

PROTEZIONE DEI DATI PERSONALI E RICERCA SCIENTIFICA

Ugo Trivellato (*)

1. PREMESSA

La tensione fra accesso ai dati per scopi scientifici e riservatezza è al crocevia di due ordini di considerazioni. L'uno attiene alla consolidata attenzione dei ricercatori per la *confidentiality*: detto altrimenti, per il riserbo al quale essi si attengono nel trattare dati riferiti (o riferibili) a singole persone. L'altro riguarda la più recente, e più ampia, preoccupazione per la *data protection*, cioè a dire per la protezione legale degli individui con riguardo al trattamento dei loro dati personali.

È soprattutto questo secondo aspetto che, a partire dalla metà del secolo scorso e in particolare negli ultimi vent'anni, è venuto assumendo rilievo. L'interrogativo che, naturalmente, si pone riguarda le ragioni di questa evoluzione.

L'impegno alla *confidentiality* nell'uso di dati per scopi scientifici si configura come un dovere che fa capo ai ricercatori: tradizionalmente affidato all'etica professionale, rafforzato da codici deontologici, in qualche caso da obblighi di legge per specifiche categorie di operatori¹. Perché questo impegno, a lungo giudicato una garanzia sufficiente contro i rischi di violazione della *privacy*, è stato progressivamente percepito come uno schermo inadeguato?

La ragione sta nell'emergere di sensibilità largamente nuove e di problemi altrettanto nuovi, o connotati in termini più acuti, che portano all'affermazione di un diritto individuale alla *privacy*. In sintesi, operano tre ordini di ragioni.

- (a) C'è una forte attenzione al tema dei diritti umani. A valle del secondo conflitto mondiale – uno spartiacque di importanza decisiva nella storia del '900 e nella coscienza dell'umanità –, essi vengono riproposti con forza. È del 1950 la *Convention for the Protection of Human Rights and Fundamental Freedoms*. In questo quadro, vi è un'enfasi sulla dimensione delle libertà individuali, su una loro lettura ampia, che si estende al diritto alla *privacy*.

(*) Questo testo riproduce, con non trascurabili integrazioni e modifiche, la relazione che ho tenuto al 2° Congresso Nazionale della Società Italiana di Statistica Medica ed Epidemiologia Clinica, Università di Brescia, 1-4 ottobre 2003.

¹ Tipicamente per gli statistici ufficiali vi è l'obbligo di legge al segreto d'ufficio (per l'Italia, vedi gli artt. 8 e 9 del decreto legislativo n. 322/1989, rispettivamente sul “segreto d'ufficio” e sul “segreto statistico”).

- (b) Con la ‘rivoluzione informatica’, l’impetuoso sviluppo delle *Information and Communication Technologies* e l’avvento della ‘società dell’informazione’, emerge una diffusa preoccupazione: ai benefici che essa offre si accompagna il grave rischio di intrusione nelle vite private. L’incubo totalitario tratteggiato da Orwell in *1984* prende le forme della *surveillance society*, in cui gran parte dei nostri comportamenti – lavorare, fare la spesa, divertirsi, viaggiare, comunicare – lascia segni permanenti, facilmente identificabili, forse identificati.
- (c) La ricerca scientifica² utilizza in maniera crescente insiemi di dati individuali, per l’effetto congiunto di fattori operanti dal lato della domanda – gli interrogativi della ricerca e il modo con cui vengono formulati (su questo tornerò brevemente nel seguito) – e da quello dell’offerta – la crescente disponibilità di basi di microdati, spesso di origine amministrativa³, e le accresciute capacità di trattarli –. A fronte di ciò, il dovere di *confidentiality* dei ricercatori è avvertito come una protezione fragile, e il tema del diritto alla *privacy* investe anche l’uso di dati personali per finalità scientifiche.

Proprio in relazione a queste più acute sensibilità e al mutato contesto con cui si confrontano, la portata stessa del diritto alla *privacy* è venuta espandendosi in misura significativa: dai confini tutto sommato circoscritti del diritto a non vedere violata la sfera privata – il diritto “*to be let alone*” della dottrina statunitense di oltre un secolo fa⁴ – all’accezione progressivamente più ampia di una sorta di diritto di proprietà che si estende a pressoché tutti i tipi di dati personali⁵ e a gran parte dei loro usi⁶.

Questo processo si è tradotto, tra l’altro, in raccomandazioni internazionali e in una normativa comunitaria, che configurano in termini nuovi il diritto alla *privacy* così come il suo temperamento a fronte di scopi di ricerca. A livello internazionale hanno rilievo soprattutto due atti del Consiglio d’Europa:

- la *Convention for the protection of the individual with regard to the automatic processing of personal data*, del 1981 (Council of Europe, 1981), e

² Nel seguito, spesso userò semplicemente il termine ‘ricerca’.

³ Il crescente utilizzo di archivi amministrativi per finalità scientifiche ha implicazioni di rilievo. Confrontati con *surveys* campionarie, infatti, gli archivi amministrativi tipicamente presentano caratteristiche che favoriscono l’identificabilità delle persone e insieme possono metterne a rischio i diritti in materia di *privacy* (primo fra tutti, quelli di essere informate delle finalità e modalità di trattamento dei dati). In particolare, gli archivi amministrativi: (i) sono di tipo censuario, cioè a dire si riferiscono a tutte le unità della popolazione di interesse; (ii) hanno un elevato grado di accuratezza; (iii) comportano l’inserimento obbligatorio di dati personali.

⁴ L’affermazione è del giudice Cooley e risale al 1888. La prima trattazione sistematica del diritto alla *privacy* è di Warren e Brandeis (1890). Per una recente rassegna della legislazione e della dottrina statunitense in materia, vedi Stevens (1998).

⁵ L’eccezione riguarda le cosiddette ‘variabili pubbliche’, definite dal legislatore italiano come “*dati provenienti da pubblici registri, elenchi atti o documenti conoscibili da chiunque*” (decreto legislativo n. 196/2003, art. 24, comma 1, lettera c).

⁶ Limitazioni di vario grado al diritto alla protezione dei dati personali sono poste per trattamenti in specifici settori o per specifiche finalità, per contemperarlo con altri diritti o interessi ‘meritevoli’ (ad esempio, con il prevalente interesse collettivo che contraddistingue i trattamenti in ambito giudiziario, da parte delle forze di polizia, per la difesa e la sicurezza dello Stato, ecc.).

– la Raccomandazione R(97)18 *on the protection of personal data collected and processed for statistical purposes*, accompagnata da un illuminante memorandum esplicativo (Council of Europe, 1997b)⁷.

A livello dell'Unione Europea (UE), il testo fondamentale è la Direttiva n. 95/46/CE *relativa alla tutela delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati* (Parlamento Europeo e Consiglio dell'UE, 1995)⁸.

Si tratta di documenti che, va da sé, non sono oro colato. L'autorevole *The Economist*, in un numero del 1999 largamente dedicato a “*The end of privacy*”, ha espresso argomentate riserve sulla normativa europea: sul nuovo concetto sul quale essa poggia – gli individui hanno un diritto di proprietà sull'informazione che li riguarda –; sul dettaglio e il rigore delle protezioni previste; sulla ragionevolezza, e la stessa possibilità, di applicarle in maniera stringente⁹. Sono rilievi in buona misura fondati.

Ai nostri fini la normativa e le raccomandazioni appena richiamate rivestono comunque un'importanza cruciale, per motivi di merito e giuridici. Nel merito, perché i testi del Consiglio d'Europa – che si collocano alla conclusione di approfondite istruttorie di gruppi di esperti – forniscono contributi ragionati, convincenti. Sul terreno normativo, perché per gli stati membri dell'UE vi è l'obbligo di adeguare la legislazione nazionale alla direttiva comunitaria – la legge nostrana sulla *privacy*, n. 675/1996 con le integrazioni e modifiche che l'hanno seguita, risponde appunto a questo scopo –, e perché la normativa nazionale rinvia alle raccomandazioni del Consiglio d'Europa per la disciplina in tema di statistica e ricerca.

Prima di entrare *in medias res*, è bene che precisi la prospettiva con cui affronto la questione dei rapporti fra ricerca scientifica e *privacy*. Innanzitutto, muovo dal punto di vista di un analista, cioè a dire di una persona interessata a (raccogliere e soprattutto) trattare dati individuali per scopi scientifici. In secondo luogo, porto l'interesse sugli aspetti basilari della tematica e sulle implicazioni che essi hanno per politiche di accesso ai/diffusione dei dati per scopi di ricerca, lasciando sullo sfondo i dettagli tecnici, pur non trascurabili. Non potrò evitare, tuttavia, di misu-

⁷ Specificamente per il trattamento di dati medici vi è poi la Raccomandazione R(97)5 *on the protection of medical data* (Council of Europe, 1997a).

⁸ Specificamente in materia di statistica e ricerca si aggiungono poi il Regolamento del Consiglio n. 322/97 *relativo alle statistiche comunitarie* (Consiglio dell'UE, 1997) e il Regolamento della Commissione n. 831/2002, di applicazione del precedente *per quanto riguarda l'accesso ai dati riservati per fini scientifici* (Commissione delle Comunità Europee, 2002).

⁹ È il numero del 1 maggio 1999, che dedica al tema copertina, articolo d'apertura e sezione speciale. Nell'articolo di apertura si afferma, tra l'altro: “*The EU data protection directive, the most sweeping recent attempt [to protect privacy], gives individuals unprecedented control over information about themselves. This could provide remedies against the most egregious intrusions. But it is doubtful whether the law can be applied in practice, if too many people try to use it. ... Policing the proliferating number of databases and the thriving trade in information would not only be costly in itself; it would also impose huge burdens on the economy. Moreover, such laws are based on a novel concept: the individuals have a property right in information about themselves. Broadly enforced, such a property right would be antithetical to an open society. It would pose a threat not only to commerce, but also to a free press and to much political activity, not to say of everyday conversation*” (p. 13).

rarmi con i risvolti giuridici della questione: vedrò di argomentare al meglio, anche se servirebbero competenze che non sono nel mio bagaglio.

2. UNO SGUARDO AI FONDAMENTALI

Comincio dai fondamentali. Il punto dal quale occorre prendere le mosse è il riconoscimento di una tensione tra ‘valori’ e diritti diversi, e la conseguente necessità di trovare un contemperamento fra gli stessi. Ciò vale, innanzitutto, per la legislazione generale sulla protezione dei dati personali, per la quale il *trade-off* e l’esigenza di bilanciamento si pongono tra il diritto alla *privacy* e il diritto fondamentale alla libertà di informazione: che è libertà di ricevere e di fornire informazione, e dove la libertà di riceverla, l’informazione, implica anche la libertà di cercarla¹⁰.

Questa stessa logica, di un bilanciamento fra ‘valori’ diversi, sta esplicitamente alla base della previsione di una specifica normativa in materia di dati personali utilizzati per scopi scientifici. Quando entrano in gioco statistica e ricerca, il bilanciamento è modificato nel senso che si permettono possibilità di trattamento dei dati personali più ampie di quelle consentite in via generale, in ragione delle intrinseche necessità della ricerca, vitali per il suo svolgersi e per la sua funzione sociale.

Ma a che cosa ci riferiamo quando parliamo di finalità statistiche e di ricerca, in relazione alla tutela della *privacy*? E perché queste finalità richiedono di trattare basi di dati individuali, non solo anonimi ma – quando necessario – personali, cioè a dire relativi a unità identificate o identificabili? Rispondere, sia pure in maniera sintetica, a questi interrogativi è essenziale, perché dalle risposte discendono le linee-guida per la disciplina della *privacy* nel settore.

L’ambito al quale ci riferiamo, innanzitutto. È un ambito ampio. Esso comprende entrambe le dimensioni nelle quali tradizionalmente si articola l’attività di ricerca e che insieme ne configurano la funzione sociale: la dimensione analitica, ‘di base’, orientata al progredire delle conoscenze; quella ‘applicata’, di supporto a *policies*, più direttamente mirata a sviluppare e utilizzare le conoscenze per migliorare le nostre condizioni di vita.

In questo quadro, ciò che caratterizza le finalità statistiche e di ricerca può essere riassunto in tre argomentazioni.

In primo luogo, quel che contraddistingue la finalità statistica è l’‘uso collettivo’ dei dati personali. Il memorandum che accompagna la Raccomandazione R(97)18 del Consiglio d’Europa chiarisce il punto in maniera nitida: “*Individuals are the necessary medium for the background information but they are not regarded as significant in their own right. In fact, starting with the basic material represented by individual information about many different people, the statistician elaborates results designed to ‘characterise a collective phenomenon’ . . . In other words, the statistical result separates the information from the person. . . . This is the distinct characteristic of statistical purpose: per-*

¹⁰ Vedi l’art. 10 della citata Convention for the Protection of Human Rights and Fundamental Freedoms.

sonal data are collected and processed with a view to producing consolidated and anonymous information?” (Council of Europe, 1997b, Explanatory memorandum, pp. 27-28; il grassetto è nell’originale). Certo, questa caratterizzazione può avvenire – e sovente avviene – in maniera articolata: in termini non solo medi, ma anche di variabilità; con riguardo non solo a statistiche di distribuzioni empiriche, ma anche a parametri di modelli; con riferimento anche a risultati non numerici, ad esempio in forma di tipologie o di conferma/smentita di ipotizzate relazioni causali. In ogni caso, alle finalità statistiche è estraneo ogni ‘uso individuale’ dei dati personali, ogni interesse ad utilizzarli vuoi per giungere a identificare una persona, vuoi per prendere decisioni che a una determinata persona si riferiscano.

In secondo luogo, sotto il profilo della tutela della *privacy* – quindi con riguardo ad attività di ricerca che trattano materiale di osservazione costituito da dati personali –, possiamo ben dire che le finalità scientifiche coincidono con quelle statistiche. Anche la ricerca scientifica, infatti, fa dei dati personali un ‘uso collettivo’, mirato a trovare regolarità e leggi generali, mai puntato sul caso singolo. “*Scientific research uses statistics as one of a variety of means of promoting the advance of knowledge. Indeed, scientific knowledge consists in establishing permanent principles, laws of behaviour or patterns of causality which transcend all the individuals to whom they apply. Thus it is aimed at characterising collective phenomena, this being the very definition of statistical results. It could be said, therefore, that research becomes statistical at a certain stage in its development.*” (Council of Europe, 1997b, Explanatory memorandum, p. 30). E argomenti analoghi valgono per la ricerca applicata, di supporto a politiche.

Si tratta di chiarificazioni che hanno implicazioni di rilievo. Il tema della tutela della *privacy* si pone in maniera sostanzialmente comune per l’insieme degli scopi statistici e di ricerca, estesi ai diversi campi delle scienze sociali, del comportamento, epidemiologiche e biomediche. Inoltre, è immediato riconoscere che il rispetto della *privacy* è un ‘valore’ per il ricercatore, e non solo per norma deontologica. In negativo, egli non ha alcuna ragione, nello svolgimento della propria attività, che possa indurlo a sacrificare o anche soltanto a mettere a rischio la *privacy*. In positivo, la salvaguardia della riservatezza dei dati personali di cui dispone è condizione perché egli possa continuare a utilizzarli, e possa acquisirne altri: in altre parole, per poter svolgere il proprio mestiere.

Da queste chiarificazioni viene anche una prima risposta al secondo interrogativo: davvero le finalità scientifiche richiedono di trattare dati personali? A ben vedere, nella comunità scientifica un tale interrogativo neppure si pone, tanto la risposta positiva è banalmente scontata. Ciò non significa, tuttavia, che questa consapevolezza sia condivisa dall’opinione pubblica, sia patrimonio della società. Anzi, gran parte delle difficoltà che abbiamo incontrato – e incontriamo – sta proprio qui: nel fatto che vi è una inadeguata percezione dell’importanza sociale della ricerca, delle sue esigenze interne, in particolare della necessità che essa ha di trattare insiemi di dati individuali – se necessario personali –. È nella nostra responsabilità di ricercatori contribuire a far crescere questa consapevolezza. Se, dunque, per lettori statistici – e più in generale ricercatori – non serve certo che mi dilunghi sull’argomento, due scarse riflessioni tornano utili avendo come riferimento l’opinione pubblica e gli *opinion makers*.

Sul piano analitico, siamo avvertiti dell'importanza che rivestono fenomeni quali l'eterogeneità degli agenti (individui, famiglie, imprese, istituzioni sociali), la molteplicità delle interrelazioni fra gli agenti stessi, le dinamiche che ne caratterizzano i comportamenti. Nel campo degli studi economici, lo sviluppo delle *household panel surveys* e di basi di dati integrate lavoratore-impresa è forse la testimonianza più evidente delle esigenze informative della ricerca, e insieme delle potenzialità che essa manifesta¹¹. L'analogo accade sul fronte degli studi epidemiologici, con l'integrazione di informazioni sulle storie di vita estese dalle condizioni di salute/malattia alla dimensione demografica e a quella lavorativa – e agli aspetti di contesto –.

Se spostiamo l'attenzione sull'azione pubblica, un tratto distintivo è la crescente importanza che assumono politiche mirate a gruppi relativamente ristretti di soggetti. Ciò vale per le politiche industriali, per le politiche attive del lavoro, per le politiche sociali. Ora, a fini di valutazione dell'impatto di una politica occorre isolare gli effetti dell'intervento dagli effetti di altri fattori, il che richiede di comparare i risultati delle persone che sono state sottoposte all'intervento in questione con quelli di persone che presentano le loro stesse caratteristiche, ma che non sono state destinatarie dell'intervento. E per questo è necessario, appunto, disporre di informazioni sui singoli individui¹².

In definitiva, soltanto studi di tipo microanalitico, che poggiano su basi di dati individuali – sovente longitudinali –, offrono il supporto informativo adeguato per rispondere a cruciali interrogativi conoscitivi e per monitorare/valutare (e, indirettamente, per disegnare) politiche di rilevante importanza.

Un terzo punto, infine, merita di essere messo in luce. La scienza è un processo incrementale, che poggia sulla discussione aperta e sulla competizione fra spiegazioni alternative. La possibilità che siano condotte rianalisi sugli stessi dati – i cosiddetti *replication studies* – è quindi fondamentale per la scienza¹³. A questi fini, è vitale che l'accesso alle basi di microdati sia ragionevolmente libero. Recentemente questo obiettivo è stato enunciato in maniera efficace nei seguenti termini: “*universal access ...for all 'bona fide' non-commercial users (subject to registration and standard undertakings of non-abuse), at no cost*” (Jenkins, 1999, p. 81)¹⁴.

3. DOVE FERMARE IL PENDOLO?

Ho indugiato sui fondamentali. Ma avere ben chiare motivazioni e criteri-guida è importante. Essi portano a fermare il pendolo, nell'esercizio di bilanciamento di cui dicevo, in una posizione che assegna al ricercatore una sorta di ‘credito di riservatezza’. Naturalmente un credito ragionato, quindi dimensionato in maniera

¹¹ Per due esempi tra i molti, vedi rispettivamente Jenkins (1999) e Contini (2002).

¹² In proposito vedi, tra gli altri, Rettore e Trivellato (1999) e Rettore, Trivellato e Martini (2003).

¹³ Ciò vale non solo per la ricerca di base, ma anche per la ricerca applicata, di supporto a politiche. Heckman e Smith (1995, p. 93) argomentano in maniera convincente che “*evaluations build on cumulative knowledge*”.

¹⁴ Alle stesse conclusioni portano argomentazioni di differente natura, sul ruolo dell'informazione statistica in una società democratica. Per queste rimando a Trivellato (2000).

parsimoniosa e definito con attenzione alle dimensioni tecniche coinvolte. E, per un altro verso, accompagnato da garanzie – e da sanzioni – che valgano ad evitare trascuratezze o usi impropri dei dati.

Rilevanti indicazioni al riguardo vengono dalla normativa comunitaria e, soprattutto, dalla più volte citata Raccomandazione R(97)18 del Consiglio d'Europa. Raggruppo queste indicazioni attorno a quattro punti: (i) la specificazione della nozione di 'dato personale' e di individuo 'identificabile'; (ii) i principi positivi specificamente rilevanti per il trattamento di dati personali per scopo statistici e di ricerca; (iii) le deroghe salienti a norme generali sulla tutela della *privacy*; (iv) la valorizzazione delle associazioni scientifiche e professionali, segnatamente tramite l'adozione di codici deontologici.

Una prima delucidazione ha a che vedere con la stessa nozione di 'dato personale'. La definizione che ne dà la Raccomandazione R(97)18, è la seguente: “*Personal data' means 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.*”. Il punto è chiaro, e piuttosto importante. Esso sta a significare che (i) per decidere dell'identificabilità o meno di una persona si fa riferimento al criterio dei 'mezzi ragionevoli', e che (ii) vi è una netta distinzione fra dati personali e dati (in origine o resi) anonimi, la disciplina sulla tutela della riservatezza valendo soltanto per i primi.

Quanto ai principi positivi della disciplina specificamente rilevante per i trattamenti a fini statistici, meritano di essere segnalati i seguenti.

- Il principio dell'uso proprio di dati raccolti e trattati per scopi statistici, nel senso che essi debbono servire solo a tali scopi, e non possono essere usati per prendere decisioni o misure nei confronti dell'interessato.
- Una sorta di 'principio di parsimonia' – non saprei come chiamarlo altrimenti –, collegato al fatto che l'interesse di statistica e ricerca è in un 'uso collettivo' dei dati individuali. Quando siano sufficienti per gli specifici scopi di una ricerca, saranno utilizzati soltanto dati anonimi. Altrimenti, l'utilizzazione di dati personali sarà limitata alle operazioni per le quali essi siano necessari (ad esempio, per la raccolta, per raffronti/abbinamenti fra archivi, per la conduzione di indagini longitudinali), valendo in generale il principio della separazione dei dati identificativi¹⁵ dai dati di interesse per le analisi.
- Un insieme di principi in materia di 'trattamenti intermedi' (conservazione, sicurezza, comunicazione di dati personali), che combina chiarezza con sobrietà. Saggiamente, il testo della Raccomandazione non si addentra in prescrizioni dettagliate, lasciando lo spazio per i necessari adattamenti degli strumenti e delle procedure – tecniche e organizzative –.
- Un fermo divieto, invece, a che i risultati statistici siano pubblicati o diffusi, quando, sulla base degli stessi, le persone siano identificabili, “*unless dissemination or publication manifestly presents no risk of infringing the privacy of the data subjects*”.

¹⁵ Per 'dati identificativi' si intendono quei dati che permettono l'identificazione diretta dell'interessato, ma che non sono successivamente usati per produrre risultati statistici.

Un diverso insieme di disposizioni riguarda, poi, deroghe a prescrizioni generali in materia di protezione dei dati personali. Esse includono:

- semplificazioni e deroghe ai principi dell'informazione all'interessato e del consenso dell'interessato;
- la possibilità di trattare per scopi statistici o di ricerca dati originariamente raccolti e trattati per altri scopi (essenzialmente per usi amministrativi), con un rilassamento dell'obbligo di informare gli interessati;
- la possibilità di estendere la conservazione di dati personali “*per periodi più lunghi [di quanto necessario per gli scopi per i quali i dati erano stati raccolti ... o ulteriormente trattati]*” (Direttiva n. 95/46/CE, art. 6).

La naturale conclusione di questo insieme di disposizioni sui trattamenti per scopi statistici e di ricerca è, come già ho anticipato, la valorizzazione del ruolo delle associazioni scientifiche, della capacità di autodisciplina delle stesse e dei singoli ricercatori, imperniata su codici deontologici. È questa una tessera importante del mosaico: indispensabile per fornire risposte adeguate a un quadro di problemi segnato dallo spessore delle dimensioni tecniche e dal forte dinamismo degli strumenti che via via si rendono disponibili – per il trattamento dei dati così come per la loro protezione –.

4. IL PERCORSO DELLA NORMATIVA ITALIANA

Come è intervenuto il legislatore italiano sullo snodo fra *privacy* e ricerca? Come ha interloquito la comunità scientifica? Come si è giunti alla definizione delle regole in proposito?

Il cammino è stato lungo. È iniziato, infatti, a cavallo fra il 1996 e il 1997. E si è sostanzialmente concluso soltanto nel maggio 2004, con la sottoscrizione, da parte della Conferenza dei Rettori delle Università Italiane (CRUI) e di dieci società scientifiche, del codice di deontologia e buona condotta¹⁶: l'ultimo tassello, ma imprescindibile, perché la disciplina più favorevole dettata per i trattamenti per scopi scientifici diventa operante soltanto quando il codice è efficace.

Ho riassunto le tappe salienti dell'accidentato itinerario nel Prospetto 1. Vedo di percorrerle rapidamente.

La prima tappa – l'iniziale normativa generale sulla *privacy* – si colloca agli sgoccioli del 1996 e ha dei tratti inusuali. Nella storia della legislazione italiana non è frequente trovare leggi approvate il 31 dicembre. Tra Natale e Capodanno anche i parlamentari e il Presidente della Repubblica, giustamente, fanno vacanza. Singolare, poi, è imbattersi in due leggi sulla stessa materia approvate nello stesso giorno. Ciò è accaduto il 31 dicembre 1996, con le legge n. 675 sulla tutela delle per-

¹⁶ Di proposito parlo di conclusione ‘sostanziale’. Per il completamento formale, infatti, sono richiesti alcuni ulteriori passi, ancora in corso quando questo testo è stato scritto. Segnatamente, il codice deve essere “*pubblicato nella ‘Gazzetta Ufficiale’ ... a cura del Garante, [previa verifica del]la conformità alle leggi e ai regolamenti anche attraverso l'esame di osservazioni di soggetti interessati?*” (Decreto legislativo n. 196/2003, art. 12).

sone rispetto al trattamento dei dati personali e con la legge n. 676 di delega al Governo in materia.

Il perché di questa peculiarità è presto detto. La legge n. 675/1996 è stata approvata con ritmi concitati, sotto l'incalzare di una scadenza – il 31 dicembre 1996 per l'appunto – che, in sostanza, non si poteva eludere. La legge in questione è, infatti, una sorta di Giano bifronte. Afferma e regola il diritto alla *privacy*. Ma, insieme, disciplina il flusso di comunicazioni fra le forze di polizia dei diversi paesi, condizione perché nell'ottobre successivo l'Italia potesse entrare nell'area di Schengen.

PROSPETTO 1

Iter della normativa sul trattamento di dati personali per scopi statistici e scientifici

Data	Disposizioni legislative	Iniziative e proposte della comunità scientifica
12.1996	Legge n. 675 sulla tutela delle persone e di altri soggetti rispetto al trattamento di dati personali Legge n. 676 di delega al Governo in materia di trattamento di dati personali	
05.1998		<i>Workshop del CSS (Consiglio italiano per le Scienze Sociali) su "Ricerca scientifica, accesso ai dati personali e tutela della riservatezza"</i>
09.1998		<i>Documento-appello del CSS sul tema (adesioni: 20 società e istituti; scientifiche; 70 studiosi)</i> <i>Formazione, presso il CSS, di un gruppo di lavoro di rappresentanti di società scientifiche</i>
05-07.1999		<i>Proposte del gruppo di lavoro per il decreto leg.vo (+ interventi di singoli sulla stampa)</i>
07.1999	Decreto lvo n. 281 in materia di trattamento di dati personali per finalità storiche, statistiche e di ricerca scientifica	
02.2000	Provvedimento del Garante che promuove l'adozione dei pertinenti codici di deontologia e di buona condotta	
03.2000		<i>CSS e società scientifiche presentano la proposta di codice di deontologia</i>
07-09.2000		<i>Riunioni presso l'ufficio del Garante</i>
03.2001	Codice di deontologia per il trattamento di dati personali per scopi storici Legge n. 127 sul differimento del termine per l'esercizio della delega previsto dalla legge n. 676/1996	
05-06.2001		<i>Ulteriori riunioni presso l'ufficio del Garante</i>
04-05.2002		<i>Ulteriori riunioni presso l'ufficio del Garante</i>
07.2002	Codice di deontologia per il trattamento di dati personali a scopi statistici e di ricerca scientifica effettuati nell'ambito del Sistan	
06.2003	Testo unico in materia di protezione dei dati personali (entrato in vigore il 01.01.2004)	
02-04.2004		<i>Ulteriori riunioni presso ufficio del Garante</i>
05.2004		<i>CRUI e dieci società scientifiche sottoscrivono il codice di deontologia per il trattamento di dati personali per scopi statistici e scientifici</i>

Nel merito, la disciplina dettata nel 1996 si segnala per un'enunciazione del diritto alla *privacy* particolarmente forte: esso viene collegato alle "*libertà fondamentali, [alla] dignità delle persone fisiche*" e si configura come "*diritto alla riservatezza e all'identi-*

tà personale” (art. 1, comma 1). Procede poi con la puntuale specificazione di diritti, obblighi e divieti, in parecchi casi più circostanziata e stringente della direttiva comunitaria.

Non c'è di che sorprendersi, dunque, se la normativa del 1996, anche per lo scotto pagato alla fretta, non sempre è felice: presenta imprecisioni e angolosità, si distingue per bardature regolatorie, eccede – è stato scritto – “*sul piano dello zelo e anche del buon senso*”¹⁷. Un esempio ben si presta per illustrare questo impasto di imprecisione ed eccessivo zelo. La legge definisce 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 compreso un numero di identificazione personale*” (art. 1, comma 2, lettera c); l'enfasi è mia). L'imprecisione è palese: si chiama in causa il numero di identificazione personale quasi fosse un'estrema risorsa – o comunque una tra le tante utilizzabili –, e non l'ovvio strumento principe per assicurare l'identificazione. L'eccesso di zelo sta nel fatto che, per decidere dell'identificabilità di una persona, si fa riferimento a “*qualsiasi altra informazione*”: formulazione larghissima, priva del richiamo ai mezzi che possono essere utilizzati (mentre, come abbiamo visto, quelli dei ‘mezzi ragionevoli’ è il criterio-guida delle raccomandazioni internazionali e della disciplina dell'UE), rivelatrice di una propensione a dilatare in maniera impropria la nozione di dato personale e quindi la sfera di applicazione della legge¹⁸.

È, del resto, lo stesso legislatore a mostrarsi consapevole, da subito, di manchevolezze della normativa. Lo testimonia il fatto che la affianca con una legge, la n. 676/1996, la quale delega il Governo a integrare e correggere il testo appena approvato. Tra gli specifici oggetti della delega vi è “*il trattamento di dati personali utilizzati a fini storici, di ricerca e di statistica*” (art. 1, lettera a)).

Nella comunità scientifica prende avvio la riflessione. Inizialmente statistici, scienziati sociali, biomedici ed epidemiologi, storici e archivisti procedono ciascuno per proprio conto, con l'elaborazione di prime, specifiche proposte. Nel maggio del 1998, uno *workshop* promosso dal Consiglio italiano per le Scienze Sociali (CSS) è l'occasione per un più largo confronto e per una significativa convergenza. Essa si traduce in due iniziative: da un lato l'elaborazione di un documento di “*considerazioni e raccomandazioni*” su *privacy*, statistica e ricerca, sottoscritto da venti società e istituzioni scientifiche e da oltre settanta studiosi (CSS, 1998); dall'altro lato la costituzione di un gruppo di lavoro incaricato di seguire l'*iter* di elaborazione del decreto legislativo.

¹⁷ Ciancarella (2003).

¹⁸ Altri esempi sono nelle deroghe che la legge n. 675/96 prevede per l'uso di dati personali a fini statistici e di ricerca scientifica – una all'obbligo del consenso dell'interessato (art. 12, comma 1, lettera d)), l'altra al divieto di comunicazione e diffusione (art. 21, comma 4, lettera a)) –, accompagnandole peraltro con l'aggiunta che esse valgono “*limitatamente a dati anonimi*”. Qui l'eccesso di zelo è solare e sfocia in disposizioni che sono *tout court* irrilevanti. L'intera legge, infatti, riguarda il trattamento di dati personali e quindi, semplicemente, non si applica a dati anonimi, qualunque sia lo scopo per il quale vengono usati. Queste imprecisioni, e parecchie altre, sono state corrette dalla legislazione successiva.

Nella fase iniziale di preparazione del decreto non mancano rischi di sordità e incomprensioni, che chiamano in causa la diffusa sottovalutazione delle esigenze e della funzione sociale della ricerca scientifica, alla quale già ho accennato. Ci si misura con la propensione per un accesso ai dati fortemente ristretto, per estesi adempimenti burocratici, per diffusi controlli amministrativi, senza che ne siano considerati i probabili effetti negativi sulla ricerca e sui suggerimenti che essa può offrire per politiche¹⁹. L'attività del gruppo lavoro e una serie di interventi di studiosi sulla stampa²⁰ valgono a superare, almeno in parte, questi rischi. Si fa strada una più equilibrata percezione dei termini del problema. E il confronto di opinioni col Garante, col Governo, con le Commissioni parlamentari si fa positivo.

Il 30 luglio 1999 – ancora giusto entro il tempo massimo (la delega scadeva il giorno dopo) – il Governo approva il decreto legislativo n. 281 sul trattamento di dati personali “*per finalità storiche, statistiche e di ricerca scientifica*”. Il cammino aperto due anni e mezzo prima dalla legge delega è, tuttavia, lungi dall'essere compiuto. Si completa una tappa importante, per molti versi decisiva. Ma non l'ultima. La logica che informa il decreto è infatti la seguente: porre alcuni punti fermi, sia nel senso di dettare norme di indirizzo sia soprattutto nel senso di allargare le maglie della disciplina generale sulla *privacy*; e, insieme, rinviare gran parte della disciplina positiva ai codici deontologici, che dovranno essere sottoscritti da società e istituzioni scientifiche.

Il decreto è in discreta sintonia con le raccomandazioni internazionali. Le scelte fatte possono essere riassunte in quattro punti. Innanzitutto, detta una disciplina generale per l'insieme delle attività di statistica e di ricerca, senza confini angusti e, opportunamente, senza distinguere se chi fa ricerca sia un soggetto pubblico o privato. Entro questo quadro comune trovano poi spazio, com'è ragionevole, norme particolari per due specifici campi: la statistica ufficiale, cioè a dire l'Istat e gli altri soggetti del Sistema statistico nazionale (Sistan); la ricerca storica e gli Archivi di Stato.

Un secondo aspetto di rilievo attiene all'ambito di applicazione della normativa. Con riguardo ai trattamenti per scopi statistici e di ricerca, che cosa si intende per “*dati personali*”? L'interrogativo non è di poco conto, per quanto ho già detto sulla definizione estremamente lata di identificabilità accolta dalla legge n. 675/1996. Ebbene, il decreto è netto in proposito: sono considerati tali i dati che con-

¹⁹ L'evidenza sugli effetti della legislazione in tema di *data protection* sulla ricerca è scarsa. (La legislazione è recente. Inoltre, i suoi effetti dipendono largamente dai modi in cui è posta in essere e fatta osservare.) A mia conoscenza, vi è un solo studio in ambito medico, condotto negli Stati Uniti, mirato a determinare gli effetti della legislazione degli Stati che richiede il previo consenso informato del paziente per l'accesso a *record* medici da parte di ricercatori esterni. Queste le conclusioni: “*Legislation requiring patient informed consent to gain access to medical records for a specific study was associated with low participation and increased time to complete the observational study. Efforts to protect patient privacy may come into conflict with the ability to produce timely and valid research to safeguard and improve public health*” (McCarthy *et al.*, 1999, p. 417).

²⁰ Tra maggio e settembre 1999 ne hanno scritto, su *Repubblica*, *La Stampa* e soprattutto *Il Sole-24 Ore*, Tito Boeri, Mario Braga, Agar Brughiavini, Valerio Castronovo, Cesare Cislighi, Bruno Contini, Guido Martinotti, Chiara Saraceno, Ugo Trivellato e Guglielmo Weber.

sentono l'identificazione di una persona “*utilizzando mezzi ragionevoli*” (art. 10, comma 6, lettera c)). La nozione di dati personali risulta così sensatamente circoscritta. Ciò è di vitale importanza per le attività di ricerca, che – l'ho sottolineato più volte – necessitano sì di trattare insiemi di dati individuali, ma per un uso ‘collettivo’, senza alcun bisogno di riferire i risultati a (tanto meno di diffondere le informazioni su) una singola persona. Buona parte delle basi di dati utilizzate nella ricerca – in ambito economico-sociale così come in quello biomedico-epidemiologico – possono essere considerate anonime: con il duplice vantaggio di non gravarle di adempimenti burocratici e di concentrare la tutela della *privacy* sulle aree davvero esposte a rischi di identificabilità.

Quanto al merito della normativa sulla ricerca, il decreto è decisamente parco. Non fornisce una disciplina articolata. Imbocca invece una strada diversa, che poggia su due pilastri.

- Per un verso, allenta le norme generali sulla *privacy* in varie direzioni: eliminando vincoli, dettando disposizioni più flessibili, prevedendo che dati originariamente raccolti per altri fini possono essere trattati per scopi di ricerca.
- Per un altro verso, affida un'ampia funzione ai codici deontologici. Essi dovranno specificare la disciplina alla luce di alcune indicazioni e, in generale, delle raccomandazioni del Consiglio d'Europa.

Ho sommariamente riassunto queste disposizioni, da un lato di allentamento dei vincoli e dall'altro di rinvio ai codici, nel Prospetto 2, al quale rimando. I due ordini di indicazioni sono naturalmente collegati. Il rinvio alla disciplina dei codici presuppone infatti che vi siano significativi gradi di libertà. Com'è ovvio, ciò comporta anche che una parte non trascurabile delle questioni rimane aperta e si decide appunto sul terreno dei codici deontologici.

Nel merito, vi è tuttavia una disposizione – o meglio, innanzitutto un silenzio – del decreto sul quale si impone una riflessione. Malgrado le sollecitazioni delle società scientifiche, esso non rimuove un serio intralcio che si frappone alla ricerca scientifica su materie e “*dati sensibili*”, in primo luogo quelli sullo stato di salute e la vita sessuale. Non modifica infatti la regola del silenzio-rigetto, stabilita nella legge n. 675/1996 in via generale per i dati sensibili: una scelta, questa, che è francamente difficile condividere e della quale neppure è agevole comprendere la *ratio*²¹. Il decreto apre, tuttavia, uno spiraglio. Per il trattamento di dati sensibili per scopi di ricerca, è esplicitamente richiamata la possibilità per il Garante di rilasciare autorizzazioni relative a categorie di titolari o di trattamenti: di indole generale, preventive.

²¹ L'obbligo di ottenere l'autorizzazione previa del Garante è accompagnato dalla disposizione per cui, se entro 45 giorni (con la normativa attuale, erano 30 nella legge n. 675/1996) il Garante non risponde, la richiesta si intende respinta. Ora, l'istituto del silenzio-rigetto è a dir poco inopportuno in via generale, per l'insieme dei rapporti fra cittadini – e imprese – e amministrazioni dello Stato (tutte, Autorità garanti incluse). Che le amministrazioni dello Stato siano tenute a motivare le loro decisioni entro tempi certi, è regola basilare di un moderno stato di diritto, che comincia a essere introdotta in maniera diffusa anche nel nostro paese. Né si riesce a vedere come tra le circostanze particolari, che potrebbero giustificare eccezioni a questa regola, possano sensatamente rientrare i trattamenti di dati sensibili a fini di ricerca.

PROSPETTO 2

Sinossi delle disposizioni salienti del decreto legislativo n. 281/1999 sui “trattamenti a scopi di statistica o di ricerca scientifica”

Disposizioni o deroghe	Rinvio ai codici deontologici
<i>(A) Generali</i>	
Uso lecito dei dati personali, per soli scopi statistici e di ricerca scientifica	
Compatibilità del trattamento per scopi di statistica/ricerca (i) di dati raccolti o trattati per altri scopi e (ii) oltre il periodo necessario a questi altri scopi	
Presupposti e procedimenti per documentare che i trattamenti siano svolti per scopi di statistica/ricerca	<i>Individuati dai codici</i>
Insieme dei mezzi che possono essere ragionevolmente utilizzati per identificare l'interessato	<i>Individuati dai codici</i>
<i>(B) Specifiche</i>	
Notificazione semplificata	<i>In conformità [anche] ai codici*</i>
Informativa all'interessato con scopi di statistica/ricerca chiaramente determinati e resi noti	
Informazioni a interessati per dati raccolti presso terzi	<i>Individuate nei codici</i>
Esclusione del consenso	<i>Nel rispetto dei codici (con individuazione di garanzie da osservare)</i>
Regole di correttezza nella raccolta dei dati	<i>Individuate dai codici</i>
Garanzie per ‘trattamenti intermedi’ (trattamento di dati identificativi, conservazione dei dati, sicurezza, diritto di rettifica)	<i>Individuate dai codici</i>
Impegno al rispetto di regole di condotta da parte di incaricati di trattamenti	<i>Individuate dai codici</i>
Cessato il trattamento, conservazione o cessione dei dati ad altro titolare	<i>In conformità [anche] ai codici</i>
Comunicazione o diffusione	<i>Effettuate nel rispetto dei codici (con individuazione di garanzie da osservare)</i>
Trasferimento di dati all'estero	<i>Effettuato nel rispetto dei codici</i>
Per dati sensibili:	
- consenso con modalità semplificate	<i>Individuate dai codici</i>
- autorizzazione tipo del Garante	

* L'obbligo di notificazione è stato successivamente eliminato, salvo che per trattamenti riguardanti particolari categorie di dati (decreto legislativo n. 196/2003, art. 37).

Un ultimo aspetto, e di non poco conto, occorre infine mettere in luce: il ruolo, la natura che assumono i codici deontologici. Già nella legge n. 675/1996 essi appaiono come uno strano animale: un ircocervo o un centauro, a mezzo fra strumento di autodisciplina di un'associazione scientifica o professionale e fonte normativa. Il decreto legislativo n. 281, dilatando la funzione dei codici deontologici, li fa a tutti gli effetti ‘codici’: testi di legge, sia pure ‘secondaria’²². Capacità e

²² Significativamente, i codici deontologici già pubblicati – fra i quali quello per i trattamenti di dati personali a scopi statistici e di ricerca scientifica effettuati nell'ambito del Sistan – fanno parte, come allegati, del testo unico in materia di protezione dei dati personali (decreto legislativo n. 196/2003), di cui dirò tra poco. La scelta si iscrive nella tradizione nostrana di una legislazione estesa, minuziosa. Non era, tuttavia, obbligata. Sull'argomento, le raccomandazioni internazionali e la direttiva europea sono piuttosto aperte. Assegnano infatti (o, comunque, permettono di assegnare) una significativa autonomia alle associazioni professionali nella redazione dei codici deontologici,

responsabilità di autodisciplina di istituzioni, società scientifiche e ricercatori sono così mortificate, in favore di una normativa più penetrante²³, sia pure elaborata col loro coinvolgimento. Di conseguenza, anche il grado di intervento del Garante nella definizione dei codici viene potenziato, al di là di quanto non dica, sobriamente, l'art. 6 del decreto²⁴.

5. LA PROPOSTA DI CODICE DEONTOLOGICO: SCELTE BASILARI E NODI IRRISOLTI

A valle del decreto legislativo, il gruppo di lavoro costituito presso il CSS, che riunisce ormai i rappresentanti di nove tra le maggiori società scientifiche²⁵, prosegue la propria attività con un nuovo obiettivo: elaborare la proposta di codice deontologico per i trattamenti di dati personali per scopi statistici e di ricerca scientifica²⁶. Questo impegno, condotto in maniera coordinata con l'Istat e con storici e archivisti – a loro volta attivi nel preparare parallele proposte di codici –, si conclude nel marzo 2000 (dunque entro i termini previsti), con la presentazione della proposta di codice al Garante.

La proposta è discussa con l'ufficio del Garante in varie occasioni, tra luglio del 2000 e maggio del 2002. Alle società scientifiche che l'hanno presentata si affiancano la CRUI e, successivamente, l'ASSIRM (Associazione tra Istituti di Ricerche di Mercato, Sondaggi di Opinione, Ricerca Sociale).

Sono chiariti molti punti, vengono concordate revisioni di parecchi altri. Ma restano alcuni nodi che non si riesce a districare. La connotazione normativa dei codici deontologici – di cui ho appena detto –, combinata con perduranti rigidità del quadro legislativo, è la principale responsabile di queste difficoltà. E dello stallo che, conseguentemente, si determina.

che dunque restano (o possono restare) tali: dei “*codes of professional ethics*”. In particolare, la Direttiva 95/46/CE lascia gradi di libertà agli Stati membri in materia: “*Gli Stati membri ... incoraggiano l'elaborazione di codici di condotta ... Dispongono che le associazioni professionali e gli altri organismi rappresentativi di altre categorie di responsabili del trattamento, che hanno elaborato i progetti di codice [deontologico] ..., possono sottoporli all'esame dell'autorità nazionale*” (art. 27; l'enfasi è mia).

²³ Per misurare la distanza di questa scelta rispetto a legislazioni più sobrie – quella inglese, ad esempio –, è istruttivo considerare uno specifico caso: le disposizioni sulla possibilità, per soli scopi scientifici, di conservare dati personali per periodi più lunghi di quanto necessario per gli scopi originari. Come prevedere i possibili, futuri usi di una base di dati personali per scopi scientifici? O meglio, formulando l'interrogativo nella maniera appropriata: come escludere che una base di dati personali possa, un domani, tornare utile per scopi scientifici? In Italia, i codici deontologici si addentrano in una problematica casistica, che finisce poi includere – e non potrebbe essere altrimenti – “*altri casi in cui ciò risulti essenziale e adeguatamente documentato per le finalità perseguite*” (così, ad esempio, il codice di deontologia per i trattamenti nell'ambito del Sistan). La norma inglese è essenziale, e convincente nella sua liberalità: “*Personal data which are processed only for research purposes ... may be kept indefinitely*” (UK Data Protection Act 1998, Section 33 (3); Lloyd, 1998).

²⁴ “*Il Garante [ne] promuove la sottoscrizione I codici sono pubblicati nella Gazzetta Ufficiale... a cura del Garante*”.

²⁵ Associazione Italiana di Epidemiologia, Associazione Italiana di Sociologia, Consiglio italiano per le Scienze Sociali, Società Italiana degli Economisti, Società Italiana di Biometria, Società Italiana di Demografia Storica, Società Italiana di Igiene Medicina Preventiva e Sanità Pubblica, Società Italiana di Statistica, Società Italiana di Statistica Medica ed Epidemiologia Clinica.

²⁶ Nel seguito, spesso userò semplicemente il termine ‘codice’.

Della proposta di codice, così come si configura alla metà del 2002, alla fine di questa prima serie di incontri col Garante, mi limito a delineare i tratti essenziali: il suo impianto, le scelte salienti che la informano, gli intoppi residui. Sarò sintetico, perché la proposta si ispira alle raccomandazioni internazionali, sulle quali mi sono già diffusamente soffermato.

La proposta di codice²⁷ si apre con un preambolo, che si propone di metterne in evidenza finalità e logica ispiratrice. Due sono i passaggi di rilievo.

In primo luogo, la funzione del codice è individuata nel temperamento di ‘valori’ e ‘diritti’ diversi. A fronte del diritto alla riservatezza, sono messe in luce “*le necessità della statistica e della ricerca scientifica*” e le ragioni che ne sono alla base: “*il principio della libertà della ricerca, costituzionalmente garantito, e le esigenze di sviluppo della ricerca per migliorare le condizioni della società*”.

In secondo luogo, Vi è un’indicazione in favore del principio dell’“*universal access to all ‘bona fide’ analytical users*”. Per dati che siano di proprietà pubblica, o comunque risultino da un’attività sostenuta da risorse pubbliche (più precisamente, “*che siano depositati in archivi pubblici o che siano stati trattati sulla base di finanziamenti pubblici*”), valgono le regole di non discriminazione e di imparzialità nell’accesso nei confronti di tutti i ricercatori.

Il nucleo principale della proposta di codice è dedicato all’ambito di applicazione e ai principi generali. Il primo, cruciale punto che viene affrontato è, per dirla con il decreto legislativo n. 281/1999, la specificazione dei “*presupposti e [de]i procedimenti per documentare e verificare che i trattamenti siano svolti per idonei ed effettivi scopi statistici e di ricerca*” (art. 10, comma 6, lettera a)). La scelta fatta è che debbano valere due ordini di presupposti, riferiti rispettivamente al soggetto e alla materia.

- (a) In ragione del soggetto, deve trattarsi di un’università o altro ente di ricerca o società scientifica, oppure di un singolo ricercatore che operi in un’università o ente di ricerca o sia socio di una società scientifica²⁸.
- (b) In ragione della materia, si richiede al ricercatore responsabile di predisporre un progetto di ricerca, redatto secondo gli standard dei protocolli di ricerca del pertinente settore disciplinare. Il progetto è depositato presso l’ente titolare della ricerca (o la società scientifica di appartenenza), che ne cura la conservazione per cinque anni dalla conclusione programmata della ricerca.

L’ambito di applicazione del codice è inoltre precisato con riguardo a un tema particolarmente delicato: quello del trattamento di dati sullo stato di salute. La linea di demarcazione che viene tracciata distingue due insiemi di “*trattamenti per scopi statistici e scientifici*” di dati personali idonei a rivelare lo stato di salute: quelli “*connessi con attività di tutela della salute svolte da esercenti le professioni sanitarie od organi-*

²⁷ Vale la pena di sottolineare che in questa sezione mi riferisco alla (e cito dalla) proposta definita dalle società scientifiche e dalla CRUI nel maggio 2002, non al codice sottoscritto dalle stesse il 13 maggio 2004. L’impianto del testo e il contenuto delle disposizioni restano, peraltro, in larga misura invariati. D’altra parte, dal confronto col testo finale del codice – riportato in questo stesso numero della rivista –, è possibile cogliere varianti redazionali anche non trascurabili, di interesse per valutare gli affinamenti apportati.

²⁸ Ovviamente, le università, gli enti e le società scientifiche “*assicurano il rispetto del codice fra tutto coloro che sono coinvolti nel trattamento dei dati personali realizzati nell’ambito delle ricerche*”, mentre i ricercatori “*si obbligano al rispetto del codice in ogni fase dei trattamenti di dati personali*” svolti in tale ambito.

smi sanitari” e i restanti trattamenti. Il codice deontologico si applica a questi ultimi, e non ai primi²⁹. La *ratio* della disposizione è evidente.

- Per il primo insieme di trattamenti la finalità di tutela della salute ha un rilievo dominante. Conseguentemente, per essi vale la disciplina dettata dal Titolo V del testo unico in materia di protezione di dati personali (emanato il 30 giugno 2003, del quale dirò nella sez. 6³⁰), relativo appunto al “*Trattamento di dati personali in ambito sanitario*”. Essa specifica, tra l’altro, che i trattamenti in questione riguardano “*dati e operazioni indispensabili per perseguire una finalità di tutela della salute o dell’incolumità fisica dell’interessato ... [o di] un terzo o [del]la collettività*” (decreto legislativo n. 196/2003, art. 76, comma 1).
- L’altro insieme di trattamenti è definito per residuo. Esso è caratterizzato, dunque, dal fatto di non essere strettamente funzionale a – o una mera conseguenza di – azioni volte alla tutela della salute (e, ovviamente, di rispondere a finalità di ricerca scientifica). Lo spettro di tali trattamenti è ampio. Si estende ad attività di ricerca medica, biomedica ed epidemiologica, le quali, naturalmente, possono essere svolte anche da ricercatori che non esercitano la professione medica. È a questo insieme di trattamenti che si applica il codice deontologico, con le disposizioni sui dati sensibili e – come vedremo – con norme particolari.

Un secondo ordine di disposizioni generali attiene alla specificazione del criterio dei “*mezzi che possono essere ragionevolmente utilizzati*” per l’identificazione. La scelta è nel senso di disposizioni articolate, ma comunque sufficientemente larghe e flessibili, non ancorate a puntuali standard tecnici, che rischierebbero di diventare rapidamente obsoleti a fronte degli sviluppi nella tecnologia del trattamento delle informazioni. Inoltre, si pone l’accento sul fatto che gli avanzamenti in questo campo non agiscono solo nel senso di aumentare i rischi di identificabilità. Sistemi di sicurezza e *software* apposito³¹ offrono anche crescenti possibilità di prevenzione e controllo dei tentativi illeciti di identificazione³². È il caso di ricordare, poi, che vi è un altro alleato contro i rischi di identificazione, potente ed efficace, anche se inopinatamente trascurato: le incompletezze³³ e gli errori nei dati³⁴.

²⁹ Alle attività di tutela della salute svolte da esercenti le professioni sanitarie sono assimilate “*attività comparabili in termini di significativa ricaduta personalizzata sull’interessato*”. Anche ai trattamenti connessi a queste ultime, dunque, il codice deontologico non si applica.

³⁰ La denominazione esatta del decreto legislativo n. 196/2003 è “*Codice in materia di dati personali*”. Per evitare possibili confusioni, e coerentemente con la natura del decreto, qui e altrove mi riferisco ad esso col termine ‘testo unico’, mentre continuo a riservare il termine ‘codice’ per designare il codice deontologico per i trattamenti di dati personali per scopo statistici e scientifici.

³¹ Mi riferisco al *software* in materia di *database protection, audit systems for statistical databases e statistical data confidentiality methods*. Lo stato dell’arte in materia è ben documentato in Eurostat (1999). Vedi anche Fienberg e Willenborg (1999) e McClean (1998).

³² Queste disposizioni, e quelle della terza parte, sono simili, sovente uguali, alle corrispondenti norme del codice deontologico per i trattamenti di dati personali per scopi statistici nell’ambito del Sistan.

³³ Tipicamente, le incompletezze si manifestano nella non partecipazione di alcune unità statistiche all’indagine (*unit nonreponse*) oppure nella mancata risposta ad alcune domande, potenzialmente utili a fini di identificazione, da parte di unità statistiche che vi partecipano (*item nonresponse*).

³⁴ Wagner (2003, p. 134) riporta i risultati di un illuminante studio di caso: “*Experiments with the German Microcensus (which is a survey like the Current Population Survey in the U.S. [e quale l’indagine sulle*

La seconda parte del codice disciplina condizioni particolari per la liceità dei trattamenti per scopi scientifici, nell'ordine in materia di informativa all'interessato, consenso, comunicazione e diffusione dei dati, trattamento di dati sensibili – in generale e specificamente per la ricerca biomedica ed epidemiologica –. La logica che le ispira è semplice: si tende a dare criteri chiari, senza tuttavia bloccarli in formule rigide, che risulterebbero contraddittorie con le necessità di adattamento alle diverse, e mutevoli, tecniche di raccolta dei dati. Un esempio può bastare: il consenso per il trattamento di dati sensibili deve sì essere “*free, informed and unambiguous*”³⁵ (salve motivate deroghe per indagini biomediche o epidemiologiche, sulla quali tornerò tra breve), ma non deve essere necessariamente manifestato in una forma particolare, per iscritto. In una società dove la comunicazione si realizza in misura crescente tramite canali diversi dallo scritto, e dove la ricerca sociale ed epidemiologica avviene tramite *surveys* che poggiano sempre più sul telefono, su modalità *computer assisted* e su *Web*, è del tutto ragionevole chiedere semplicemente che il consenso dell'interessato sia “*documentato*” a cura di chi raccoglie le informazioni.

La terza parte del codice, infine, dà regole da seguire nella raccolta dei dati, detta disposizioni in materia di conservazione dei dati – ispirate a quello che ho chiamato ‘principio di parsimonia’ – e di misure di sicurezza, enuncia le regole propriamente deontologiche.

Ho lasciato per ultimi i punti sui quali, verso la metà del 2002, il dialogo col Garante si arena, e che concordemente vengono accantonati in vista di auspicabili novità legislative (nell'ambito del testo unico sul trattamento di dati personali, in corso di definizione, del quale dirò tra breve). Le questioni che restano aperte sono, tra l'altro, dei casi esemplari, che meglio di qualsiasi argomentazione di indole generale si prestano per illustrare le difficoltà che incontriamo (fortunatamente, almeno in parte possiamo dire: che abbiamo incontrato) nel far valere le buone ragioni della ricerca scientifica. Eccole.

- (a) Si conviene che certi dati personali possano essere raccolti tramite un rispondente *proxy*. (Un esempio tipico sono le informazioni individuali di natura demografica e sociale rilevate in *surveys* sulle famiglie.) Si osserva, tuttavia, che ciò non esime il ricercatore dall'obbligo di legge di informare l'interessato. D'altra parte, è palese che quest'obbligo – così come formulato – ha tratti di irragionevolezza: se si affrontano i tempi ed i costi per raggiungere l'interessato ed informarlo, non vi è infatti alcun motivo per non rilevare dallo stesso i dati di interesse³⁶.

forze di lavoro in Italia: NdA]) *show that the risk of effective disclosure, matching external information with a survey record, is close to zero ... Furthermore, the probability of a 'false match' is about 80 percent. This means that even in the rear event of a match of two records, the probability that these records are not from the same person is very high*".

³⁵ Così la Raccomandazione R (97)18 del Consiglio d'Europa. La direttiva comunitaria richiede che il consenso sia “*manifestato in maniera inequivocabile*” (Direttiva n. 95/46/CE, art. 7, lettera a)).

³⁶ L'accettazione del rispondente *proxy*, in generale meno affidabile e quindi un *second best* per il ricercatore, sta proprio nel contenimento dei tempi e costi della *survey*, che si dilaterrebbero di molto se si dovessero intervistare tutti i membri della famiglia.

- (b) Si conviene che possano essere trattati per scopi scientifici dati personali raccolti per altri scopi, tipicamente amministrativi (come previsto, del resto, dall'art. 3 del decreto legislativo n. 281/1999). Anche in questo caso, tuttavia, si fa presente che vige l'obbligo di informare gli interessati. Certo, si riconosce che gli interessati possono essere migliaia, centinaia di migliaia, milioni di individui: con gli ovvi problemi organizzativi e di mezzi che l'obbligo di informarli comporta. Non si riescono a individuare, tuttavia, soluzioni plausibili entro la cornice normativa vigente.
- (c) In tema di ricerca medica, biomedica ed epidemiologica – che ovviamente coinvolge dati sensibili e, d'altra parte, può ben essere svolta anche da ricercatori che non esercitano la professione medica –, vi sono indagini per le quali può essere necessario prescindere dal consenso dell'interessato. Ciò può accadere vuoi per l'impossibilità di informare l'interessato per motivi organizzativi, vuoi per l'esigenza di **non** informarlo per motivi etici (rispettare l'ignoranza dell'interessato sulla sua condizione) o metodologici (non comunicare al soggetto le ipotesi sotto studio e/o la sua posizione entro il disegno della ricerca). In questi casi, si propone che la ricerca possa essere legittimamente svolta, purché abbia ottenuto la previa, motivata approvazione del pertinente Comitato etico del Servizio sanitario nazionale e sia stata autorizzata dal Garante con un'autorizzazione di indole generale. A questa proposta, della quale pure si riconosce la sensatezza, si oppone l'inammissibilità perché *contra legem*³⁷.

6. IL TESTO UNICO IN MATERIA DI PROTEZIONE DEI DATI PERSONALI E LA DEFINIZIONE DEL CODICE DENTOLOGICO

Mentre si svolge il confronto col Garante sul codice deontologico, con la legge n. 127/2001 il Governo è delegato ad emanare un testo unico in materia di protezione dei dati personali. Il testo unico è emanato col decreto legislativo n. 196 del 30 giugno 2003 ed entra in vigore il 1° gennaio 2004.

Oltre a unificare e razionalizzare le disposizioni di più di una decina di leggi e decreti succedutisi dal 1996 al 2002, esso integra e corregge la normativa. Per la parte attinente ai trattamenti per scopi statistici o scientifici, ciò avviene in maniera tutto sommato rispondente alle principali preoccupazioni e ai suggerimenti della comunità scientifica. Mi limito qui a riportare le norme rilevanti per le questioni aperte di cui ho appena detto.

- (a) In tema di rispondente proxy, il testo unico prevede che, “*quando specifiche circostanze individuate dai codici [deontologici] siano tali da consentire ad un soggetto di rispondere in nome e per conto di un altro, in quanto familiare o convivente, l'informativa*

³⁷ Le sole deroghe esplicitamente previste alla necessità del consenso dell'interessato sono nell'art. 85, comma 1, del decreto legislativo n. 282/1999 sulla riservatezza dei dati personali in ambito sanitario: “*qualora la ricerca sia prevista da un'espressa previsione di legge o rientri nel programma di ricerca biomedica o sanitaria di cui all'art. 12-bis del decreto legislativo 30 dicembre 1992, n. 502, e successive modificazioni ed integrazioni*”.

- all'interessato può essere data anche per il tramite del soggetto rispondente*" (art. 105, comma 3).
- (b) In tema di trattamento per scopi scientifici di dati raccolti per altri scopi, stabilisce che *"l'informativa all'interessato non è dovuta quando richieda uno sforzo sproporzionato rispetto al diritto tutelato, se sono adottate le idonee forme di pubblicità individuate dai codici [deontologici]"* (art. 105, comma 4).
- (c) In tema di ricerca medica, biomedica ed epidemiologica, stabilisce che *"il consenso dell'interessato per il trattamento dei dati idonei a rivelare lo stato di salute ... non è necessario quando a causa di particolari ragioni non è possibile informare gli interessati e il programma di ricerca è oggetto di motivato parere favorevole del competente comitato etico a livello territoriale ed è autorizzato dal Garante ai sensi dell'art. 40"*, cioè con un'autorizzazione generale (art. 110, comma 1).

Si danno dunque le condizioni per un conclusivo sforzo cooperativo delle società scientifiche e della CRUI – tramite il gruppo di lavoro del quale ho ripetutamente detto – e del Garante. Detto altrimenti, alla luce del nuovo quadro normativo definire e sottoscrivere il codice deontologico senza ulteriori ritardi è un obiettivo ragionevole. Ed è un obiettivo pressante: per permettere, finalmente, di utilizzare la normativa *ad hoc*, passabilmente agile, per i trattamenti di dati personali per scopi scientifici.

Tra febbraio e aprile 2004 vi è una serie di incontri tra il gruppo di lavoro e l'ufficio del Garante, con serrati approfondimenti che portano alla bozza conclusiva. Il testo finale del codice viene definito in un'ultima riunione presso il Garante il 13 maggio 2004, e sottoscritto dalla CRUI e dalle dieci società scientifiche che hanno partecipato alla sua elaborazione.

7. UNA CONSIDERAZIONE CONCLUSIVA

L'entrata in vigore del codice è una cruciale condizione permissiva perché la nostra attività di ricercatori si possa svolgere senza irragionevoli ostacoli, cioè a dire potendo produrre e analizzare basi di dati individuali, vitali per la ricerca economica e sociale così come per quella biomedica ed epidemiologica, entro una ragionevole cornice normativa.

Tra l'altro, essa ha anche un'importante effetto indiretto: rende pienamente operanti le disposizioni del parallelo codice deontologico per i trattamenti per scopi statistici effettuati nell'ambito del Sistan (vigente dall'agosto 2002) sulla comunicazione di dati personali al di fuori del Sistan. Le due modalità di comunicazione lì previste, *"nell'ambito di specifici laboratori"*³⁸ e *"nell'ambito di progetti congiunti"* (art. 7, commi 2 e 3), sono infatti riservate ai soggetti – ricercatori di università o enti di ricerca e soci di società scientifiche – ai quali si applica il codice deontologico in questione.

³⁸ Il termine "laboratori" designa quelli che sono correntemente noti come *safe settings*: aree protette di deposito di dati personali e di lavoro, alle quali ricercatori esterni sono autorizzati ad accedere per elaborarli, sotto il controllo dell'agenzia statistica nazionale.

A questa condizione permissiva vanno peraltro affiancate, sollecitamente, condizioni positive di effettiva disponibilità di basi di microdati per la ricerca: prodotte in via corrente, mantenute e distribuite entro una cornice organizzativa adeguata, rese accessibili secondo il principio dell’*“universal access to all ‘bona fide’ analytical users”*.

Su questo terreno, i ritardi del nostro paese rispetto agli standard di paesi avanzati sono preoccupanti³⁹. La distanza da esperienze quali l’*UK Data Archive* o la *Data Liberation Initiative* canadese⁴⁰ è così vistosa da rischiare di essere scoraggiante. Non è certo questa la sede per trattarne. Ma la produzione e distribuzione di adeguate basi di microdati per la ricerca è – a mio giudizio – il prossimo, non eludibile impegno che attende le società scientifiche e l’intera comunità scientifica.

Dipartimento di Scienze Statistiche
Università di Padova

UGO TRIVELLATO

RIFERIMENTI BIBLIOGRAFICI

- E. BOYKO, W. WATKINS (2003), *Safe data, safe settings: Not either/or solution*, in CEIES (2003), op. cit..
- CEIES (2002), *CEIES opinion: Access to microdata*, adopted at the 12th meeting of the CEIES, Brussels, 22 November 2002 (draft CEIES 2002/12/6.2/EN).
- CEIES (2003), *Proceedings of the nineteenth Seminar: Innovative solutions in providing access to microdata, Lisbon-Portugal, 26-27 September 2002*, Eurostat, Theme 0 - Miscellaneous, Luxembourg, Office for Official Publications of the European Communities.
- A. CIANCARELLA (2003), *Tanto ‘sensibili’, ma così irragionevoli*, “Il Sole-24 Ore”, 15.05.2003.
- COMMISSIONE DELLE COMUNITÀ EUROPEE (2002), *Regolamento (CE) N. 831/2002 della Commissione del 17 maggio recante attuazione del Regolamento (CE) N. 322/97 del Consiglio relativo alle statistiche comunitarie, per quanto riguarda l’accesso ai dati riservati per fini scientifici*, “Gazzetta Ufficiale delle Comunità Europee”, 18.05.2002, N. L 133/7-9.
- CONSIGLIO DELL’UE (1997), *Regolamento (CE) N. 322/97 del Consiglio del 17 Febbraio 1997 relativo alle statistiche comunitarie*, “Gazzetta Ufficiale delle Comunità Europee”, 22.02.1997, N. L 52/1-7.
- CSS (1998), *Ricerca scientifica, accesso ai dati personali e tutela della riservatezza: considerazioni e raccomandazioni*, Roma, Consiglio italiano per le Scienze Sociali (mimeo).
- B. CONTINI (a cura di) (2002), *Osservatorio sulla mobilità del lavoro in Italia*, Bologna, Il Mulino.
- COUNCIL OF EUROPE (1981), *Convention for the protection of the individual with regard to the automatic processing of personal data*, adopted by the Committee of Ministers on 28 January 1981, Strasbourg [with Explanatory memorandum].
- COUNCIL OF EUROPE (1997a), *Recommendation No. R (97) 5 concerning the protection of medical da-*

³⁹ Per una breve rassegna di esperienze in altri paesi, rimando a Trivellato (2000). Vedi inoltre molti dei contributi raccolti in CEIES (2003). Per recenti orientamenti a livello dell’UE, espressi dal CEIES (Comitato Europeo consultivo sull’Informazione Statistica nei campi Economico e Sociale), che tra l’altro suggeriscono la realizzazione di *safe settings* virtuali sulla rete *Web*, vedi Dale e Trivellato (2002) e CEIES (2002).

⁴⁰ Per la prima vedi Schürer (2003); per la seconda Statistics Canada (1998) e Boyko e Watkins (2003).

- ta, adopted by the Committee of Ministers on 13 February 1997 at the 584th meeting of the Ministers' Deputies, Strasbourg [with Explanatory memorandum].
- COUNCIL OF EUROPE (1997b), *Recommendation No. R (97) 18 concerning the protection of personal data collected and processed for statistical purposes*, adopted by the Committee of Ministers on 30 September 1997 at the 602nd meeting of the Ministers' Deputies, Strasbourg [with Explanatory memorandum].
- A. DALE, U. TRIVELLATO (2002), *Access to microdata for scientific purposes*, Background Paper to the 19th CEIES Seminar 'Innovative Solutions in Providing Access To Micro-Data', Lisbon, 26-27 September 2002.
- EUROSTAT (1999), *Statistical data protection*. Proceedings of the conference, Lisbon, 25 to 27 March 1998, Luxembourg, Office for Official Publications of the European Communities.
- S.E. FIENBERG, L.C.R.J. WILLENBORG (Eds.) (1999), *Disclosure limitation methods for promoting the confidentiality of statistical data*, special issue of the Journal of Official Statistics, 14 (4).
- J.J. HECKMAN, J.A. SMITH (1995), *Assessing the case for social experiments*, "Journal of Economic Perspectives", 9 (2), pp. 85-110.
- S.P. JENKINS (1999), *Measurement of the income distribution: an academic user's view*, in CEIES, Proceedings of the seventh Seminar: Income distribution and different sources of income, Cologne-Germany, 10th-11th May 1999, Eurostat, Theme 0 - Miscellaneous, Luxembourg, Office for Official Publications of the European Communities.
- I. LLOYD (1998), *A guide to the Data Protection Act 1998*, London, Butterworths.
- D.B. MCCARTHY et al. (1999), *Medical records and privacy: empirical effects of legislation*, "Health Services Research", 34 (1), pp. 417-425.
- S. MCCLEAN (1998), *Statistical microdata, macrodata and metadata on the Web: perspectives*, in "Proceedings of the 21st SCORUS Conference", Belfast, 8-11 June 1998, Conference Compendium, University of Ulster, pp. 4.5.1-4.5.8.
- PARLAMENTO EUROPEO E CONSIGLIO DELL'UE (1995), *Direttiva 95/46/EC del Parlamento Europeo e del Consiglio del 24 ottobre 1995 relativa alla tutela delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati*, "Gazzetta Ufficiale delle Comunità Europee", 23.11.1995, N. L 281/31-50.
- E. RETTORE, U. TRIVELLATO (1999), *Come disegnare e valutare politiche attive del lavoro*, Il Mulino, 48 (385), 1999, pp. 891-904.
- E. RETTORE, U. TRIVELLATO, A. MARTINI (2003), *La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati?*, "Politica Economica", 19 (3), 2003, pp. 301-341.
- K. SCHÜRER (2003), *Solutions to provide access to microdata: the experience of the UK Data Archive*, in CEIES (2003), op. cit..
- STATISTICS CANADA (1998), *Report to the Data Liberation Board on management of the second year of operation of the Data Liberation Initiative*, Ottawa, Statistics Canada (mimeo).
- D.W. STEVENS (1998), *Confidentiality revisited: motives and consequences*, in "Proceedings of the International Symposium on linked employer-employee data", 21-22 May 1998, Arlington (VA), U.S. Bureau of the Census (CD-ROM).
- U. TRIVELLATO (2000), *Data access versus privacy: an analytical user's perspective*, "Statistica", 60 (4), pp. 669-689.
- G.G. WAGNER (2003), *The case of the German Socio-Economic Panel Study (SOEP)*, in CEIES (2003), cit..
- D. WARREN, L.D. BRANDEIS (1890), *The right to privacy*, "Harvard Law Review", 4 (5), pp. 193-220.
- D. WARREN, L.D. BRANDEIS (1999), *The end of privacy. The surveillance society*, "The Economist", May 1st-7th 1999, pp. 13-14 and 19-23.

RIASSUNTO

Protezione dei dati personali e ricerca scientifica

La tensione fra accesso ai dati e tutela della riservatezza è al crocevia della recente legislazione sulla protezione dei dati, che si estende a tutti i dati personali e all'insieme dei loro usi. La nota affronta la questione dal punto di vista di un analista, cioè a dire di una persona interessata a (raccolgere e soprattutto) trattare dati individuali per scopi di ricerca scientifica. L'interesse è portato innanzitutto su aspetti basilari della tematica e sulle implicazioni che essi hanno per politiche di accesso ai/diffusione dei dati per scopi di ricerca. Si evidenziano la funzione sociale della ricerca scientifica e le intrinseche necessità della stessa: queste considerazioni sono alla base dell'assegnazione al ricercatore di un 'credito di riservatezza' nel trattamento di dati personali. In questo quadro, si richiamano alcuni basilari principi e linee-guida in materia, quali risultano da recenti raccomandazioni e direttive internazionali. Viene quindi esaminato lo stato della situazione in Italia: l'approvazione, a fine 1996, di una legge sulla protezione dei dati personali piuttosto restrittiva; i passi in avanti realizzati con successivi provvedimenti; l'adozione del codice deontologico in materia, che completa il quadro delle regole con spirito ragionevolmente liberale e rende operante la disciplina più favorevole dettata per i trattamenti per scopi scientifici. Le considerazioni conclusive sottolineano la necessità di ulteriori passi per rendere le basi di dati individuali più ampiamente disponibili a ricercatori.

SUMMARY

Data protection and scientific research

The issue of privacy *versus* data access is at the crossroads of the recent legislation on data protection, that extends to all kinds of personal data and their uses. The paper takes the point of view of an analytical user, *i.e.* a person wanting to (collect and principally) process micro data for research purposes. It focuses on some basic features of the issue and on their policy implications. First of all, the societal role of scientific research and its intrinsic needs are considered: they are the basis for motivating a 'credit of confidentiality' when personal data are processed for scientific purposes. Some basic principles and guidelines for regulating this credit of confidentiality, as they are set out in recent international recommendations and regulations, are then briefly reviewed. The state of the affairs in Italy is discussed: a fairly strict act on data protection was passed in 1996; substantial improvements were introduced by further legislation; a code of conduct has been recently adopted, which is crucial both for a liberal interpretation of the rules and for making them effective (indeed, the more favourable provisions dealing with the processing personal data for research purposes were conditional on the code to come into force). The concluding remarks stress the need for additional efforts to make micro databases widely available to researchers.