.
- S. POLETTINI (2003), *Maximum entropy simulation for microdata protection*, "Statistics and computing", 13, 307-320.
- S. POLETTINI, J. STANDER (2004), *A bayesian hierarchical model approach to risk estimation in statistical disclosure limitation*, in J. Domingo-Ferrer, V. Torra (Eds.), "Privacy in statistical databases", LNCS 3050. Springer-Verlag, Berlin, 247-261.
- G. SERI (2004), *Adele. Il Laboratorio per l'analisi dei dati elementari*, sottoposto per la pubblicazione nella collana "Documenti Istat".
- U. TRIVELLATO (2003), *Protezione dei dati personali e ricerca scientifica*, "Statistica", 4, pp. 627-648.
- A. ZULIANI (2001), *Innovazione tecnologica e informazione statistica*, "Atti della quinta conferenza nazionale di statistica", Istat, Roma.
- A. ZULIANI (2002), *Autorità e information technology. Costanza d'impegno per la cultura dell'innovazione*, "www.aipa.it", 2, 2.

#### RIASSUNTO

*I codici deontologici per il trattamento di dati personali a scopi statistici e di ricerca scientifica: qualche riflessione*

Sono svolte alcune considerazioni sui due codici deontologici per il trattamento dei dati personali a scopi statistici e di ricerca riferiti ai soggetti che operano all'interno del Sistema statistico nazionale e rispettivamente ai soggetti privati. Ci si sofferma in particolare sulle modalità di accesso dei privati ai dati prodotti dalla statistica pubblica e in particolare dall'Istituto nazionale di statistica. I due codici dettano regole ragionevoli per l'utilizzazione statistica e scientifica dei dati personali. Occorrono però ancora un forte impegno e conseguenti investimenti pubblici per la produzione e l'accesso a data set che consentano analisi significative.

#### SUMMARY

*Codes of conduct in processing personal data for statistical and scientific purposes: some considerations*

Two codes of conduct are here treated pertaining to the use of personal data for statistical and research purposes with reference respectively to the users who operate within the National Statistical System and private users. Attention is given to the modality of access on the part of private users to public statistics, and especially those issued by the National Institute of Statistics. The two codes of conduct described provide flexible rules for the statistical and scientific utilization of personal data. However, considerable efforts are required, with attendant public investment, for the production and accessing of the data sets necessary for significant analyses.