

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"

Andrea Ichino

1. INTRODUZIONE

Non sono un esperto dei problemi giuridici sollevati dalla necessità di contemperare il diritto alla riservatezza con le esigenze della ricerca statistica. Sono solo un ricercatore che si trova spesso ad analizzare per scopi scientifici dati provenienti da vari paesi del mondo e che, altrettanto spesso, osserva l'arretratezza macroscopica dell'offerta di dati, soprattutto di micro-dati, nel panorama Italiano.

Accolgo volentieri l'invito del Direttore di questa Rivista a dare un contributo per questo numero monografico, poichè sento il bisogno di esprimere le perplessità mie, e forse di altri ricercatori, riguardo al modo in cui questo problema viene affrontato in Italia. Per essere chiari fin dal principio, temo fortemente (ma spero di sbagliarmi) che il nuovo Codice Deontologico¹ non aiuterà a chiudere il divario esistente tra il nostro e gli altri paesi, e anzi potrebbe offrire, a chi i dati dovrebbe produrli e distribuirli, una giustificazione per non farlo nel modo in cui ciò viene fatto all'estero e che sarebbe auspicabile si diffondesse anche da noi. Si noti che, riferendomi all'estero, non sto parlando di Paesi in cui la "privacy" è calpestata, perchè il riferimento è ai nostri partners nella comunità Europea, oppure agli Stati Uniti o al Canada.

Dico questo pur essendo pienamente cosciente del fatto che il Codice Deontologico ha notevolmente migliorato la situazione rispetto a quanto disposto dalla disastrosa Legge 675 del 31 dicembre 1996 e dalle successive integrazioni e modificazioni ben descritte da Ugo Trivellato in questo stesso numero. Tuttavia, come quei genitori che chiedono ai figli perchè non hanno preso 10 se questi tornano a casa con un 9, così sento il dovere di segnalare che il risultato raggiunto dal Codice, per quanto migliorativo della situazione

¹ Vedi pagina ??? in questo stesso volume.

precedente, è lungi dall'essere sufficiente. Nelle pagine che seguono mi propongo di spiegare il perché di questa conclusione pessimistica.

1.1. *La logica ispiratrice del Codice Deontologico*

Il Codice Deontologico conferma il presupposto della disciplina pre-esistente – tipicamente italiano - in base al quale si ritiene che i ricercatori siano naturalmente portati ad usare i dati individuali in un modo che contrasta con la tutela della riservatezza delle persone. Viene però naturale chiedersi perché mai un ricercatore dovrebbe voler violare tale riservatezza, non essendo certo nel suo interesse. E se eccezionalmente qualche ricercatore male intenzionato lo facesse potrebbe, anzi dovrebbe, essere punito severamente, ma senza che ciò debba implicare il rendere l'attività di ricerca impossibile per tutti gli altri.

Questo è infatti il principio, radicalmente diverso da quello scelto in Italia, che prevale nei più avanzati Paesi Europei e negli USA: 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. Questa è anche la logica che ispira il progetto di legge presentato da Nicola Rossi al Parlamento Italiano nello scorso autunno (riportato in appendice) e che avrebbe potuto benissimo diventare la logica ispiratrice del Codice Deontologico. Purtroppo, temo che solo una disciplina ispirata a questa logica possa rendere possibile nel nostro paese una ricerca statistica di livello pari a quello osservato all'estero.

Mi si può obiettare che il Codice poteva muoversi soltanto all'interno dei margini concessi dalla disciplina Europea. Ma in primo luogo l'Europa è fatta dai suoi stati membri e nulla impediva all'Italia di farsi promotrice di una battaglia di avanguardia a livello Europeo volta a diffondere la logica ispiratrice alternativa sopra descritta.² E in secondo luogo, quella stessa disciplina Europea chiamata a giustificare la limitatezza dei margini di libertà concessi dal Codice, non sembra imporre limiti simili in altri Paesi dell'Unione. Una rapida occhiata agli articoli empirici pubblicati dalle principali riviste economiche internazionali offrirebbe al lettore numerosi esempi di basi dati Europee utilizzate per scrivere articoli di notevole interesse. Con tutta probabilità il nostro Garante considererebbe queste basi dati pericolose per la protezione delle riservatezze dei soggetti analizzati vista la ricchezza delle informazioni di natura longitudinale da esse offerte.³

² Del resto, lo stesso *Economist*, nel numero del 1 maggio 1999, ha sollevato numerose perplessità sulla disciplina europea in materia. E l'Eurostat, l'agenzia Europea per la raccolta, produzione e distribuzione dei dati, non rappresenta certo un modello da seguire e di cui andar fieri.

³ Ad esempio le basi dati del *UK Data Archive* nel Regno Unito, del DIW in Germania, i progetti LINDA (*Longitudinal Individual Data for Sweden*) e HUS (*Swedish Survey of Household*

Eppure, all'estero quelle banche dati esistono e sono messe a disposizione dei ricercatori.

Un esempio particolarmente significativo di ricerca possibile in un altro paese Europeo ma difficilmente realizzabile in Italia anche con il nuovo Codice, è offerto da un recente paper di Anders Bjorklund (Svezia), Mikael Lindhal (Svezia) e Eric Plug (Olanda) su *Intergenerational effects in Sweden: What can we learn from adoption data?* (Bjorklund et al., 2003) che utilizza dati sull'universo dei bambini adottati in Svezia, abbinando informazioni sulle famiglie di origine e adottive: un caso estremo in cui i ricercatori dispongono, riguardo ai soggetti analizzati, di informazioni più ampie di quelle note ai soggetti stessi. Infatti, normalmente, le famiglie adottive non dispongono di alcuna informazione per identificare le famiglie di origine e viceversa. Non credo che agli svedesi la tutela della riservatezza non interessi e, per quel che mi risulta, sono anche loro soggetti alla stessa disciplina Europea. L'impressione, tuttavia, è che essi siano in grado di tutelare la riservatezza almeno tanto quanto il nostro Garante, senza per questo rendere difficili o impossibili ricerche statistiche suscettibili di produrre risultati di notevole interesse.

1.2. La vaghezza dei "termini" usati dal Codice Deontologico

Un secondo motivo di perplessità di fronte al nuovo Codice riguarda la "vaghezza" con cui viene definito il criterio per stabilire quando un soggetto meritevole di protezione sia identificabile con i dati disponibili. Si tratta di un criterio molto importante poiché è quello da utilizzare per distinguere i dati *anonimi*, il cui utilizzo è libero, dai dati *personali*, per i quali vale la disciplina sulla tutela della riservatezza.

Il Codice accoglie la raccomandazione R(97)18 del Consiglio d'Europa che definisce dati personali "... any information relating to an identified or identifiable individual ... An individual shall not be regarded as 'identifiable' if identification requires an unreasonable amount of time and manpower. Where an individual is not identifiable, data are said to be anonymous.". Come ben ci ricorda Ugo Trivellato nel suo articolo in questo numero monografico, il fare riferimento alla *ragionevolezza dei mezzi* per verificare l'identificabilità è un fondamentale passo avanti rispetto alla normativa del 1996 secondo la quale era da considerarsi dato personale "qualunque informazione relativa a persona fisica, persona giuridica, ente o associazione identificati o identificabili anche indirettamente, mediante riferimento a qualsiasi altra informazione, ivi

Market and Non-Market activities, Swedish Level of Living Survey) per la Svezia, la DNB Household Survey per l'Olanda, la Continuous Family Expenditure Survey (ECPF) per la Spagna, i Quadros de Pessoal, per il Portogallo, solo per limitarsi agli esempi più significativi. Si noti inoltre che non tutti questi esempi riguardano paesi del nord d'Europa.

compreso un numero di identificazione personale".

Questo importante passo avanti, non è tuttavia un passo sufficiente e, anzi, l'indeterminatezza delle sue conseguenze rischia di generare una situazione non favorevole ad un incremento della disponibilità di dati nel nostro paese. Ciò perchè gli effetti reali del Codice verranno a dipendere in modo cruciale dal giudizio di chi deciderà quali *mezzi* siano da considerarsi *ragionevoli* per la identificazione dei soggetti meritevoli di tutela della riservatezza. In altre parole è proprio dalla interpretazione del concetto di *ragionevolezza* che dipenderanno gli effetti concreti della nuova disciplina.

Si dice che saranno le società scientifiche firmatarie del codice a stabilire i criteri per definire la *ragionevolezza dei mezzi*, ma dubito che ciò possa accadere senza il controllo del Garante. Se infatti questo controllo mancasse, società poco sensibili al problema della riservatezza potrebbero stabilire criteri in base ai quali il Codice diventerebbe di fatto irrilevante, ed è difficile pensare che il Garante voglia lasciare che questo accada. Se invece il Garante interverrà nella definizione dei criteri, temo che ben poco cambierà nella disponibilità di dati in Italia, o, per lo meno, poco cambierà in meglio.

La scelta di questo modo di procedere ricorda molto la scelta del legislatore in tema di protezione del lavoratore contro il licenziamento, un atto che l'impresa può compiere in presenza di un *"giustificato motivo"*. Chi poi decide nel concreto sulla giustificabilità del motivo è il giudice, cosicché, in aggiunta ai rilevanti costi sociali del contenzioso che in questo modo si viene a generare, se improvvisamente i giudici del lavoro in Italia cambiassero idee sui motivi ammissibili di licenziamento, i lavoratori non godrebbero più della stessa protezione oggi offerta dal nostro ordinamento.

In tema di tutela della riservatezza e ricerca statistica, accadrà qualcosa di analogo grazie a questo Codice: i risultati concreti dipenderanno dagli orientamenti di chi definirà quali siano i *mezzi ragionevoli* per l'identificazione di un soggetto meritevole di tutela. Con l'aggravante del costo sociale generato dalle pratiche burocratiche necessarie per influenzare chi deciderà della *ragionevolezza* o *sproporzionalità* dei mezzi utilizzati. E così anche in questo contesto, come in molti altri di questo disastroso paese, non sarà avvantaggiato chi ha buone idee ma chi saprà fare buona attività di *lobbying*.

Qualcuno potrebbe dire che non si poteva fare diversamente, ma, come suggerito nel paragrafo precedente, una soluzione alternativa c'era: adottare la logica ispiratrice del progetto di Nicola Rossi che risolve il problema senza alcun bisogno di stabilire cosa si debba intendere per *mezzi ragionevoli* di identificazione dei soggetti meritevoli di tutela.

1.3. Macchinosità burocratiche del Codice e la realtà della ricerca

In molti punti, poi, il Codice sembra essere stato scritto senza molto riguardo a come la ricerca statistica applicata si svolge nel concreto. Per fortuna, grazie, immagino, a Ugo Trivellato, è stata tolta la richiesta di distruggere i dati a fine ricerca, una imposizione la cui assurdità è talmente palese da risultare sorprendente. Ma rimangono prescrizioni veramente singolari. Ad esempio si richiede che i dati vengano utilizzati solo per gli scopi indicati nel "progetto di ricerca" originario da presentare preventivamente presso la propria università o società scientifica. Ogni trasferimento di dati ad una diversa ricercatrice (che ad esempio voglia verificare la replicabilità dei risultati precedentemente ottenuti) o il ri-utilizzo dei dati per un progetto di ricerca diverso da quello originale sembrerebbe richiedere la presentazione e archiviazione di un nuovo progetto di ricerca. E ciò per limitare l'incontrollato proliferare di ricercatori che detengano i dati indefinitamente.

Tuttavia, ancora una volta, all'estero il problema o non si pone o è risolto senza la stessa macchinosità burocratica che caratterizza la soluzione prevista dal nuovo Codice Deontologico Italiano. Ad esempio, i dati della *Current Population Survey*, analoghi a quelli della *Indagine Trimestrale sulle Forze di lavoro* dell'Istat sono scaricabili facilmente dal sito <http://www.bls.census.gov/cps/cpsmain.htm> (inclusa la struttura longitudinale che in Italia rende questi dati di natura *personale*) senza presentazione di alcun progetto di ricerca o altra formalità burocratica particolarmente gravosa. Per venire invece ad un esempio Europeo, lo *UK Data Archive*, una volta pagati i costi (per altro minimi) di accesso, consente di mettere i suoi dati a disposizione dell'intera comunità dei ricercatori dell'istituzione universitaria in cui insegno (lo European University Institute di Fiesole) richiedendo solo la comunicazione anche in forma elettronica del nome degli utilizzatori effettivi e del titolo della loro ricerca.

Questa è una modalità agile di diffusione dei dati che permette di registrare chi li sta utilizzando senza imporre inutili gravami burocratici nei casi in cui i dati vengano consultati non con un progetto ben definito in testa, ma, ad esempio, con scopi meramente esplorativi per verificare la fattibilità di una data ricerca. Penso ad esempio agli studenti del mio dottorato che nello scegliere il loro argomento di tesi spesso esplorano numerose basi dati, al fine di delineare un progetto di ricerca preciso. Ma non molto diverso è il modo di procedere per ricercatori con maggiore esperienza. Se per ogni articolo che essi producono fosse necessario scrivere e depositare preventivamente presso la propria università un progetto di ricerca dettagliato capace di prevedere ex-ante quali dati verranno utilizzati e come, la perdita di tempo sarebbe enorme, senza evidenti benefici per la collettività rispetto al molto più agile modo di procedere del *UK Data Archive*. A meno che non ci si voglia illudere sulla possibilità che qualcuno esamini effettivamente i progetti di ricerca che, secondo i dettami del Codice,

verranno archiviati presso le società scientifiche!

Poichè accade sistematicamente che nella realtà della ricerca i dati vengano in corso d'opera utilizzati diversamente da come inizialmente previsto, leggendo il Codice viene poi naturale chiedersi che cosa accadrà al povero ricercatore che si trovi nella situazione di aver sottomesso un progetto di ricerca non esattamente corrispondente al prodotto finito? C'è il sospetto, in tutto questo, di un intento di controllo da parte del Garante, il che non fa certo ben sperare per le sorti della libertà di ricerca in Italia.

Tra l'altro, se il problema fosse veramente solo quello di conoscere chi usa una particolare base dati, ai giorni nostri, in cui è prassi normale mettere sul *web* i risultati dei propri studi, una ricerca con *Google* fornirebbe probabilmente in pochi secondi una lista completa degli utilizzatori.

La scarsa attenzione alla realtà della ricerca, è poi forse ancor più evidente nel Codice, quando esso richiede che i ricercatori i quali notoriamente navigano nei fondi di ricerca, utilizzino le loro ampie risorse per comprare pagine di quotidiani mediante le quali informare "i soggetti tutelati" del fatto che su di essi si sta facendo ricerca!

Difficile non vedere un eccesso di "burocrazia" fine a se stessa in molte delle prescrizioni del Codice, che, oltre a complicare la vita ai ricercatori, rischiano comunque, proprio per la loro macchinosità, di rimanere lettera morta, come spesso accade in casi analoghi nel nostro paese. Son curioso ad esempio di verificare se, come e in quanto tempo le società scientifiche e le università si attrezzeranno per la ricezione e archiviazione (spero in formato elettronico) dei progetti di ricerca relativi all'utilizzo di dati personali.

Ma il punto fondamentale è un altro: posto che il 99% dei ricercatori non sembra avere alcun interesse a violare la privacy dei cittadini, perchè non dare fiducia ai ricercatori medesimi evitando gran parte di queste macchinosità burocratiche, ma punendo severamente ogni eventuale violazione del diritto alla riservatezza?

1.4. *Manca di esempi concreti*

Non poteva certo essere il Codice Deontologico a dare esempi concreti, ma è evidente che il giudizio sui suoi effetti reali dipenderà da come queste norme verranno implementate e interpretate e, soprattutto, da quali dati verranno considerati *personali* e quindi soggetti a questa disciplina.

Ora, però, gli esempi concreti sono necessari affinché i ricercatori sappiano come comportarsi nel futuro, e sarebbe quindi auspicabile che le società scientifiche che hanno sottoscritto il codice si esprimessero in proposito. Ad esempio, i dati della *Indagine sui Bilanci delle Famiglie Italiane* raccolta dalla *Banca d'Italia*, una delle pochissime fonti di informazione per la ricerca micro-

economica in Italia, saranno considerati di natura *personale* e quindi soggetti alla disciplina di tutela della riservatezza? O cosa dovrà essere cancellato in quei dati perchè non vengano considerati *personali*? E se l'Istat consentisse il collegamento longitudinale delle informazioni relative ad individui intervistati in trimestri successivi dalla *Indagine sulle Forze di Lavoro*, diventerebbero per questo *personali* i dati raccolti da tale *Indagine*? Se così fosse, sarebbe un disastro per la ricerca statistica e sociale in Italia, visti i macchinosi vincoli burocratici che il Codice impone a chi vuole utilizzare dati *personali*.

Gli esempi potrebbero continuare a lungo, soprattutto se prendessimo in considerazione quelle basi dati che già esistono all'estero, dove consentono ricerche di interesse notevolissimo, ma che ancora non esistono in Italia e che molti vorrebbero iniziare a costruire. Oltre agli esempi già citati nella nota 3 per l'Europa, si aggiungono le basi dati Statunitensi che per anni hanno fornito informazioni di interesse fondamentale ai ricercatori di tutto il mondo, quali il *Panel Study of Income Dynamics*, i dati del *US Census Bureau*, le diverse *waves* della *National Longitudinal Survey*, la *Highschool and Beyond Survey*, la *National Educational Longitudinal Survey*, la *Consumer Expenditure Survey*, la *Survey of Income and Program Participation* o i dati di *Statistics Canada* per il Canada, per limitarmi di nuovo solo ad alcuni degli esempi più famosi.

Ha ragione Trivellato: il "prossimo ineludibile impegno che attende ... l'intera comunità scientifica" italiana è la "produzione e distribuzione di adeguate basi dati". Purtroppo, però, il presupposto ispiratore di questo nuovo Codice vede sostanzialmente la produzione di basi dati adeguate come un "minaccia" per la riservatezza delle persone. Da questo presupposto è impensabile che possa nascere una spinta al potenziamento della raccolta e distribuzione dei dati statistici in Italia.

Il messaggio forte da lanciare doveva essere un altro. Ossia: i dati non rappresentano una minaccia; al contrario, sono un bene pubblico che stimola la ricerca consentendo una migliore comprensione della realtà e una più facile identificazione degli strumenti necessari per migliorarla nell'interesse della collettività.

Firenze, 12 luglio 2004

RIFERIMENTI BIBLIOGRAFICI

A. BJORKLUN, M. LINDHAL, E. PLUG, (2003), *Intergenerational effects in Sweden: What can we learn from adoption data?*, IZA Discussion Paper no. 1194 Bonn, Germany: IZA, July 2004.

APPENDICE

CAMERA DEI DEPUTATI

PROPOSTA DI LEGGE

d'iniziativa dei deputati

Nicola ROSSI

Disposizioni per la tutela della libertà di accesso alle informazioni statistiche per chi svolge attività di ricerca scientifica

Onorevoli Colleghi!

E' opinione comune che il sistema italiano della ricerca sia, salvo importanti eccezioni, collocato in una posizione marginale nel campo della ricerca mondiale e che, al tempo stesso, si dimostri spesso incapace di rispondere efficacemente alla domanda di innovazione proveniente dal sistema produttivo nazionale. E' altrettanto diffusa l'idea che la crescita della qualità scientifica del sistema italiano della ricerca passi per una sua crescente internazionalizzazione, per la definitiva adozione di criteri di selezione fondati sul merito e per una sempre maggiore accettazione di sistemi effettivi di valutazione. Il sistema dell'innovazione è più ampio di quello della ricerca ed è in questo quadro che vanno collocate la progressiva liberalizzazione e la semplificazione dei rapporti fra il mondo dell'industria e quello della ricerca.

Come in altri campi, peraltro, anche nel campo della ricerca permane un *gap* fra l'Italia e gli altri paesi industrializzati in buona misura riconducibile ad un eccesso di regolazione. Ad esempio, la semplice visita per via telematica del sito <http://www.ipums.umn.edu> consente ad ogni ricercatore, qualunque sia la sua nazionalità, di avere accesso a dati individuali dettagliati – ma assolutamente anonimi - sulla popolazione vivente negli Stati Uniti d'America così come rilevati dai censimenti effettuati tra il 1850 e il 1990 (il 2000 sarà disponibile tra breve). In sostituzione di complesse restrizioni volte a tutelare la *privacy* e dei relativi complicati formulari cartacei che tipicamente caratterizzano la distribuzione dei pochissimi dati disponibili nel nostro Paese, il sito americano invita semplicemente l'utente a usare le informazioni “*for good, never for evil*”.

In Svezia, il progetto LINDA (*Longitudinal Individual Data for Sweden*), nato da una collaborazione tra l'Università di Uppsala, il Ministero delle finanze svedese,

l'Agenzia svedese per le pensioni e la sicurezza sociale e *Statistics Sweden* (il corrispondente svedese del nostro ISTAT), mette a disposizione dei ricercatori dati individuali per una serie rilevante di anni estratti da un campione di dimensioni pari al 3% circa della popolazione svedese e al 20% circa della popolazione immigrata. I ricercatori hanno così accesso ad un insieme particolarmente ricco di informazioni che integra tra loro archivi informativi diversi ed in gran parte pubblici: dichiarazioni dei redditi a fini fiscali, rilevazioni dei censimenti della popolazione e delle abitazioni, dichiarazione dei redditi pensionabili, rilevazioni dei trattamenti pensionistici, indagini sulle forze di lavoro, rilevazione dei trattamenti di disoccupazione e non occupazione, indagini sulle carriere scolastiche, sulle assenze dal lavoro per malattia e per motivi familiari, rilevazioni dei redditi da lavoro. Le domande di accesso ai dati sono valutate in relazione alla validità scientifica delle ricerche proposte, senza limitazioni relative alla tutela della *privacy* che vadano oltre le “*normal confidentiality provisions*”, ossia l'impegno a non diffondere i dati personali senza autorizzazione, ad utilizzarli solo per l'attività scientifica e a comunicare i risultati ottenuti.

La ricerca scientifica utilizza in modo crescente insiemi di dati individuali. Sul piano analitico fenomeni come l'eterogeneità degli agenti, la molteplicità delle interrelazioni fra gli agenti stessi, la natura delle dinamiche che caratterizzano i loro comportamenti sono fra i campi di attività più rilevanti per l'odierna ricerca tanto teorica quanto applicata. Sul fronte della *policy*, invece, crescente importanza assumono le politiche mirate a gruppi relativamente ristretti di soggetti. La loro valutazione richiede la comparazione di situazioni individuali affette o meno dalle misure stesse. Ad esempio, se in Italia fossero disponibili, per qualità e quantità, microdati simili a quelli citati in precedenza, il dibattito sulle riforme economiche e sociali (dalla scuola, al mercato del lavoro, all'assistenza sanitaria, alle pensioni, all'immigrazione etc.) avrebbe un contenuto assai meno ideologico e potrebbe, invece, in misura assai maggiore essere basato su informazioni statistiche affidabili riguardo alla natura dei problemi e alla dimensione degli effetti delle proposte di riforma in discussione. Ultimo, ma non meno importante, ogni risultato scientifico dovrebbe essere, in linea di principio, replicabile e ciò è possibile solo se l'accesso alle basi di microdati è ragionevolmente libero.

Purtroppo tutto ciò in Italia non è oggi possibile anche perché i dati raccolti dalle amministrazioni pubbliche e archiviati presso l'ISTAT così come presso altri istituti non sono generalmente disponibili per la ricerca. Non solo nel modo mirabilmente integrato che caratterizza la Svezia e altri Paesi del Nord Europa, ma nemmeno separatamente archivio per archivio. Ciò non accade in parte per motivi di ordine tecnico (peraltro facilmente superabili nell'odierno ambiente

tecnologico anche mediante la costruzione di ambienti sicuri accessibili attraverso la rete), ma anche e forse soprattutto a ragione delle disposizioni della legge 31 dicembre 1996, n. 675 in tema di “Tutela delle persone e di altri soggetti rispetto al trattamento dei dati personali”. Una legge di cui a suo tempo si disse che eccede “sul piano dello zelo e anche del buon senso”.

La logica della legge 675/1996 sarà interamente recepita, a partire dal 1 gennaio 2004, del testo unico “Codice in materia di protezione dati” pubblicato sulla G. U. n. 123 del 29 luglio 2003, che pure contiene – fermo restando l’impianto concettuale della legge 675/1996 - qualche timido passo in avanti in particolare nella stesura dei Codici deontologici. In particolare, è da considerarsi positivamente il fatto che il Testo unico circoscriva la nozione di dati personali (e cioè riferiti a persona identificata o identificabile e per questo motivo oggetto di una sorta di diritto di proprietà a favore della persona stessa) considerando tali solo i dati che consentano l’identificazione attraverso l’utilizzo di “mezzi ragionevoli”. Permangono, peraltro, tanto nel Testo unico quanto nei Codici deontologici bardature regolatorie che in larga misura finiscono per vanificare i deboli segnali di apertura. In alcuni casi, la distanza fra l’impianto logico delle legislazioni occidentali e la nostra minuziosa casistica è straordinariamente eloquente. Valga per tutti quanto il Codice deontologico prescrive a proposito della possibilità di conservare, per soli scopi scientifici, dati personali per periodi più lunghi di quelli inizialmente preventivati. Se il nostro legislatore disciplina dettagliatamente ogni possibile situazione salvo poi includere tutti gli “altri casi in cui ciò risulti essenziale e adeguatamente documentato per le finalità perseguite”, il legislatore anglosassone – con ben altra eleganza e realismo – stabilisce che che *“personal data which are processed only for research purposes ... may be kept indefinitely”*.

In totale contrasto con la filosofia prevalente in altri paesi, la legge n. 675/1996 è essenzialmente fondata sul presupposto – tipicamente italiano - secondo cui i ricercatori sono naturalmente portati a usare i dati individuali (e quindi anche se anonimi o resi tali) in un modo che contrasta con la tutela della riservatezza delle persone. Sulla base di questo presupposto, i dati individuali raccolti da qualunque ente pubblico o privato sono resi preventivamente indisponibili per l’attività di ricerca a meno che il ricercatore non accetti di passare attraverso le forche caudine di un insieme di vincoli burocratici tali da rendere i dati stessi di difficilissimo accesso e spesso inutilizzabili a fini statistici. Non è raro, ad esempio, che in omaggio alla legge 675/1996 si rendano non casuali i campioni accessibili da parte dei singoli ricercatori con il risultato di renderli inutilizzabili per indagini statistiche che ambiscano a risultati rappresentativi. Né è raro, per fare un secondo esempio, che - sempre per ottemperare per ottemperare alla legge 675/1996 – vengano oscurate

informazioni essenziali ai fini dell'analisi scientifica e particolarmente rilevanti per la discussione di tematiche di grande importanza (valga per tutti il riferimento alle analisi sulle disparità regionali).

Il tema della ricerca non chiede di essere affrontato solo in termini di risorse. Altrettanto importante è che muti l'ambiente nel quale operano i nostri ricercatori nel senso di garantire loro la libertà di azione senza la quale la ricerca non è più tale. Nei più avanzati Paesi Europei e negli USA l'atteggiamento è radicalmente diverso: 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. A questa trasparente filosofia si ispira il progetto di legge qui illustrato il quale parte da una semplice, elementare, considerazione. Il rispetto della privacy è un valore per il ricercatore e non solo per norma deontologica. In negativo, egli non ha motivo alcuno per metterla a rischio nello svolgimento della propria attività. In positivo, egli sa che la salvaguardia della riservatezza dei dati personali è condizione primaria per continuare ad utilizzarli e per acquisirne di nuovi e quindi per operare nel proprio campo. In sintesi, è ragionevole presumere che per chi svolge attività di ricerca scientifica l'identificazione degli individui da cui originano i dati elementari oltre ad essere del tutto inutile comporti, per definizione e salvo controprova *ex post*, mezzi irragionevoli.

L'approvazione di questo progetto, oltre a consentire un'attività di ricerca scientifica di enorme utilità per il dibattito politico e sociale ed a mettere i ricercatori italiani su un piano di parità con i loro colleghi statunitensi ed europei, costituirebbe un passo significativo, anche se piccolo, verso una riscrittura dei rapporti fra Pubblica amministrazione e cittadini. Non di sole leggi è fatta, infatti, la riforma della Pubblica amministrazione ma anche e soprattutto di un cambio di mentalità di quest'ultima fino ad ora fin troppo abituata a presumere che tutti i cittadini mirino – salvo prova contraria – a violare la legge. Salvo, poi, condonarne i reati.

Infine, tutelare il diritto di accesso dei nostri ricercatori alle informazioni statistiche riporterebbe il nostro ordinamento in linea con l'indirizzo Costituzionale volto a favorire la libera circolazione delle idee e delle informazioni, il cui fondamento va individuato oltre che nell'art. 21 Cost., anche e soprattutto nell'art. 3, dato lo stretto nesso esistente tra la circolazione delle informazioni e la effettiva partecipazione del cittadino all'organizzazione della società, e nell'art. 2, poiché tra i diritti inviolabili della persona deve essere ricompreso quello sancito dall'art.19 della *Dichiarazione universale dei diritti dell'uomo* di “cercare, ricevere e diffondere informazioni”.

ART.1

I vincoli e i limiti alla conoscenza e alla trattazione dei dati personali di cui alla legge 31 dicembre 1996, n. 675 e successive modifiche e integrazioni, nonché alle disposizioni legislative e regolamentari attuative della predetta legge n. 675 del 1996, non si applicano per lo svolgimento di attività di ricerca scientifica svolta da soggetti pubblici o privati finalizzata all'acquisizione di nuove conoscenze o alla dimostrazioni di tesi scientifiche per mezzo della utilizzazione di dati statistici. I soggetti pubblici comunque competenti per le rilevazioni statistiche sono tenuti a fornire, su richiesta motivata, ai ricercatori, e agli istituti di ricerca pubblici o privati i dati in loro possesso. Ricercatori ed enti di ricerca pubblici o privati sono comunque tenuti a non trasferire a terzi i dati stessi ed a non identificare gli individui cui gli stessi dati fanno riferimento né a presentare i dati stessi in modo tale da consentire l'identificazione. Restano coperti da riservatezza i dati trattati dai soggetti di cui all'articolo 4, comma 1, della legge n. 675 del 1996 e successive modificazioni e integrazioni.

Fermo restando quanto stabilito dal comma 1 ai ricercatori, ai responsabili dei progetti di ricerca, ai direttori delle strutture pubbliche e private di ricerca si applicano nelle attività di utilizzazione e divulgazione dei dati, e del trattamento dei dati personali ovvero della loro detenzione, le disposizioni di cui alla legge n. 675 del 1996, e successive modificazioni e integrazioni, aumentate le sanzioni penali e amministrative ivi previste di un terzo, nonché, in quanto compatibili, i codici di deontologia e di buona condotta per il trattamento dei dati personali di cui alle deliberazioni del Garante per la protezione dei dati personali.

RIASSUNTO

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"

Con il nuovo Codice è stata persa una occasione per sancire in modo chiaro che i dati per la ricerca statistica non rappresentano una grave minaccia alla riservatezza delle persone. Al contrario, prima che una minaccia, essi sono un bene pubblico che stimola la ricerca consentendo una migliore comprensione della realtà e una più facile identificazione degli strumenti necessari per migliorarla nell'interesse della collettività.

La logica che avrebbe dovuto ispirare il Codice doveva essere quella di favorire un accesso ampio e facile ai dati personali, punendo severamente ogni eventuale violazione della riservatezza. Si è preferito invece rendere difficile la vita ad ogni ricercatore con incombenze burocratiche macchinose, che probabilmente avranno solo costi senza benefici. Lascia inoltre perplessi la vaghezza del concetto di *mezzi ragionevoli* per la identificazione dei soggetti meritevoli di tutela. Gli effetti reali del Codice dipenderanno da quale interpretazione verrà data nel concreto di questo concetto, da cui dipende la distinzione cruciale tra dati *anonimi* e dati *personali*.

SUMMARY

The doubts of a data user concerning the "Code of practice for the use of personal data for statistical and scientific research"

The new Code misses an important occasion to state clearly that data for statistical research are not a major threat for personal privacy. They are instead a public good that stimulates research allowing for a better understanding of reality and for an easier identification of the instruments that are necessary to improve collective welfare.

The goal of the code should have been to allow an easy access to data, punishing severely any violation of privacy. The preference has been instead to impose on individual researchers complicated bureaucratic requirements, which will probably have only costs without benefits. Moreover, the use of the concept of *reasonable means* to define the cases in which the data allow for the identification of protected subjects raises concern because of its vagueness. The real consequences of the Code will depend on how this concept will be concretely interpreted. This interpretation will in fact determine which data should be considered *personal* and which ones can instead be considered *anonymous*.