

## PROBLEMI DI PRIVACY NELLA SELEZIONE DI PICCOLI CAMPIONI

Loredana Cerbara

## 1. INTRODUZIONE

Il problema del rispetto della legge sulla privacy nella selezione di piccoli campioni è piuttosto rilevante perché oggi non sembra pensabile di poter fare a meno di indagini campionarie per supportare la ricerca sociale. Esse sono molto utilizzate, per l'immediatezza con cui forniscono dati, in relazione a molti fenomeni sociali contingenti o per il monitoraggio di questioni ritenute socialmente importanti. D'altra parte, anche l'istituto centrale di statistica italiano (ISTAT) e molti organismi di rilievo internazionale come l'EUROSTAT hanno tra i loro fini istituzionali l'utilizzo di queste tecniche per scopi conoscitivi e interpretativi della realtà contemporanea.

Tuttavia negli ultimi anni si è forse un poco abusato di queste tecniche di campionamento e sono proliferati sulla stampa e su gli altri media i risultati di diversi sondaggi senza il necessario fondamento statistico, o, perlomeno, senza che si parlasse a sufficienza della loro affidabilità. Per questo l'Autorità Garante delle Comunicazioni (<http://www.agcom.it/>) ha introdotto una serie di norme<sup>1</sup> che riguardano la pubblicazione dei risultati di sondaggi per limitare l'abuso di certe pratiche che avevano l'effetto di diffondere informazioni di dubbio reperimento con tutte le conseguenze negative che possono verificarsi<sup>2</sup>.

Ma a parte il ricorso a indagini a fine non scientifico, le tecniche di campionamento sono largamente usate per la produzione di informazioni a supporto dell'economia e della politica, al punto che moltissimi progetti di ricerca anche a livello internazionale e trasversalmente a più Paesi, soprattutto in ambito economico e sociale, raramente prescindono da informazioni prodotte tramite sondag-

---

<sup>1</sup> Delibera n. 153/02CSP modificata dalla delibera n. 237/03/CSP, [www.agcom.it/sondaggi/sondaggi.htm](http://www.agcom.it/sondaggi/sondaggi.htm). A onor del vero, con questa delibera si chiede di esplicitare il margine di errore al livello di fiducia del 95% anche quando i campioni non sono probabilistici, e ciò è quantomeno illogico. Tuttavia, in questo modo si costringe chi diffonde i dati di un sondaggio a dare anche informazioni sulla qualità degli stessi attraverso una serie di indicatori che possono essere di aiuto per comprendere la fondatezza metodologica di un sondaggio senza poterla però garantire.

<sup>2</sup> Ad esempio si spaccia per sondaggio anche l'approccio di certe vendite effettuate telefonicamente in modo che chi risponde si ben disponga verso l'operatore che lo chiama e sia maggiore la probabilità di concludere la vendita.

gi. Molto spesso le statistiche ufficiali, pur molto ricche specialmente nei Paesi più avanzati in cui il flusso di informazioni è garantito ed è assicurato per tutte le attività di primaria importanza, non sono sufficienti per particolari scopi conoscitivi. In questi casi le indagini campionarie rappresentano il necessario complemento per il reperimento delle informazioni mancanti. È chiaro perciò che la ricerca scientifica in ambito sociale non può più fare a meno di questo strumento di indagine e che si rende necessario un intervento legislativo per migliorare la disponibilità di liste ai fini del campionamento scientifico.

Un intero numero della rivista 'Statistica' (2003, 4) dal titolo '*I microdati, tra scienza e privacy*' è stato dedicato al problema del trattamento dei dati personali per scopi di ricerca scientifica. Fra i vari articoli, tutti di rilievo e di grande attualità, uno in particolare scritto da Schizzerotto (2003) descrive i problemi campionari che si verificano quando per la selezione dei campioni si fa riferimento alle liste di microdati in possesso della Pubblica Amministrazione, liste che sono generalmente inaccessibili per effetto della legge sulla privacy (Decreto legislativo 30 giugno 2003, n. 196) oltre che della legge della trasparenza della pubblica amministrazione (l. 241/1990). Ma la legge sulla privacy "...non riconosce al ricercatore alcun potere coercitivo nei confronti di chi si rifiutasse ancora di trasmettere l'elenco dei nominativi contenuto in archivi pubblici. Essa rende, però, assai meno agevole, di quanto non fosse in passato, la legittimazione di eventuali dinieghi..." (Schizzerotto, 2003). "Ma la produzione e distribuzione di adeguate basi di microdati per la ricerca è ... il prossimo non eludibile impegno che attende le società scientifiche e l'intera comunità scientifica." (Trivellato, 2003).

A tutt'oggi, e dopo la pubblicazione del Codice Deontologico della Statistica (G.U. n. 190 del 14 agosto 2004)<sup>3</sup>, le uniche liste *pubbliche* utilizzabili per la selezione dei campioni a scopi di ricerca scientifica sono le liste elettorali. Vedremo nel seguito come anche queste liste possano creare problemi soprattutto per la selezione di campioni di piccole dimensioni. Anche le liste che chiamiamo *ad hoc*, di cui si parlerà nel seguito, possono essere usate per la selezione dei campioni, che generalmente però risultano distorti e quindi non utilizzabili ai fini della ricerca scientifica.

Nel volume citato non si fa però menzione della possibilità, ampiamente utilizzata fino ad oggi, di estrarre campioni a partire dagli elenchi telefonici. Il vantaggio che questo tipo di lista presenta risiede soprattutto nel fatto che si tratta di liste 'pubbliche', già disponibili per tutti e che garantiscono una certa copertura della popolazione presente. La legge però impone di utilizzare gli elenchi telefonici solo per i 'contatti interpersonali' e non per altri scopi. In questo lavoro, a partire da considerazioni di ordine pratico, si vuole evidenziare l'utilità dei sondaggi ef-

---

<sup>3</sup> È a tutti gli effetti un testo di legge, anche se non era obbligatorio, secondo la direttiva europea e le raccomandazioni internazionali. È stato sottoscritto dalle società scientifiche e dalla crui (Conferenza dei Rettori delle Università Italiane) e, successivamente dall'assirm (Associazione tra Istituti di Ricerche di Mercato, Sondaggi di Opinione, Ricerca Sociale) e disciplina il trattamento dei dati personali a scopi statistici e di ricerca scientifica, che la legge 675/96 e poi 196/2003 avevano lasciato in deroga (art. 1, lettera a) l. 675/96) in attesa, appunto di un codice specifico della materia.

fettuati con l'ausilio di sistemi automatici a partire dalle liste telefoniche ponendo l'accento sulla necessità di selezionare campioni a partire da queste liste.

## 2. VANTAGGI E SVANTAGGI DELLE TECNICHE DI SONDAGGIO ASSISTITE DA COMPUTER

Una delle tecniche più utilizzate per i sondaggi nell'ambito della ricerca sociale è quella del sistema CATI<sup>4</sup>. Essa prevede la somministrazione telefonica di un questionario da parte di un intervistatore, il quale inserisce immediatamente le risposte ottenute in un terminale che è di supporto sia perché può essere dotato di software di controllo della congruità delle risposte inserite, sia perché i dati sono immediatamente disponibili in formato digitale, pronti per le elaborazioni successive (per ulteriori dettagli operativi si veda Brusati (2003), oppure Fabbris (2000). Tutta la famiglia dei CAI<sup>5</sup>, accomunati dal fatto che si tratta di modalità di somministrazione dei questionari 'assistita da computer' ha avuto un enorme sviluppo negli ultimi tempi e deve la propria fortuna alle caratteristiche di affidabilità (abbattono l'errore non probabilistico) e di immediatezza (i dati sono immediatamente disponibili).

Naturalmente il CATI è un sistema non esente da critiche metodologiche, prima fra tutte il fatto che non consente la produzione di sondaggi probabilistici, se no in casi particolari in cui si ha a disposizione una lista completa di contatti.

Con opportune tecniche di bilanciamento delle modalità di approccio alla rilevazione campionaria<sup>6</sup> e con tecniche di ponderazione a posteriori dei dati, l'affidabilità di questi campioni non è lontana da quella di campioni fatti con il metodo tradizionale (tipo quelli effettuati dall'ISTAT) e le differenze in termini di affidabilità dipendono essenzialmente dalla dimensione campionaria.

Certo è che una indagine telefonica *non è probabilistica in senso stretto*. L'uso delle tecniche di correzione di queste imperfezioni rende però queste metodiche di indagine molto più affidabili (anche se questa affidabilità non è statisticamente quantificabile) di quanto si pensi, e per questo molto appetibili. In presenza di risorse economiche limitate esse rappresentano la sola via percorribile per effettuare un sondaggio. In verità si effettuano interviste telefoniche soprattutto quando si vuole fare uso di *piccoli campioni*, mentre le grandi indagini vengono progettate con maggiore rigore scientifico e con l'uso di risorse economiche molto più consistenti.

---

<sup>4</sup> CATI è l'acronimo di *Computer Assisted Telephon Interviewing*.

<sup>5</sup> CAI sta per *Computer Assisted Interviewing*

<sup>6</sup> Il problema principale in questi casi è relativo al fatto che occorre trovare gli intervistandi a casa, disponibili a rispondere al telefono. Perciò è molto frequente che si intervistino casalinghe o pensionati ed è più difficile intervistare coloro che lasciano l'abitazione per lavoro o studio. Allora si effettuano le telefonate in orari serali, si conducono le interviste anche nei giorni non lavorativi o si ripetono più tentativi durante il giorno in modo da coprire tutte fasce orarie di possibile presenza a casa di chi deve essere intervistato. Inoltre si stabiliscono delle quote per le quali si cerca intervistare un numero minimo di persone che presentano determinate caratteristiche: ad esempio si può scegliere di avere una composizione finale del campione costituita da metà maschi e metà femmine, o un numero di persone anziane proporzionale alla popolazione residente anziana, e così via.

Un altro limite delle indagini campionarie telefoniche è riconducibile al fatto che le domande devono essere chiaramente formulate e i questionari devono essere piuttosto brevi, perché le interviste che durano molto rischiano maggiormente di venire troncate prima della fine. Per questo motivo i sondaggi telefonici sono riservati a tutte le consultazioni snelle e per le quali si chiede immediata disponibilità di informazione.

Si capisce quindi come sia vasto l'ambito di applicazione, in un contesto tecnologico e sociale, sia nazionale che internazionale, in cui le informazioni diventano obsolete in breve tempo e necessitano di continui e rapidi aggiornamenti, specialmente per questioni legate all'efficienza e all'efficacia di azioni effettuate su larga scala (come le scelte di politica sociale, ad esempio, ma anche in molti altri settori). Le imprese private hanno avuto immediato sentore di queste potenzialità e da anni fanno uso di queste tecniche di reperimento delle informazioni a corredo di più importanti operazioni di acquisizione dati per studi di mercato.

In definitiva, nel settore delle indagini campionarie ci sono casi per cui si fanno scelte che possono apparire non rigorose. In effetti la teoria del campionamento aiuta a progettare campioni efficienti consentendo il calcolo della probabilità di errore e quindi dando al ricercatore la possibilità di realizzare disegni campionari affetti dal minor errore di campionamento possibile. Ma quando i vincoli di budget e quelli di ordine temporale impongono delle forti limitazioni, si opta per campioni non probabilistici (come le indagini telefoniche) e si fa ricorso ad aspetti diversi della teoria per minimizzare gli errori. In effetti sono proprio gli errori non campionari, provocati da tutte le possibili imprecisioni e inaccuratezze commesse o subite durante un'indagine, a costituire molto spesso la parte più rilevante dell'errore totale, cosicché è molto importante che siano questi ultimi ad essere controllati e limitati oltre che l'errore di campionamento. Così non è raro che si opti per sondaggi telefonici con una stratificazione territoriale più estesa possibile o caratterizzati da una numerosità campionaria più elevata possibile, sempre entro i limiti che le risorse economiche e tecnologiche (e, si potrebbe aggiungere, normative) consentono. Queste caratteristiche affiancate dal fatto di poter limitare al minimo le mancate risposte oppure gli errori di imputazione dei dati, rendono i sondaggi telefonici una delle tecniche di rilevazione campionaria tra le più vantaggiose.

### 3. ASPETTI DELLA LEGGE SULLA PRIVACY

Da un decennio circa la tematica della tutela della privacy è emersa fortemente anche in Italia, anzi, forse anche a causa del ritardo con cui il nostro Paese l'ha affrontata - dal 1996 con vari interventi successivi fino al 2003 - lo sviluppo di normative nel settore è stato uno dei più prolifici. L'Italia, da ultima della lista in Europa, è diventata leader in tema di privacy perché la normativa vigente è tanto ricca di dettagli da risultare caso unico in un contesto internazionale. Il Codice della Privacy è oggi un Testo Unico, cioè una raccolta monotematica di tutte le norme vigenti sulla tutela dei dati personali. Ad esso sono allegati vari codici de-

ontologici, tra cui il già citato Codice di deontologia e di buona condotta per i trattamenti di dati personali per scopi statistici e scientifici che chiarisce come alla ricerca scientifica e al trattamento statistico dei dati sia riservata dalla normativa vigente tutta una serie di regolamentazioni specifiche. In particolare si asserisce che se il trattamento dei dati personali è funzionale ad un progetto di ricerca approvato dagli enti competenti non sono dovuti neanche una serie di adempimenti obbligatori in altri ambiti (ad esempio il consenso), pur rimanendo l'obbligo dell'*informativa* (che deve contenere accenni agli scopi della raccolta dei dati) e di altre restrizioni in sede di diffusione dei risultati della ricerca. Si garantisce cioè un *adeguato* trattamento dei dati in modo che sia quantomeno molto difficile risalire al proprietario dei dati personali. La garanzia che ciò avvenga è assicurata dalla professionalità di chi effettua la ricerca in quanto costituiscono deroga alla legge della privacy in particolare le *ricerche di rilevante interesse pubblico*<sup>7</sup> che generalmente sono affidate a professionisti della statistica.

Naturalmente, in ogni caso, di uso lecito o illecito dei dati personali, restano salvi i diritti sanciti dall'art. 7 del codice sulla privacy (il già citato D.L. 196/2003) per i quali ogni individuo può chiedere l'indicazione sull'origine dei dati, sulle finalità e modalità del trattamento e sul titolare. Inoltre ciascuno ha diritto a chiedere l'aggiornamento, la rettifica, l'integrazione, la cancellazione dei suoi dati personali ed opporsi al trattamento degli stessi anche solo in modo parziale. Perciò, chi non vuole essere oggetto di attenzione ha gli strumenti, sanciti dalla legge, per farlo.

#### 4. PROBLEMI DI PRIVACY NELL'USO DI LISTE PER LA SELEZIONE DEI CAMPIONI

Questa in somma sintesi è la situazione della normativa vigente. Sebbene sembra che non vi sia nulla da eccepire, ci sono casi in cui la tutela della privacy rende meno semplici certe procedure di accesso ai dati, fonte primaria di informazione per gli statistici o gli studiosi e i politici, e perfino inattuabili certe procedure consolidate che solo in apparenza violano la privacy di chi viene coinvolto nella ricerca. In effetti è diventata più onerosa, e in qualche caso impossibile, la realizzazione di campioni probabilistici, o anche di campioni non probabilistici ma ad alto potere di rappresentatività statistica, come quelli che si possono ottenere potendo disporre liberamente dei numeri di telefono delle famiglie italiane. La normativa vigente, però, indica esplicitamente tre sole possibilità per la selezione dei campioni da parte di chi volesse effettuare una indagine campionaria:

- le *liste ad hoc*, una collezione di nominativi di individui da cui si è ottenuto un consenso (scritto o soltanto orale a seconda dei casi regolamentati dal codice vigente sulla privacy);

---

<sup>7</sup> Secondo l'art. 22 d.leg. n. 135 del 1999, «si considerano di rilevante interesse pubblico i trattamenti svolti dai soggetti pubblici che fanno parte del sistema statistico nazionale ai sensi del decreto legislativo 6 settembre 1989, n. 322».

- i nominativi di chi ha dato il consenso alle compagnie telefoniche<sup>8</sup> (di cui si parlerà più ampiamente nel seguito);
- le liste elettorali, che rappresentano la sola fonte utilizzabile in assenza di liste specifiche e sono liberamente consultabili per l'estrazione dei campioni (oltre che per le altre attività, tra le quali la pubblicità elettorale) da parte degli enti afferenti al SISTAN.

Perciò chiunque, non appartenente al SISTAN o che non si sia rivolto ad un ente del SISTAN per un supporto metodologico ad una ricerca di tipo scientifico, volesse condurre una indagine campionaria, avrebbe solo le prime due possibilità: le liste ad hoc o quelle delle compagnie telefoniche.

Tutte queste liste comportano una serie di inconvenienti all'atto della selezione di campioni a partire da esse.

Infatti nel caso delle liste ad hoc occorre reperire i consensi, con grande dispendio di energie, anche se molte di queste liste sono oggi disponibili sul mercato e si possono acquistare. Esse però non consentono la selezione di campioni rappresentativi della popolazione italiana perché contengono individui dalle caratteristiche particolari non direttamente assimilabili a quelle degli italiani in generale. Perciò da queste liste non si possono che estrarre campioni distorti (ai fini della ricerca sociale a livello nazionale).

È importante sottolineare che anche i campioni estratti a partire dalle liste elettorali presentano alcuni problemi in fase di progettazione e realizzazione di una indagine campionaria. Ad esempio si possono selezionare da queste liste solo individui *che hanno compiuto i 18 anni di età* perché solo questi sono inclusi in tali liste e vale il vincolo di progettare campioni su base territoriale con il conseguente obbligo di utilizzare campioni complessi, sia dal punto di vista della selezione che da quello del calcolo della probabilità di errore. Infatti a partire dalle liste elettorali dei comuni italiani si possono progettare campioni a più stadi. Questo tipo di disegno campionario, tra l'altro, impone un aggravio nei calcoli rispetto ad uno più semplice, sia per ciò che riguarda le stime dei parametri di interesse che per la valutazione della loro affidabilità, dovendo far ricorso a indicatori complessi che tengano conto del disegno campionario adottato.

## 5. USO DI LISTE TELEFONICHE

Le liste di numeri telefonici che normalmente (a parte casi particolari) si hanno a disposizione per un sondaggio con il sistema CATI non offrono la corrispon-

---

<sup>8</sup> Secondo la normativa vigente, solo chi ha dato un esplicito consenso ad essere contattato telefonicamente o per posta per azioni di telemarketing può essere raggiunto da qualunque promozione o iniziativa che usi il mezzo telefonico o postale per raggiungere il suo scopo. Tutti gli abbonati hanno ricevuto, nei mesi di gennaio e febbraio 2005 presso la propria residenza, da parte dei propri operatori di telefonia, un modulo da compilare con le informazioni che vogliono siano inserite nel Data Base Unico, cioè l'elenco telefonico unificato di tutti gli utenti facenti capo ad un qualunque operatore di telefonia fissa o mobile.

denza uno ad uno tra numero telefonico e unità di rilevazione. Ad esempio, nel caso dei numeri telefonici fissi italiani, essi corrispondono di regola ad unità abitative, quindi al massimo si può pensare una coincidenza con i nuclei familiari, ma non certo con singoli intervistandi. Inoltre è chiaro che tutte le famiglie senza telefono fisso non possono essere raggiunte da un sondaggio telefonico che usa le liste residenziali come base per l'estrazione del campione e che le famiglie con più di una linea telefonica sono presenti nella lista più di una volta e dunque hanno una diversa probabilità di essere incluse nel campione.

Tuttavia, anche in presenza di questi limiti, molto spesso si è fatto uso di questa tecnica che garantisce buoni risultati, in quanto il telefono fisso copre la maggior parte delle famiglie italiane; è da rilevare però che nel tempo questa copertura va diminuendo poiché aumentano le famiglie senza telefono fisso e che fanno uso della sola telefonia mobile. In realtà oggi sono previsti elenchi telefonici unificati per tutti gli operatori telefonici e che comprendono sia telefonia fissa che mobile. Questo database potrebbe perciò rivelarsi molto utile per la progettazione di indagini telefoniche perché migliorerebbe la copertura delle famiglie italiane sul territorio nazionale, pur con i conseguenti problemi di duplicazione delle unità da campionare. In realtà, se la maggior parte degli utenti proprietari di un numero telefonico mobile, esprimeranno il consenso ad essere inseriti nell'elenco telefonico unificato, si otterrà un database che può offrire grandi possibilità di realizzazione di campioni affidabili.

Le liste raccolte dalle compagnie telefoniche hanno più o meno gli stessi inconvenienti di quelle ad hoc. Ad oggi risulta che oltre 684.000 utenti hanno dato il consenso alle compagnie telefoniche per il trattamento dei propri dati personali per esprimere la loro volontà di ricevere informazioni commerciali, promozionali o interviste per sondaggi e ricerche di mercato, via posta o via telefono. Un numero piuttosto esiguo se si intende progettare un'indagine campionaria, sia pure di natura non probabilistica, basata su queste liste. Ciò vale anche se si nota il fatto che l'ammontare totale di chi ha dato il consenso con queste modalità a livello regionale è pressoché proporzionale all'ammontare delle famiglie residenti (tavola 1).

Infatti la percentuale di utenti che hanno dato il consenso in base alle liste disponibili presso gli operatori telefonici calcolata sul totale delle famiglie per regione non si supera il valore del 4% ed è in media del 3%. Un valore piuttosto basso che non può mettere al riparo dall'elevata distorsione nelle stime che deriva dal fatto di non aver potuto includere nel campione anche utenti meno disponibili di quelli che sono presenti in queste liste. Ed è anche un valore molto lontano dal 68% che rappresenta la percentuale di utenze telefoniche fisse sul totale delle famiglie italiane. Tale valore è stato calcolato semplicemente considerando che le utenze telefoniche fisse sono circa 15 milioni su un totale di 22 milioni di famiglie residenti in Italia (ISTAT, 2005). Però questa percentuale non può rappresentare esattamente la percentuale di famiglie con telefono fisso, ma ne è al massimo una approssimazione, perché non possiamo stabilire con certezza (come si è già detto) la corrispondenza tra nucleo familiare e utenza telefonica.

TAVOLA 1

*Distribuzione e percentuale per regione degli utenti che hanno dato il consenso al trattamento dei propri dati personali alle compagnie telefoniche, anno 2006*

Regione	Percentuale di consensi rispetto al numero di famiglie residenti
PIEMONTE	3,9
VALLE D'AOSTA	3,4
LOMBARDIA	3,4
TRENTINO ALTO ADIGE	2,8
VENETO	3,6
FRIULI VENEZIA GIULIA	4,2
LIGURIA	3,1
EMILIA ROMAGNA	3,9
TOSCANA	3,6
UMBRIA	4,1
MARCHE	3,4
LAZIO	2,3
ABRUZZO	2,6
MOLISE	2,6
CAMPANIA	2,0
PUGLIA	3,0
BASILICATA	2,6
CALABRIA	1,7
SICILIA	2,3
SARDEGNA	3,0
Totale	3,1

Fonte: Nostre elaborazioni su dati SEAT

I campioni ottenuti invece dalle liste telefoniche generali teoricamente non sono neanche confrontabili con quelli ottenuti a partire dalle liste elettorali, per le enormi differenze che le caratterizzano (*in primis* il fatto dalle liste telefoniche difficilmente si può ottenere un campione probabilistico). Tuttavia essi sono stati largamente utilizzati per la semplicità e l'immediatezza delle fasi progettuali delle ricerche, fino a quando la legge sulla privacy ne ha relegato l'uso al solo contatto interpersonale. Inoltre, fatto non di secondaria importanza e di cui si è già discusso nei paragrafi precedenti, si tratta di inchieste molto economiche rispetto a quelle fatte a partire dalle liste elettorali, che per la sola selezione del campione presuppongono la visita di un incaricato presso i comuni per l'esame delle liste elettorali. Il sistema CATI consente invece di eseguire tutta l'indagine da una postazione fissa, qualunque sia la dislocazione territoriale delle unità da intervistare.

E' anche da sottolineare che negli ultimi anni è mutato l'atteggiamento della popolazione nei riguardi del sondaggio e questo può limitare la realizzazione delle indagini: se la diffusione su larga scala dei sondaggi eseguiti per le più svariate motivazioni ha avuto un impatto iniziale positivo sull'atteggiamento degli intervistandi perché ha contribuito ad eliminare le barriere di diffidenza che potevano in qualche modo influenzare le risposte, oggi al contrario è sempre più frequente il rifiuto a sottoporsi ad una intervista telefonica. Ciò probabilmente per effetto dell'uso smodato di questi strumenti e del conseguente abuso laddove eseguite senza le necessarie basi di rigore scientifico, senza scopi di particolare rilievo e spesso senza l'osservanza delle norme vigenti. Così, contrapposto al vantaggio che tutti oggi sanno cosa sia un *sondaggio*, una *percentuale di risposte* e perfino il *livello*

*di confidenza* non è più un concetto del tutto ignoto, si pone lo svantaggio che i sondaggi sembrano giochetti o manovre pubblicitarie o, peggio, approcci che servono a mascherare altre intenzioni. È ovvio quindi che si vada sempre più diffondendo una crescente diffidenza associata perfino ad una certa *stanchezza*.

Quest'ultimo termine non è inappropriato se si considera che, proprio per effetto della normativa sulla privacy le aziende di telemarketing si sono viste ridurre le liste di persone *contattabili* ed hanno dovuto *scambiarsi* (probabilmente a pagamento!) le liste di contatti disponibili le une con le altre per poter ampliare il proprio bacino di utenza. Così firmando un consenso ad una sola di queste aziende tra loro collegate, quando il consenso include anche il consenso alla cessione dei dati a terzi, si viene inseriti in un circuito per il quale si fa parte delle liste di utenza di molteplici aziende. Per questo anche se non si è dato esplicito consenso a ciascuna di esse, si può comunque venire contattati anche molto di frequente.

## 6. CONCLUSIONI

L'introduzione della normativa sulla tutela della privacy è da considerarsi come un grande passo avanti del progresso civile e sociale in Italia. Il ritardo con cui nel nostro Paese si è arrivati ad affrontare questo tema è ormai storia, dal momento che oggi possediamo uno dei codici in materia di trattamento dei dati personali più all'avanguardia a livello internazionale. Ciò nonostante continuano a porsi, come è naturale per un codice tanto complesso, numerosi problemi di applicazione della legge, ogniqualevolta essa si rivela in contrasto con esigenze contrapposte. Non entriamo qui in questioni relative al bilanciamento di opposti diritti sanciti da leggi di peso equivalente. Quello che si vuole invece mettere in evidenza è il caso della selezione dei campioni per sondaggi di opinione nella ricerca sociale che incontra molti ostacoli che andrebbero eliminati, o quantomeno minimizzati a favore del progresso della conoscenza. In particolare, il settore delle indagini telefoniche, tanto care alle aziende che ne fanno la principale fonte di guadagno per le loro caratteristiche di parsimonia e di efficacia comunicativa, soffre del fatto che occorre tutelare la privacy di chi non vuole essere disturbato per telefono.

Perciò è necessario, oltre che urgente, un intervento legislativo che renda utilizzabili le liste telefoniche, al pari di quelle elettorali già considerate dalla legge vigente, per la selezione di campioni da utilizzare nei sondaggi telefonici a scopi di ricerca scientifica. Questo perché si tratta di una modalità di selezione dei campioni ormai consolidata che consente la disponibilità di molti contatti per arrivare a risultati attendibili e rilevanti al tempo stesso. Così basterebbe fare ancora qualche distinzione e qualche eccezione in più nella legge, come è già in parte accaduto durante gli ultimi dieci anni cui abbiamo assistito ad una notevole evoluzione della legge, per evitare che non sia più possibile usare questo strumento di ricerca.

E non si lede il diritto alla privacy di alcuno facendo ricerca scientifica, specialmente se a farlo sono le università o gli enti pubblici di ricerca nazionali, istituzionalmente preposti a questo tipo di attività e tali da poter garantire un corretto uso dei dati personali necessari per il raggiungimento dei propri scopi, al punto

da aver sottoscritto un codice di deontologia in tema di privacy. Per dirla con Ichino (Ichino, 2003) “i dati non rappresentano una minaccia; al contrario, sono un bene pubblico che stimola la ricerca...nell’interesse della collettività”.

#### RINGRAZIAMENTI

Ringrazio il professor Enrico Pugliese per il supporto datomi durante la fase iniziale di questo lavoro e per avermi messa in contatto con massimi esperti in fatto di legge sulla privacy.

Un ringraziamento particolare va al professor Giuseppe Leti per il suo continuo supporto e per i suoi preziosi suggerimenti.

*Istituto di Ricerche sulla Popolazione e le Politiche Sociali*  
Roma

LOREDANA CERBARA

#### RIFERIMENTI BIBLIOGRAFICI

- L. BIGGERI (1999), *Diritto alla ‘privacy’ e diritto all’informazione statistica*, in Sistan-Istat, “Atti della Quarta Conferenza Nazionale di Statistica”, Roma, 11-13 novembre 1998, Roma, Istat, Tomo 1, pp. 259-279.
- F. BILLARI, A. ROSINA (2006), *La guerra dei sondaggi, Quando e perché fidarsi di un sondaggio*, [http://www.lavoce.info/news/view.php?id=10&cms\\_pk=2010&from=index](http://www.lavoce.info/news/view.php?id=10&cms_pk=2010&from=index).
- T. BOERI, M. PELLIZZARI (2003), *La deontologia di chi produce e detiene dati statistici: dalla possibilità alla certezza dell’accesso*, “Statistica”, n. 4, pp. 649-662.
- E. BRUSATI (2003), *Come si fanno i sondaggi*, “Induzioni”, n. 26/2003, <http://matematica.uni-bocconi.it/statistica/sondaggi.htm>.
- A. DI TARANTO (2006), *Seminario cnr su ‘tutela della privacy e pubblica amministrazione – Ricerca della SPSSA sulla Privacy*, Presidenza del consiglio dei ministri, Scuola Superiore della Pubblica Amministrazione, sede di Roma, 7-13 aprile.
- L. FABBRIS (1983), *L’indagine campionaria. Metodi, disegni e tecniche di campionamento*, NIS, Roma.
- L. FABBRIS (1990), (a cura di) *Rilevazione per campione delle opinioni degli italiani*, SGE Editoriali, Padova.
- L. FABBRIS (1997), *Statistica multivariata. Analisi esplorativa dei dati*, McGraw-Hill, Milano.
- L. FABBRIS (2000), (a cura di) *Il questionario elettronico*, CLEUP, Padova.
- A. ICHINO (2003), *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 “Statistica”, n. 4, pp. 673-683.
- N. ROSSI, A. ICHINO (2003), *Se la privacy non tutela la ricerca*, [http://www.lavoce.info/news/download.php?cms\\_pk=718](http://www.lavoce.info/news/download.php?cms_pk=718).
- A. SCHIZZEROTTO (2003), *Significati sociali e culturali del codice deontologico per il trattamento di dati personali a scopi scientifici e alcune sue implicazioni per la ricerca sociologica*, “Statistica”, n. 4, pp. 703-711.
- U. TRIVELLATO (2000), *Privacy, statistica e ricerca scientifica: verso i codici deontologici*, Relazione al Seminario organizzato dalla Facoltà di Scienze Statistiche, Università di Roma “La Sapienza” - CGIS - CSS, Roma.
- U. TRIVELLATO (2003), *Protezione dei dati personali e ricerca scientifica*, “Statistica”, n. 4, pp. 627-648.

- U. TRIVELLATO (2004), *Un Codice per tutelare privacy e ricerca*, [http://www.lavoce.info/news/download.php?cms\\_pk=1080](http://www.lavoce.info/news/download.php?cms_pk=1080).
- A. ZULIANI (2003), *I codici deontologici per il trattamento di dati personali a scopi statistici e di ricerca scientifica: qualche riflessione*, "Statistica", n. 4, pp. 713-722.

## RIASSUNTO

*Problemi di privacy nella selezione di piccoli campioni*

Il versante della ricerca sociale che fa uso di piccoli campioni per la produzione di microdati, trova oggi alcune difficoltà operative dovute alla legge sulla privacy. La legge sulla privacy è senza dubbio importante e necessaria perché siano garantiti i diritti dei cittadini italiani, come già accade in altri Paesi del mondo. Tuttavia non sembra opportuno limitare ulteriormente le possibilità di produzione dei dati da parte dei centri di ricerca nazionali, possibilità già peraltro compromesse dalla scarsa disponibilità economica che è sempre più diffusa nel settore della ricerca. Occorrerebbe quindi che si potesse includere nella legge la possibilità di usare le liste telefoniche per la selezione dei campioni per attività di ricerca scientifica di diretto interesse dei cittadini stessi, come accade nel caso della ricerca sociale effettuata sulla base di sondaggi di opinione dai centri di ricerca del CNR e di alcune università italiane.

## SUMMARY

*Privacy problems in the small sample selection*

The side of social research that uses small samples for the production of micro data, today finds some operating difficulties due to the privacy law. The privacy code is a really important and necessary law because it guarantees the Italian citizen's rights, as already happens in other Countries of the world. However it does not seem appropriate to limit once more the possibilities of the data production of the national centres of research. That possibilities are already moreover compromised due to insufficient funds is a common problem becoming more and more frequent in the research field. It would be necessary, therefore, to include in the law the possibility to use telephonic lists to select samples useful for activities directly of interest and importance to the citizen, such as the collection of the data carried out on the basis of opinion polls by the centres of research of the Italian CNR and some universities.