Lettori: n.d.
CORRIERE DELLA SERA laLettura
07-DIC-2014
Diffusione: n.d.
da pag. 4

Aiuto, la nostra salute non è un segreto

Una falla del sistema permette di identificare i pazienti Così datori di lavoro e industrie potrebbero servirsene

Il caso I dati su sesso, giorno di nascita e comune di residenza sono accessibili dalle schede di dimissione ospedaliere. Combinati con le liste elettorali, reperibili a poco prezzo, consentono di risalire con precisione quasi assoluta a nomi e cognomi dei ricoverati. Una pacchia per chi voglia conoscere la storia sanitaria di una persona prima di assumerla, erogarle un mutuo, venderle prodotti mirati

di SERENA DANNA *e* SIMONA RAVIZZA

I dati personali sono l'oro del terzo millennio. Il nuovo petrolio, come li definisce il World Economic Forum report. Li vogliono e li cercano — attraverso l'uso delle nuove tecnologie — imprese, società di marketing, banche e assicurazioni desiderose di conoscere gusti, abitudini e stato di salute di ciascuno di noi per venderci prodotti e prendere decisioni sulla nostra vita. Nel mercato dei Big Data, le più ambite sono senza dubbio le informazioni sanitarie. Ma i dati sulla nostra salute sono anche i più sensibili e protetti: quale datore di lavoro ci assumerebbe sapendo che siamo malati? Quale istituto di credito ci concederebbe un mutuo? E chi ci farebbe sottoscrivere una polizza? Finora abbiamo creduto che almeno i segreti sulla nostra salute fossero al riparo da occhi indiscreti.

Ci sbagliavamo.

In Italia il diritto alla riservatezza delle informazioni sanitarie può essere violato. I dati contenuti nelle cartelle cliniche, i risultati delle analisi mediche, i giorni di ricovero in ospedale — tutte informazioni tanto delicate quanto appetibili — possono finire nelle mani sbagliate. Lo stato di salute dei cittadini è una miniera d'oro per i broker data, gli analisti che vendono alle aziende di marketing i profili dei consumatori. Il «Financial Times» ha quotato il mercato delle informazioni e rivelato che la presenza delle condizioni mediche è una costante: un uomo ricco, obeso, in attesa di un figlio è un'informazione che per una società di marketing può valere 1,05 dollari; l'informazione su un professionista sposato, affetto da diabete e senza figli, è quotata 0,6. Possiamo essere raggruppati in base alle

nostre condizioni fisiche e «venduti» di conseguenza: i diabetici alle imprese alimentari di cibi a basso contenuto calorico; gli incontinenti ai produttori di pannoloni; le donne incinte alle aziende di latte artificiale. Ma non è solo una questione di marketing. Il rischio è di trovarsi discriminati. Senza lavoro, polizze assicurative, prestiti bancari.

Uno scenario da fantascienza? Niente affatto.

Tutto può partire dalle schede di dimissione ospedaliere (Sdo). Ne viene compilata una ogni volta che un malato, ricoverato in strutture pubbliche o private, lascia l'ospedale. Il loro uso è strettamente riservato al paziente e l'ospedale è tenuto a produrle per i rimborsi pubblici. Le schede di dimissioni ospedaliere, però, restano negli archivi e possono essere rilasciate dietro richiesta anche ai ricercatori per fini di studio. Questi documenti dovrebbero essere anonimi per garantire la nostra privacy. In realtà — ed è questo il problema — i malati possono essere identificati. Tra i dati che gli studiosi possono consultare ci sono, infatti, i cosiddetti «quasi-identificatori»: sesso, data di nascita e comune di residenza. Sono tre elementi che, combinati con le banche dati dell'anagrafe, consentono di risalire al paziente. A ciascuno di noi.

I tre «quasi-identificatori» dovrebbero essere chiusi in cassaforte. Così non è. Ricercatori di importanti istituti italiani confermano a «la Lettura» di essere in possesso dei tre dati. È quanto ammette per email anche la Direzione generale per la programmazione sanitaria del ministero alla Salute: «Se il progetto di studio per cui viene effettuata la richiesta giustifica la necessità di disporre di questa variabile e assicura il trattamento corretto del dato secondo la normativa sulla protezione dei dati personali e sensibili, data di nascita, comune di residenza e sesso dei pazienti possono essere rilasciati insieme».

La pericolosità dei «quasi-identificatori» per la privacy è nota dal Duemila, anno in cui Latanya Sweeney — all'epoca laureanda del Mit di Boston, oggi a capo del laboratorio di privacy della prestigiosa università e consulente del governo — pubblicò uno studio che ha ispirato la nuova legislazione americana sulla privacy dei dati sanitari. Negli anni Novanta lo Stato del Massachusetts, in una delle prime iniziative di rilascio di dati della storia, decide di rendere pubbliche alcune informazioni mediche, cancellando gli «identificatori espliciti» (nome, indirizzo e numero di previdenza sociale), ma lasciando lo zip code a 5 cifre (codice di avviamento postale), la data di nascita e il sesso del paziente. I dati vengono acquisiti dalla studentessa, che, per avere accesso a nomi e indirizzi di 54 mila residenti di Cambridge, con soli 25 euro compra le liste elettorali al municipio. Incrociando i due database, Sweeney arriva a identificare la maggior parte dei ricoverati della città, dove hanno sede le università Mit e Harvard, compreso il governatore dello Stato William F. Weld, curato qualche tempo prima per un malore. Successivamente la studentessa applica lo stesso metodo fuori dai confini della città. Con risultati sorprendenti: utilizzando data di nascita, zip code a cinque cifre e sesso, Sweeney identifica in maniera univoca l'87% della popolazione americana.

¶

A distanza di 14 anni dallo studio americano, in Italia, le schede che vengono compilate dagli ospedali e che — passando dalle Regioni — arrivano al ministero della Salute contengono ancora le tre informazioni, che possono essere rilasciate per fini di ricerca. Evidentemente, nonostante il clamore suscitato dal caso Sweeney e gli spettri di discriminazione e business illeciti che si nascondono dietro l'accesso ai dati sanitari dei cittadini, per il Garante italiano della privacy «i quasi-identificatori» non costituiscono un problema. Eppure i comportamenti di ciascuno di noi possono essere liberi solo finché la privacy è assicurata. È il motivo per cui a chiunque entri in contatto con medici per cure, interventi chirurgici e acquisto di medicine dev'essere garantita la più assoluta riservatezza.

Com'è possibile «bucare» il sistema dei dati sanitari? Il processo è ricostruito da Valeria Cattolica, ex studentessa del dipartimento di informatica dell'Università degli Studi di Milano, nella tesi di laurea *Privacy e ricoveri ospedalieri*. Cattolica ha chiesto e ottenuto i dati di ricovero dell'ospedale di Legnano; poi ha comprato — con poco più di 100 euro — le liste elettorali del comune lombardo per avere accesso a una base dati anagrafica. Incrociando i dati ospedalieri di Legnano con quelli ricavati dalle liste elettorali, lo studio di Cattolica ha portato all'identificazione univoca con nome e cognome di 3.005 ricoverati, ovvero del 90,38% del totale degli assistiti dell'ospedale nel 2012 (il restante 9% riguarderebbe pazienti trasferiti o deceduti oppure casi di omonimia).

¶

Quanti segreti sulla nostra salute sono nelle mani di ricercatori e tecnici che potrebbero utilizzarli a nostra insaputa? E come mai i problemi, sebbene noti da tempo, non sono stati affrontati dalle autorità? È convinzione comune tra gli esperti di privacy che le tecniche italiane di schermatura dei dati nella pubblica amministrazione siano arretrate. Più che di «anonimizzazione», si tratta di «pseudonimizzazione»: per garantire l'anonimato dei pazienti non basta eliminare gli identificatori espliciti (nome, cognome, codice fiscale), ma bisogna camuffare in maniera definitiva e irreversibile il dato, in modo che sia impraticabile qualsiasi confronto con informazioni in chiaro o con dataset pubblici.

Prendiamo il codice fiscale: la sua conversione in un codice «non parlante» attraverso un'operazione di cifratura comporta il rischio che, a un certo punto, si crei una tabella di transcodifica, la cui accessibilità comprometterebbe la sicurezza dell'informazione. Basterebbe avere accesso a quella tabella per decodificarlo. Per spiegare le ragioni del problema sarebbe tuttavia sbagliato prendersela con la lenta burocrazia italiana. La questione va letta nel contesto più generale della mutazione del concetto stesso di privacy, che deve adattarsi alle scoperte tecnologiche e ai nuovi contesti.

Paul Ohm, docente di legge all'Università del Colorado e tra i maggiori esperti di anonimizzazione dei dati personali, ha scritto in un articolo della «Ucla Law Review» del 2010: «La nuova scienza della re-identificazione ha rivoluzionato il panorama delle politiche di privacy, minacciando la fede che abbiamo messo nell'anonimizzazione. Non è cosa da poco se consideriamo tutti i tecnologi che giustificano la condivisione e l'archivio dei dati, assicurando agli utenti che stanno proteggendo la loro privacy». Secondo Mauro Alovisio, direttore del Centro studi di informatica giuridica di Ivrea-Torino, «l'impatto dei dati sanitari è sottovalutato in Italia». L'avvocato crede che siano due i problemi principali: da un lato la poca efficacia del processo di anonimizzazione del ministero della Salute, dall'altro la mancanza di verifiche «per garantire che i dati siano effettivamente anonimi». Alovisio ricorda che dal 2002, quando fu introdotto dal Garante della privacy il «Codice di deontologia e di buona condotta per i trattamenti di dati personali a scopi statistici e di ricerca scientifica effettuati nell'ambito del Sistema statistico nazionale», non c'è stata alcuna presa di consapevolezza effettiva da parte degli addetti ai lavori. «Eppure i dati — soprattutto quelli sanitari — creano valore e dovrebbero essere trattati con rigore e cura».

Raccolta dati sulle malattie
Ogni volta che un malato, ricoverato in strutture pubbliche o private, lascia l'ospedale viene compilata una scheda di dimissioni ospedaliere (Sdo)

Ricerca e privacy
Per motivi di studio e ricerca scientifici, il ministero della Salute è disposto a rilasciare le schede di dimissioni ospedaliere con data di nascita, comune di residenza e sesso dei pazienti. Sono i cosiddetti «quasi-identificatori» che, incrociati ai dati anagrafici, permettono di arrivare all'individuazione dei pazienti

Liste elettorali
Chi ha in mano data di nascita, sesso e comune di residenza, può attingere, come base anagrafica, alle liste elettorali. Sono, infatti, acquistabili pagando poco più di 100 euro

Bibliografia
Il primo studio sulla pericolosità dei «quasi-identificatori» risale al Duemila e ha ispirato la nuova legislazione americana sulla privacy dei dati sanitari. L'autrice è Latanya Sweeney, all'epoca laureanda del Mit di Boston. Oggi la Sweeney è a capo del laboratorio di privacy della prestigiosa università e consulente del governo

Il mercato dei Big Data
In un articolo del 12 giugno 2013 il «Financial Times» ha quotato il mercato delle informazioni personali. Tra le più ambite, quelle sanitarie. Un uomo ricco, obeso, in attesa di un figlio è un'informazione che per una società di marketing può valere 1,05 dollari; mentre l'informazione su un professionista sposato, affetto da diabete e senza figli, è quotata 0,6