Presidenza del Consiglio dei Ministri

TECNOLOGIE DELL'INFORMAZIONE E DELLA COMUNICAZIONE E *BIG DATA*: PROFILI BIOETICI

25 novembre 2016

INDICE

Introduzione

L'avanzamento delle tecnologie dell'informazione e della comunicazione e gli sviluppi dell'informatica stanno aprendo nuove possibilità di conoscenza, uso e applicazione in diversi ambiti, tra questi anche la salute. Con l'espressione *big data* si indicano massivi dati digitali costituiti dalle tracce delle informazioni personali rilasciate nell'ambiente mediante l'uso di tecnologie informatiche: si tratta di dati di enorme quantità, provenienti da fonti eterogenee, generati con grande velocità e pervasività.

Il documento si sofferma sulle enormi opportunità di sviluppo che si dischiudono, in particolare nell'ambito sanitario, con la telemedicina, la medicina di precisione, l'elaborazione di politiche sanitarie. Il Comitato delinea anche alcune criticità nella difficoltà a governare l'enorme massa di dati nella raccolta, analisi e uso dei dati, in modo particolare quando sono usati e applicati in modo diverso dalla raccolta iniziale o senza la consapevolezza dell'utente. Le tecnologie informatiche possono presentare dei rischi che riguardano soprattutto il trattamento e la protezione della vita privata e dei dati a carattere personale, la trasparenza, la qualità dell'informazione, la dipendenza, il principio di giustizia partecipativa e la governance.

Il CNB, in vista di un uso "appropriato" dei dati, rispettoso dei diritti fondamentali dell'uomo, propone alcune raccomandazioni: definizione delle responsabilità dei provider; verifica della qualità dei dati e della trasparenza degli algoritmi; identificazione di strumenti efficienti per richiedere il consenso o dissenso al trattamento dei dati all'utente; attuazione a un riconoscimento effettivo del diritto all'oblio; coinvolgimento dei cittadini ai problemi etici emergenti, con particolare attenzione ai giovani, auspicando la elaborazione di linee guida per un corretto uso delle tecnologie sociali; promozione della ricerca per l'innovazione dell'approccio etico al disegno delle tecnologie sociali; garanzia delle condizioni di accesso a tutti coloro che intendono avvalersi delle nuove tecnologie.

Il parere sollecita l'elaborazione di una normativa per la protezione dei dati personali e la tutela dei cittadini/utenti da rischi sociali dell'abuso dei dati.

In appendice è inserita una nota giuridica che ripercorre sinteticamente in modo critico la situazione attuale della regolazione europea e nazionale.

Il parere è stato proposto all'attenzione del Comitato, come tema e bozza originaria, dalla Prof.ssa Rosaria Conte nell'aprile 2015.

La Prof.ssa Rosaria Conte ha illustrato in un'audizione interna lo stato della discussione scientifica ed etica.

Il gruppo si è riunito due volte nel 2015. La redazione del parere si è interrotta a causa delle condizioni di salute di Rosaria e della sua scomparsa nell'estate del 2016, che ha lasciato un grande vuoto.

I Proff. Antonio Da Re, Lorenzo d'Avack, Laura Palazzani e Carlo Petrini hanno integrato e completato il parere.

Il parere è stato approvato all'unanimità dei presenti il 25 novembre 2016 dai Proff.: Salvatore Amato, Luisella Battaglia, Stefano Canestrari, Antonio Da Re, Lorenzo d'Avack, Riccardo Di Segni, Silvio Garattini,

Marianna Gensabella, Assunta Morresi, Laura Palazzani, Monica Toraldo di Francia, Grazia Zuffa.

Hanno espresso la loro approvazione anche i membri di diritto: Dott. Maurizio Benato, Dott. Carlo Petrini.

Hanno successivamente espresso la loro adesione i Proff.: Carlo Caltagirone, Cinzia Caporale, Carlo Casonato, Bruno Dallapiccola, Mario De Curtis, Carlo Flamigni, Paola Frati, Andrea Nicolussi, Rodolfo Proietti, Lucetta Scaraffia e i membri di diritto, Prof.ssa Maria Teresa Palamara e Dott.ssa Carla Bernasconi.

1. Premessa

In questo documento, il Comitato Nazionale per la Bioetica (CNB) intende riferirsi alle Tecnologie dell'Informazione e della Comunicazione (ICT: *Information and Communication Technologies*) o l'insieme dei metodi e delle tecnologie per la trasmissione ed elaborazione digitale di informazioni che consentono o migliorano l'interazione fra e con utenti a diversi livelli (individui, gruppi, comunità, compagnie, organizzazioni, istituzioni, ecc.), con specifico riferimento al fenomeno dei *big data* in relazione alla salute degli individui.

L'avanzamento delle tecnologie dell'informazione (con l'aumento e velocizzazione della raccolta, conservazione ed elaborazione delle informazioni) e gli sviluppi della "scienza dei dati" (ossia l'uso della informatica e della matematica con tecniche statistiche e algoritmi) stanno aprendo nuove possibilità di conoscenza, uso e applicazione in diversi ambiti, tra questi anche la salute.

In modo particolare ci si riferirà: a) alle Tecnologie Sociali, ossia le tecnologie utilizzate da utenti per condividere contenuti e tecnologie che offrono informazione su possibili partner di scambi e cooperazioni (social networks); b) ai providers; c) ai motori di ricerca; d) alle tecniche di profilazione dell'utenza, che ottengono/ricavano/estraggono (grazie all'uso di complessi algoritmi) informazioni dagli utenti, per applicazioni diversificate anche commerciali, nel contesto della c.d. "rivoluzione digitale"[1].

Con l'espressione *big data*, in questo contesto, si indicano massivi dati digitali costituiti dalle tracce rilasciate nell'ambiente mediante l'uso di tecnologie informatiche[2], caratterizzati da alcune proprietà:

- eterogeneità: diversità di tipologia di dati digitali e diversità delle fonti di provenienza (computer, telefoni cellulari, Internet, sensori e dispositivi mobili, ecc.); ai dati medici tradizionali (risultati di analisi di laboratorio, cartelle cliniche, indagini epidemiologiche) si aggiungono dati medici provenienti dalle nuove tecnologie (test genomici, risultati di body imaging o internet of the body, registri elettronici, tecnologie di monitoraggio remoto o telemedicina);
- numerosità: aumento quantitativo dei dati, di cui è sempre più difficile seguire l'aggiornamento delle unità di misura (terabyte, pentabyte, esabyte, ecc.);
- velocità: incremento esponenziale delle connessioni e della velocità di generazione dei dati (è molto probabile che nel 2016 il numero di connessioni sui social arrivi a quasi 19 miliardi);
- ubiquità: potenziale espansione illimitata dei dati nell'ambito della rete globale, e pervasività o invasività della possibile raccolta di dati in ogni ambito della vita delle persone;
- possibile non-veridicità: scarsità e possibile mancanza di un'opera di verifica dell'autenticità, della precisione e della qualità dell'informazione.

[1] S. Spiekermann, A. Acquisti, R. Boehme, K-L. Hui, *The Challenges of Personal Data Markets and Privacy*, in "Electronic Markets", 2015, 25, pp. 161-167.

[2] Non sarà oggetto di analisi la questione dei *big data* provenienti da informazioni tratte dall'uso di test genetici, *genome-wide test*, oggetto di analisi del parere *Gestione degli* incidental findings *nelle indagini genomiche con le nuove piattaforme tecnologiche*, 2016.

L'utenza è indotta e incentivata a fornire informazioni perché scambia la propria con altra informazione. E poiché qualunque azione degli utenti nel sistema si trasforma in *signaling*, cioè dà accesso alle caratteristiche dell'utente, gli utenti rilasciano una gran quantità di informazioni a beneficio dei gestori dei sistemi in oggetto. I sistemi ICT sono quindi generatori di informazione pubblica e privata prodotta collettivamente.

Non ci occuperemo invece:

= di piattaforme educative (come le MOOC, *Massive Open Online Courses* o quelle in dotazione alle specifiche istituzioni telematiche) né di sistemi di intrattenimento (*secondlife*, *World of warcraft*, ecc.), perché non si basano sull'erogazione collettiva di informazione[3];

= di *big data* in Paesi in via di sviluppo con economie emergenti, dato che molti problemi hanno connotazioni diverse;

= dei sistemi di sicurezza e dei sistemi di scambio e condivisione (come i mercati elettronici e i sistemi di *economysharing*), che pur essendo basati sull'informazione fornita dall'utenza e pur sollevando preoccupazioni etiche in senso lato non presentano problemi di natura specificamente bioetica, sui quali si concentra l'attenzione del Comitato;

= i sistemi che fanno uso di *big data* per rilevare orientamenti e sensibilità politiche collettive e che vengono utilizzati per ottenere consensi e orientare scelte elettorali.

Il CNB si è già in parte occupato di alcune di queste tematiche in precedenti pareri: *Etica, salute e nuove tecnologie dell'informazione,* 2006, *L'identificazione del corpo umano: profili bioetici della biometria,* 2010 e *Mobile-health e applicazioni per la salute: aspetti bioetici,* 2015.

2. Prospettive di sviluppo e problematiche etiche emergenti

Le tecnologie informatiche offrono notevoli opportunità di sviluppo economico e sociale, e potenziano significativamente le possibilità dell'individuo di acquisire informazioni e di entrare in relazione con altri individui e soggetti sociali, istituzionali, ecc..

L'espansione delle ICT consente e consentirà la comunicazione a livello mondiale nella società c.d. "numerica" e ai cittadini digitali di ottenere ampie informazioni, in maniera semplice, anche nella propria lingua e in quantità indefinite. I grandi sistemi ICT nel futuro ci condurranno in un'epoca di scelte e di possibilità senza precedenti con un maggiore numero di dati a disposizione. La "esplosione" dei dati e la "rivoluzione dei dati" avrà ricadute nel mondo della politica, dell'economia, dell'istruzione.

Al tempo stesso, la facilità di accesso e di uso delle ICT consente pressoché a chiunque di avvalersene. Gli unici limiti esistenti sono quelli basati sull'accessibilità alla tecnologia necessaria (cellulare, energia

[3] Inoltre, in questi sistemi, la numerosità dell'utenza non è altrettanto esplosiva quanto quella di altre tipologie di ICT (nel 2015, 1 milione su *secondlife*, 5 milioni su *world of warcraft* contro 1 miliardo e mezzo su FB). Infine l'effettiva funzionalità di questi sistemi è controversa: *secondlife* viene ad esempio considerata da qualcuno un fallimento (https://it.finance.yahoo.com/notizie/che-fine-fatto-secondlife-07524429.html), anche se l'utenza di questo sistema non cresce ma neppure arretra (http://www.downloadblog.it/post/21761/secondlife-dimenticato-dalla-stampa-ma-rimasto-solido-ed-amato).

elettrica) e un certo livello di istruzione che ne consenta l'utilizzo (la c.d. "*digital literacy*"). L'uso sempre più pervasivo delle nuove tecnologie dell'informazione e della comunicazione nella vita quotidiana (ogni nostra azione lascia tracce informatiche), ci rende, consapevolmente o inconsapevolmente, "utenti/contributori" dei *big data* e la nostra dipendenza "digitale", personale e collettiva, tende ad aumentare. Ma d'altra parte il problema non è quello di fermare tutto o tornare indietro: non solo è impossibile, ma significherebbe voler bloccare un processo di sviluppo con enormi potenzialità e promesse.

In modo particolare rilevanti percorsi di sviluppo nell'ambito della salute si delineano con la c.d. "*Data-driven precision medicine*", ossia la possibilità (ancora oggetto di studio e ricerca e al momento non estensibile a tutti gli aspetti della biologia e della medicina) di costruire sulla base della quantità di dati raccolti predizioni e simulazioni di diagnosi e trattamenti per singoli pazienti in contesti specifici o per gruppi stratificati di pazienti (la c.d. medicina personalizzata/stratificata o medicina di precisione), ma anche possibilmente estesi alla definizione di politiche sanitarie per la salute pubblica in particolare di tipo preventivo. Inoltre le ICT possono migliorare e rendere più efficienti i servizi sanitari. Si sta infatti creando una sempre più marcata compenetrazione tra assistenza sanitaria, tecnologie informatiche e decisioni nell'ambito della cura. Si tratta di un fenomeno che in termini onnicomprensivi è stato denominato *Cybermedicine* e che sta ad indicare la messa in rete di una vasta comunità di interessi tra diversi soggetti sempre più interdipendenti e con la necessità di scambio comunicativo (istituzioni nazionali e locali, aziende sanitarie pubbliche con le loro organizzazioni territoriali, le organizzazioni sanitarie private, i professionisti della salute, le farmacie e tutto il settore farmaceutico, le società di servizi e i fornitori di aziende no-profit, le associazioni del volontariato e dei cittadini).

Viviamo in un mondo profondamente interconnesso, del quale però conosciamo solo a grandi linee le modalità di interconnessione: non solo non siamo in condizione di conoscere sufficientemente le procedure e i percorsi con cui si formano i dati che riguardano la nostra persona, ma non siamo pienamente informati dell'uso effettivo che ne verrà fatto, dove saranno conservati e per quanto tempo, quali soggetti potranno utilizzarli e a quali fini, con quali percorsi, chi è responsabile e in che misura. È l'esistenza dei *big data* in quanto tali a rendere impossibile essere veramente e pienamente a conoscenza di questo nuovo universo.

Emergono comunque alcune problematicità sul piano etico in relazione ai *big data* con specifico riferimento alla raccolta, analisi e uso dei dati, in modo particolare quando sono usati e applicati in modo diverso dalla raccolta iniziale o senza la consapevolezza dell'utente[4]. Le tecnologie informatiche possono presentare dei rischi che riguardano soprattutto il trattamento e la protezione della vita privata e dei dati a carattere personale, la trasparenza, la qualità dell'informazione, la dipendenza, anche patologica, alle ICT, il principio di giustizia partecipativa e la governance. I pericoli riguardano soprattutto: l'abuso dei

[4] Sulle problematiche etiche emergenti nell'ambito delle ICT, si veda: European Group on Ethics in Science and New Technologies (EGE), *Ethics of Information and Communication Technologies*, 2012; riferimenti ai big data in *The Ethical implications of new health technologies and citizen participation*, 2015.

dati che provoca danno agli individui, sorveglianza indebita, stigmatizzazione o discriminazione. Gli scenari possono, dunque, essere complessi e sotto alcuni profili inquietanti, ancora non sufficientemente analizzati, soprattutto nell'ambito dell'impatto sulla salute dei cittadini/utenti.

2.1. Trattamento dei dati personali: sfide alla privacy e al consenso informato

I contatti con il mondo virtuale iniziano fin dalla nascita e intere fasi della vita vengono fissate nel tempo e potenzialmente rese digitalmente visibili a tutti. L'esposizione dei "dati personali" e della "vita personale" nella rete fa emergere la questione della privacy, problematica di cui, in particolare i soggetti minori, sono spesso inconsapevoli. Tale problematica diviene ancora più complicata da affrontare in una società che apprezza, valorizza e promuove la "condivisione" (*sharing*) dei dati (informazioni, immagini, video).

Si dovrebbe fin dai primi anni di vita, soprattutto con l'aiuto dei genitori e dei sistemi scolastici, fare emergere l'esigenza della tutela della riservatezza di talune informazioni e le modalità di difesa della propria privacy virtuale. Una educazione alla cittadinanza attiva, che voglia essere adeguata ai tempi (una sorta di "Educazione civica") non può affatto sottovalutare la rilevanza di questi temi. Ne consegue la raccomandazione che siano istituiti programmi educativi che permettano agli utenti di sviluppare tale consapevolezza nel contesto di una cultura tecnica sulle modalità di uso di internet, tanto più che l'estensione dei dati rischierà di togliere ai cittadini gran parte del controllo che hanno sulle loro informazioni private nello spazio virtuale. La sfida che si presenta all'utente digitale è quella di acquisire consapevolezza critica del problema, di determinare quali accorgimenti e strumenti possa e voglia utilizzare per non perdere il controllo sulla propria privacy, nella misura possibile.

Una delle prime questioni etiche emerse nella discussione etica e giuridica sulle ICT è stata l'esigenza che le persone possano avere un controllo sufficiente dei loro dati nell'utilizzare internet, in particolare dei dati c.d. sensibili, tra questi anche quelli sanitari. Le prime regolamentazioni nel settore hanno precisato le condizioni applicabili al trattamento dei dati nel momento della richiesta delle informazioni, che deve essere sempre accompagnata da un consenso informato esplicito. Ciò implica che vi sia trasparenza, completezza e semplicità nell'informazione data da coloro che utilizzano e trattano le informazioni: gli operatori dovrebbero specificare chi raccoglie e chi userà i dati, quali dati, come vengono raccolti, dove verranno conservati e per quanto tempo, per quale ragione e per quale scopo. Un consenso che deve poter essere sempre revocabile, senza conseguenze negative per l'utente; e, salvo che l'espressione del consenso non preveda diversamente, deve permanere il diritto all'accesso, alla rettificazione e alla cancellazione dei

dati a carattere personale (il c.d. diritto all'oblio)[5], di modo che essi non siano più accessibili al pubblico sotto qualsiasi forma (copie o riproduzione). Il consenso deve sempre poter prevedere informaticamente una alternativa, ossia il dissenso (a fornire informazioni), precisandone le implicazioni. Si tratta, in altri termini, di trasporre il concetto di consenso informato "tradizionale" dalla biomedicina all'informatica.

Nell'era dei *big data* stiamo assistendo ad una trasformazione digitale così radicale che esige un cambiamento anche dell'approccio etico e giuridico, con la risignificazione delle categorie tradizionali, tra queste proprio quella della "privacy". Molti parlano di "fine della privacy", o comunque di un concetto destinato alla "evaporazione". Anche il consenso informato nella pratica digitale è ben più complesso ed esige un ripensamento strutturale. Si tratta di delineare i nuovi scenari nei quali ci troviamo nell'era dei *big data*.

1) Quali informazioni sono raccolte?

Le informazioni richieste all'utente digitale sono di natura assolutamente eterogenea: ad esempio, le informazioni relative ai dati socio-anagrafici (sesso, età, stato civile, educazione, filiazione, ecc.) divergono in modo significativo da quelle relative alla religione, all'appartenenza politico-ideologica, e queste divergono a loro volta dall'informazione concernente gli stati emotivi, gli atteggiamenti, le attitudini, le preferenze ecc. A quale tipo di informazione, dunque, ci si riferisce quando si autorizza il trattamento di dati personali nell'ambiente digitale? Lo stesso concetto di privacy sta cambiando: non fa più solo riferimento alla protezione dei "dati personali", informazioni relative ad una persona fisica, identificata o identificabile, anche indirettamente, mediante riferimento a qualsiasi altra informazione, ivi compreso un numero di identificazione personale e informazioni anagrafiche, ma si allarga alla considerazione della "vita personale" o "vita privata" (che include anche informazioni su dati biologici, su comportamenti, pensieri sentimenti, comunicazioni personali, movimenti, aggregazioni ecc.). Con privacy oggi si intende ogni interferenza con la vita privata, come appello alla riservatezza e confidenzialità.

Per quanto riguarda specificamente i dati relativi alla salute, va anche considerato che i confini tra l'ambito strettamente medico e non medico si vanno sempre più sfumando, così come tra salute e società: le informazioni sugli stili di vita e sui comportamenti tendono a divenire sempre più rilevanti per la salute anche in prospettiva di prevenzione. In questo senso sono da considerarsi informazioni sulla salute non solo i

[5] Come si sostiene nel parere del CNB *L'identificazione del corpo umano: profili bioetici della biometria,* 2010, "deve diventare un aspetto del diritto fondamentale all'identità personale, il diritto a non essere schedati, classificati, eventualmente emarginati in maniera irreversibile sulla base di informazioni assunte a propria insaputa, attraverso criteri non trasparenti e per finalità in gran parte ignote. La crescita, in termini di efficienza e sicurezza, delle acquisizioni biometriche, dovrebbe quindi accompagnarsi all'aumento proporzionale delle possibilità di tutela e auspicabilmente della consapevolezza pubblica. Se non è possibile pretendere l'anonimato, è fondamentale che almeno siano garantite le condizioni per ottenere l'oblio". Il documento dell'European Group on Ethics in Science and New Technologies (EGE), *Ethics of Information and Communication Technologies*, cit., suggerisce di sostituire l'espressione prevalente "*right to be forgotten"* con "*right to data deletion"*.

risultati di analisi di laboratorio o dati epidemiologici, ma anche notizie generali che provengono dai social network.

2) Come vengono raccolte le informazioni?

Qualunque sia il tipo di informazione, è difficile garantire la piena consapevolezza del diretto interessato nel momento in cui acconsente al trattamento dei propri dati: nella quotidiana pratica dell'uso dei servizi in rete l'utente, distratto e frettoloso, non si sofferma a leggere con attenzione i contratti che sottoscrive (spesso visualizzati su uno schermo, scritti in caratteri piccoli, in molteplici pagine) per poter accedere al servizio che gli interessa. Non solo: l'utente - e questo è vero particolarmente nel caso dei gestori d'informazione in rete - non sa come questa informazione sia stata ottenuta, non sa quali specifici utilizzi essa consenta, non sa o non ricorda se e a che cosa in precedenza abbia consentito. Si consideri ad esempio il caso delle informazioni che si possono ricavare dai like che un iscritto può lasciare sulle pagine di un social network. Anche se il soggetto in questione sa che i suoi like possono rappresentare una preziosa fonte d'informazione per il provider del servizio di cui sta beneficiando, non sa come questi vengano trattati. Nel contempo l'utente non può dirsi completamente libero quando l'alternativa cui si trova di fronte è fra cedere i propri dati o rinunciare a fruire di servizi del social network. O addirittura quando in cambio della "dazione" di dati ottiene dei servizi gratuiti (es. accesso a musica o altro).

3) Chi raccoglie e conserva i dati? Per quale scopo? Dove vengono conservati?

Rispetto alla quantità della raccolta e delle modalità di elaborazione dei dati il singolo utente difficilmente è in grado di comprendere le potenzialità detenute da chi raccoglie i dati che lo riguardano. All'utente non è generalmente spiegato chi raccoglie i dati (la compagnia telefonica se raccolti dal cellulare? Il provider se raccolti dal computer?). I "gestori di informazione in rete" (sia pubblici che privati) raccolgono i dati in forma aggregata, con la possibilità di estrarre informazione (è questa la base dell'appetibilità economica dei dati). Si tenga altresì conto che, come accade ora, le piattaforme online (in particolare i gestori privati) possono inoltrare o vendere i dati come "macro-aggregati" ad aziende terze (se non è impedito sul piano legale) senza chiedere uno specifico e puntuale consenso, con il risultato che le persone finiranno per comunicare dati della propria vita più di quanto siano consapevoli.

Vi è "di principio" una differenza tra i potenziali usi dei dati e i rischi connessi: è diverso l'uso a scopo commerciale (ad es. l'uso di dati per identificare categorie a scopo pubblicitario e di marketing) o l'uso scientifico (ad es. la condivisione delle informazioni sanitarie all'interno dei servizi/percorsi di cura, a scopo di ricerca clinico-epidemiologica). Non è facile fare questa distinzione; nella maggior parte dei casi il confine tra ricerca scientifica e ricerca di mercato diviene sfumato ed indistinguibile, a discapito dell'utente digitale inconsapevole (es. si pensi alla piattaforma "*Patientslikeme*", ove la ricerca scientifica e commerciale sono strettamente connesse in modo indistinto).

D'altra parte, anche ammettendo la possibilità che tutte le informazioni siano chiaramente espresse e comunicate agli utenti (sulla raccolta,

conservazione e uso dei dati), si pone il problema della loro effettiva comprensibilità, data la quantità di informazioni e la complessità dei contenuti (basti pensare a cosa significa spiegare e capire significato e funzionalità di un "algoritmo" o spiegare i concetti di vita privata e di dignità della persona) e data la mancanza di interazione con l'utente che consenta la verifica di adeguatezza (a differenza del consenso informato che si realizza dopo l'informazione fornita dal medico al paziente).

Ne risulta, inevitabilmente, una "sfida" al consenso informato tradizionalmente inteso, che di contro esige una riformulazione seguendo nuovi percorsi, che potrebbero prevedere modalità di consenso informato differenziate a seconda dei contesti.

Il consenso informato, nell'ambito dei *big data*, non può essere specifico (così come avviene, per esempio, per il consenso medico "tradizionale"); esso esige caratteri di "ampiezza" inevitabili, data l'impossibilità di anticipare in modo preciso i percorsi della ricerca (analogamente a quanto accade nell'ambito delle biobanche): è quindi un consenso "dinamico" e "flessibile", che identifica aree simili di ricerca direttamente o indirettamente collegate al percorso originario. Ampiezza, dinamicità e flessibilità non significano un consenso "cieco" a qualsiasi ricerca.

Si potrebbe forse ripensare la stessa espressione "consenso informato" nel mondo digitale, limitandola ad una "presa di coscienza" o "presa d'atto" che i dati saranno raccolti, nella consapevolezza critica dell'impossibilità dell'anonimato (essendo sempre possibile la re-identificazione, ossia risalire mediante incrocio di informazioni all'identità dell'utente), nella non precisa determinazione a priori delle modalità d' uso, conservazione ed analisi dei dati, nell'impossibilità di garantire sicurezza e confidenzialità in ogni circostanza.

Va anche considerato che l'esigenza di introdurre in modo puntuale l'uso del consenso per ogni specifico passaggio nell'utilizzo e applicazione dei dati può provocare nell'utente digitale la reazione opposta: non di aumentare la sua attenzione sui problemi, ma di renderlo più frettoloso e distratto rispetto all'oggetto del consenso. Si pensi all'uso del consenso ai cookies: non si può non rimarcare che a fronte di una giusta preoccupazione per l'utilizzo delle informazioni ottenute "cliccando" sulle diverse pagine web, il meccanismo di richiesta è divenuto così pervasivo che l'utente generalmente non legge l'informativa, finendo con il dare un consenso indiscriminato tutte le volte che gli viene richiesto. Se così non fosse, la navigazione sul web sarebbe dedicata la maggior parte del tempo alla lettura di documenti informativi sulla privacy.

A fronte di questi scenari incerti, e in particolare dei confini sfumati tra ricerca scientifica e ricerca commerciale, si potrebbe prevedere la possibilità che un individuo possa almeno esprimersi in alcune "condizioni minime" a priori, con la richiesta di una verifica a posteriori: chiedere di cedere i dati solo a condizione di conoscere il soggetto che li acquisisce (almeno distinguendo istituzioni pubbliche e private) e lo scopo della cessione; esigere trasparenza e partecipazione interattiva (con portali on line che facilitano la comunicazione), comunicazione dinamica dei risultati della ricerca, possibilità di "opt-out" ossia di uscire dalla ricerca se prende percorsi che l'utente non condivide e, conseguentemente, di cancellare i dati da quel momento; definire ambiti di ricerca ai quali obietta (ad

esempio, obiezione all'uso di informazioni per ricerche in ambito militare o altro).

Bisogna ricordare che sono i soggetti stessi che sono disponibili a "condividere" le informazioni, a non tenere privati i dati, a partecipare in modo attivo alla ricerca. L'importante è che siano consapevoli delle possibili implicazioni di tale partecipazione attiva e condivisione e della stessa possibilità dell'eventuale uscita dalla ricerca.

Un problema specifico emergente in ambito sanitario è il rapporto tra obbligatorietà dell'invio/condivisione di dati e assicurazioni. Il cliente/utente digitale dovrebbe essere libero di inviare o non inviare informazioni sulle sue abitudini, stili di vita, preferenze: le assicurazioni tendono ad incentivare la condivisioni di dati, e a proporzionare i costi/premi sulla base della condivisione e delle abitudini di vita. Si sta spostando dunque il rapporto cliente/assicurazioni dalla rilevazione dei dati medici anche alla condivisione di abitudini di vita, che potranno tradursi, per l'utente inconsapevole, in forme nascoste di controllo, che limitano la libertà individuale sulla base di una predefinita "standardizzazione" dei comportamenti considerati "migliori" (ma chi definisce lo standard?; non dovrebbe essere verificato caso per caso?). Lo stesso problema emerge nell'ambito del lavoro, tra lavoratori o potenziali lavoratori e datori di lavoro.

2.2. Trasparenza: l'etica degli algoritmi e la neutralità della rete

Uno degli elementi di maggiore problematicità etica dei *big data*, in particolare nell'ambito di applicazioni alla salute, è costituito dall'uso di algoritmi. La raccolta massiva dei dati è generalmente casuale e basata generalmente sulla rilevazione statistica della frequenza di comportamenti, ma la selezione nell'uso dei dati/informazioni, la costruzione di correlazioni (non causali) tra le informazioni e le predizioni (ossia la configurazione dei probabili scenari futuri, di comportamenti o altro) sono costruiti da algoritmi, elaborati da matematici e informatici. Gli algoritmi calcolano in base a variabili, ricercano correlazioni, disegnano predizioni.

È importante introdurre alcune riflessioni. I comportamenti selezionati sono solo quelli frequenti: la frequenza è l'elemento chiave della rilevazione statistica. Con la conseguenza che comportamenti non frequenti (isolati o rari) rischiano di essere marginalizzati nell'era digitale, con la possibilità che le persone tendono, per essere riconoscibili nel mondo digitale, a conformarsi ai comportamenti più diffusi o ripetuti nella società. La pressione alla omologazione, conformazione dei comportamenti (così come la pressione alla condivisione) evidenzia la riduzione degli spazi della libertà nell'era dei *big data*. Bisognerebbe invece rendere consapevole l'utente digitale che gli algoritmi costruiscono correlazioni e predizioni che riguardano la "probabilità" (calcolata matematicamente) e che non coincidono che la realtà (né con l'evidenza, o la causalità o l'imprevedibilità che deve rimanere sempre possibile e pensabile). Gli algoritmi costruiscono un'identità che diviene sempre più collettiva. Si parla di group identity. È la profilazione, che identifica la maggiore e minore probabilità o propensione ad agire in certi modi, di certi gruppi stratificati di individui.

Ma sulla base di quali criteri sono costruiti gli algoritmi? La rilevanza etica della trasparenza nell'uso degli algoritmi utilizzati dai provider riguarda il rilascio e il trattamento dei dati personali.

Un aspetto importante concerne dunque la pubblicazione completa di tutti i fattori di cui gli algoritmi di ricerca tengono conto. È infatti ammissibile che il segreto industriale venga esteso a comprendere anche i fattori che entrano in gioco nell'algoritmo utilizzato dal motore di ricerca? La questione è di rilevanza etica per almeno due ragioni. In primo luogo, solo una parte dei fattori in questione sono attualmente di dominio pubblico, favorendo fra gli utenti l'illusoria convinzione che siano questi, o prevalentemente questi, i fattori utilizzati. In secondo luogo, i provider d'informazione su rete svolgono un ruolo sempre più decisivo nella formazione della c.d. opinione pubblica. Ma l'opacità dei fattori su cui si basano i motori di ricerca non permette di monitorare la qualità dell'informazione e le eventuali manipolazioni della stessa opinione pubblica attraverso la rete.

La mancata trasparenza assume poi una rilevanza maggiore sul piano etico quando si consideri il ruolo delle pubblicità nell'ordinare i risultati di una ricerca in rete. Nulla vieta di associare i messaggi pubblicitari che pagano meglio ai siti sul piano informativo più completi. Il contrario - e cioè ordinare i siti tenendo conto dei messaggi pubblicitari che assicurano migliori introiti al provider - è invece eticamente insostenibile.

Collegata alla trasparenza è una delle domande più complesse: Chi ha il potere di decidere quali informazioni siano adatte a essere condivise e che cosa debba essere emendato anche temporaneamente? Vi sono opposte opinioni sul punto. Alcuni ritengono che un certo livello di supervisione sia necessario per qualsiasi piattaforma informatica, se si vuole che svolga un ruolo positivo per la società. Altri ritengono che questo sia concretamente impraticabile e citano le attività illegali degli hacker. Ma aggiungono anche che se esistono un organismo centrale o gruppi di interesse o di pressione che stabiliscono o favoriscono il rilascio delle informazioni, allora c'è il rischio di una informativa poco trasparente, controllata dalle idee e ideologie di chi prende tali decisioni e ciò a prescindere da quanto trasparenti o neutrali siano le stesse piattaforme. Queste preoccupazioni, che crescono in regimi politici non democratici, spingono a ritenere che la permanenza dei dati nel tempo costituisca un sistema di "controllo" da parte della società nei confronti dei poteri forti, anche a fronte di conseguenze negative (minacce alla sicurezza, scarsa tutela della privacy, interessi delle organizzazioni, ecc.). Più vasta è la documentazione e più è possibile ottenere una informativa esauriente, fare delle scelte consapevoli, informare le generazioni future.

Una tesi quest'ultima che non ha trovato ad oggi accoglienza fra i governi e le società occidentali, preoccupate dei rischi che implica una carente sicurezza informatica. Ne consegue che è possibile constatare l'investimento di risorse sempre più mirato a migliorare la protezione dei dati e dei documenti.

2.3. Veridicità e qualità dell'informazione: le possibili implicazioni bioetiche

Alcuni fenomeni informatici mostrano la problematicità della veridicità e della qualità dell'informazione, mettendo in evidenza la "opacità" di alcune tecnologie. Ne sono esempi la bolla informazionale e i fattori nascosti nell'ambito dei motori di ricerca, che possono avere implicazioni negative sulla salute.

I motori di ricerca sono servizi che favoriscono l'individuazione di siti di interesse per l'utente, ottenendo in cambio la profilazione dell'utente stesso[6]. Questi algoritmi di ricerca sono oggetto di dibattito all'interno della comunità scientifica e non solo. In particolare, le loro proprietà generano quantomeno due[7] criticità fondamentali:

• Bolla informazionale (*information bubble*).

Mediante l'algoritmo di ricerca utilizzato dai provider e dalle proprietà dei social, gli utenti vengono progressivamente separati dall'informazione che contrasta con il loro punto di vista e isolati intellettualmente nella propria "bolla informazionale". Un utente che cerca su internet "BP" ottiene informazioni sugli investimenti di *British Petroleum*, mentre un altro può ottenere informazioni sul disastro della *Deepwater Horizon* del 2010 anche noto come BP *disaster*. Le pagine possono essere anche molto diverse l'una dall'altra. La bolla informazionale funziona come un ecosistema "personale" o come una cornice "ideologica"[8]. È potenzialmente dannosa per l'individuo e per la società, limitando la diffusione di controargomenti, confutazioni e prove delle opinioni ed ipotesi accettate. Poiché uno dei filtri dell'algoritmo di ricerca sono le caratteristiche personali dell'utente, questi finisce per restare intrappolato in una bolla informazionale. Si potrebbe ipotizzare che i più "deboli" sono più esposti a questo effetto poiché le loro ricerche sui motori di ricerca possono essere usate/sfruttate a fini

[6] Esiste un motore di ricerca che non traccia i nostri dati, cioè che non sfrutta commercialmente i nostri input, differentemente da quanto fa, ad esempio, Google. È un motore di ricerca (https://duckduckgo.com) già disponibile online e immediatamente fruibile dal cittadino navigatore; per ulteriori informazioni al riguardo, si veda l'articolo pubblicato in http://www.fastweb.it/web-e-digital/duckduckgo-il-motore-di-ricerca-che-tutela-la-privacy, dedicato al funzionamento del motore di ricerca e alla sua storia. Duckduckgo.com è un'interessante e concreta realtà sulla quale riflettere, una realtà che sta crescendo, che non è di nicchia (si parla di 10 milioni di navigazioni al giorno). Potrebbe essere un temporaneo rimedio offerto al cittadino per tutelare maggiormente la propria privacy e la propria navigazione, un rimedio facile da applicare e di immediato impatto.

[7] Non trova invece conferma univoca un terzo effetto, noto come googlarchia o googlecrazia. Secondo alcuni studiosi, il *PageRank* produrrebbe una distribuzione *power-law* dell'informazione: le pagine con alto *rank* avrebbero maggiori probabilità di essere visitate, creando un circolo vizioso, amplificando cioè la popolarità dei siti più popolari. L'argomento è in realtà controverso. Secondo altri studiosi, l'uso dei motori di ricerca avrebbe addirittura un effetto egalitario, se paragonato con un algoritmo di ricerca casuale sul Web o con la ricerca effettuata dagli utenti a partire dai siti conosciuti.

[8] In *The Filter Bubble* (E. Pariser, *Il filtro. Quello che internet nasconde*, Il Saggiatore, Milano 2012) si osserva come, attraverso i filtri, "i creatori della personalizzazione ci offrono un mondo su misura, ogni aspetto del quale corrisponde perfettamente ai nostri gusti. È un modo rassicurante, popolato dalle nostre persone, cose e idee preferite" (E. Pariser, *Il filtro*, p. 16). Secondo l'autore, le *filter bubbles* danneggiano la società perché rendono la gente più vulnerabile ed esposta alla propaganda e alla manipolazione.

commerciali, cioè per reclamizzare loro prodotti che compenserebbero o risolverebbero le loro "difficoltà". Possono essere "difficoltà" di ogni sorta: carenze affettive e relazionali, malattie, povertà economica, ecc.. In questo caso, "debole" è chi (i) ha tali difficoltà e chi (ii) ha meno competenze informatiche o meno malizie informatiche, cioè adolescenti, anziani, utenti adulti che hanno cominciato ad usare internet e i social network senza che questi abbiano fatto parte della loro formazione.

• Fattori nascosti.

Secondo alcune fonti, le pagine in testa all'ordinamento sono quelle che associano pubblicità per le quali la compagnia riceve più soldi[9]. Queste fonti non sono state smentite.

Il *PageRank* è trasparente, ma esiste una lista di criteri di cu le compagnie informatiche tengono conto[10] che contiene anche centinaia di fattori. Questa lista non è ufficiale. Alcuni specialisti nel campo dei motori di ricerca hanno elaborato un'ipotetica lista basandosi sulla propria esperienza e su dati empirici[11]. Ufficialmente, le compagnie informatiche si limitano a dichiarare di tener conto fattori nell'ordinamento delle pagine, ma non pubblicano l'algoritmo, in quanto segreto industriale. L'utente quindi non sa come questi dati vengano utilizzati ed aggregati. Sappiamo solo che il *PageRank* costituisce uno di questi fattori.

Quali implicazioni bioetiche presentano tali criticità? Si pensi al malato che cerca cure attraverso i suddetti motori di ricerca. Ciò consente al motore di ricerca di raccogliere tali input e personalizzare la futura navigazione online del malato stesso (ossia la dinamica della bolla informazionale), costringendolo a (i) pensare e ripensare alla propria malattia ad ogni navigazione; (ii) visualizzare annunci di cure non comprovate; (iii) indurlo a credere in maniera crescente all'affidabilità di tali annunci ("l'ho visto su Internet!", "il Dott. .... ha anche un sito personale con le testimonianze dei pazienti curati che ora sorridono!", "se quest'annuncio ricorre spesso, sarà anche affidabile") con la possibilità di rinunciare ai canali di cura a lui prossimi e soprattutto consolidati (es. medico di base, farmacista, pronto soccorso, etc.). Questo meccanismo può ritrovarsi anche in altre figure particolarmente "deboli", delineate in precedenza. In sintesi, i più "deboli" possono entrare in una sorta di circolo vizioso alimentato dalle loro "difficoltà", ma accelerato dall'indicizzazione a fine commerciale (*big data*) che i motori di ricerca e i social network utilizzano. La bolla informazionale diventa una sorta di trappola (ben nascosta) per i più "deboli".

---

[9] Si vedano, per spiegazioni generali e di dettaglio, rispettivamente: https://support.google.com/adwords/answer/2497976?hl=en&ref_topic=3121763 e https://support.google.com/adwords/answer/1752122

[10] Si veda la lista seguendo il link http://www.searchenginejournal.com/infographic-googles-200-ranking-factors/64316/).

[11] Un esempio di studio eseguito per identificare i fattori e la relativa importanza è riportato in https://moz.com/search-ranking-factors. Ogni due anni circa, Moz effettua un'indagine per identificare i fattori più importanti che influenzano le ricerche in rete https://moz.com/search-ranking-factors. Uno studio ancor più preciso sulla scala temporale è riportato in http://www.cnet.com/news/testing-googles-panda-algorithm-cnet-analysis/.

Si consideri poi che il 72% delle persone che usa Internet, utilizza questo strumento per informazioni biomediche[12], ma purtroppo il numero, la varietà, e l'impatto delle "bufale" biomediche sono piuttosto impressionanti[13]. Si tratta non solo di bufale terapeutiche - che rappresentano comunque un sottoinsieme elevato della disinformazione in campo biomedico - ma anche di leggende e tesi complottistiche, di "scoperte" farmaceutiche che si sono rivelate vere e proprie truffe; di fallacie nei ragionamenti diagnostici e nell'individuazione di malattie apparenti; di falsa informazione concernente le correlazioni fra malattie esistenti o sui metodi di prevenzione – primo fra tutti i vaccini; di crociate contro le terapie ufficiali (è il caso della chemioterapia, contro la quale spesso si esercitano guaritori e ciarlatani).

I facili guadagni ricavati da terapie fasulle e pratiche inutili, quando non pericolose, non sono, apparentemente, l'unico motivo della pronta e vasta diffusione della disinformazione in campo medico (che a quanto pare ha una storia antica[14]). Alcune proprietà delle grandi tecnologie della comunicazione favoriscono la diffusione delle bufale. Dalle proprietà degli strumenti tecnologici a nostra disposizione dipende in larga misura la qualità dell'informazione circolante, e da questa dipende la salute pubblica. Una sola, allarmante, evidenza: l'impatto del movimento anti-vaccini misurato dal numero cumulativo di nuovi casi di morbillo per mese negli USA, per ogni anno dal 2001-2014. Secondo i dati CDC, solo nello scorso anno, si sono registrati 644 nuovi casi di morbillo nei 27 Stati dell'Unione. Si tratta del più grande numero di casi mai visto da oltre un quarto di secolo[15]. Occorre vigilare sugli strumenti informativi per tutelare la qualità dell'informazione e, in ultima analisi, la salute pubblica[16].

[12] Cfr. http://slideplayer.it/slide/3649692/.

[13] http://medbunker.blogspot.it/;
cfr. anche https://www.healthonnet.org/HONcode/Italian/?HONConduct117222.

[14] Salvatore Casillo, Federico Di Trocchio, Salvatore Sica, *Falsi giornalistici. Finti scoop e bufale quotidiane*, Guida, Milano 1997.

[15] Dati tratti da http://www.gravita-zero.org/2015/01/vaccini-negazionisti-movimento-antivaccini.html.

[16] A conferma di quanto riportato, in data 10.9.2016 abbiamo fatto una verifica dei risultati che emergono dal motore di ricerca Google, digitando il binomio "vaccino/autismo". I dati emersi sono senz'altro interessanti (e preoccupanti). Considerando i primi dieci siti risultanti dalla ricerca combinata dei due termini, si può sostenere che cinque destituiscano di ogni giustificazione la presunta correlazione tra vaccinazioni e insorgere dell'autismo; tre invece siano apertamente contrari alla vaccinazione, utilizzando toni fortemente polemici e allarmistici e alludendo anche all'esistenza di possibili complotti, come tali ovviamente non verificabili; due siti si mantengono invece su un livello più descrittivo, riferendo di alcuni casi di cronaca. Se può risultare confortante la presenza maggioritaria, nelle prime dieci posizioni, di siti ben documentati, che raccomandano il ricorso alle vaccinazioni, sulla base di comprovate ricerche scientifiche e il riferimento a significativi dati epidemiologici, sconcerta tuttavia scoprire che i tre siti antivaccinazione si collochino rispettivamente al secondo, terzo e quarto posto, quindi nelle posizioni più immediate, pronte per essere cliccate. Ovviamente in tali siti si tralascia di ricordare che "le pubblicazioni scientifiche che sostenevano un nesso causale tra vaccinazioni e insorgenza di autismo sono state ritrattate e (…) l'autore principale di tale tesi è stato condannato per frode" (Comitato Nazionale per la Bioetica, *L'importanza delle vaccinazioni*, mozione del 24 aprile 2015).

2.4. Dipendenza: un'ulteriore implicazione bioetica

Le ICT possono anche creare varie forme di dipendenza, una criticità dall'evidente rilevanza bioetica. Si parla al riguardo di *Internet Addiction Disorder* (IAD). Nella comunità scientifica c'è una sostanziale convergenza nel riconoscere che la dipendenza da internet ha delle ripercussioni negative sulla vita dell'individuo; ad essere colpita è soprattutto la sua dimensione relazionale, che viene fortemente compressa e di fatto ostacolata nelle concrete condizioni della vita quotidiana, dal lavoro alla famiglia, alla presenza più in generale in società. Più problematico è stabilire se tali effetti, certamente negativi, possano configurare una vera e propria forma di patologia, paragonabile per esempio alla dipendenza da sostanze. In tal senso va ricordato che l'ultima edizione del *Diagnostic and Statistical Manual of Mental Disorders* (DSM V) non include l'*Internet Addiction Disorder* all'interno della categoria diagnostica delle dipendenze comportamentali (*behavioral addictions*), lasciando pertanto intendere che al momento non vi sono sufficienti evidenze al riguardo. È stato però inserito l'*Internet Gaming Disorder*, con la precisazione che comunque è necessario sviluppare ulteriori ricerche per stabilire se esso possa essere considerato a tutti gli effetti un disturbo patologico.

Sta di fatto che fino a un certo momento gli effetti negativi dell'uso compulsivo di internet sono stati valutati in riferimento ad attività specifiche quali la pornografia online e il gioco d'azzardo; successivamente si è avanzata l'ipotesi che tali effetti interessino anche l'uso incontrollato della navigazione in generale, condotta nei più diversi ambiti[17]. Sette sono i sintomi che potrebbero indicare una qualche forma di disturbo legato all'utilizzo di internet: soddisfazione personale ricavata dal tempo trascorso in navigazione, scarso interesse per altre attività, ansia e depressione quando non si ha accesso alla rete, imperiosa necessità di controllare gli eventi sul web, frequenza crescente del ricorso alla rete rispetto alle solite abitudini, prolungati tempi di connessione e uso della rete anche in presenza di impedimenti fisici, lavorativi e sociali.

Vengono individuati cinque profili a seconda della dipendenza: *cyber-sexual addiction* (legato alla pornografia), *net-compulsion* (gioco d'azzardo e shopping), *information overload* (ricerca spasmodica di informazioni), *cyber-relation addiction* (abuso di social network) e *computer addiction* (utilizzo eccessivo di giochi online).

2.5. Giustizia partecipativa

Un altro problema meritevole di essere segnalato è quello dell'accesso a tali tecnologie. Esiste, oggi, un "divario digitale" a causa dell'età, della condizione socio-economica, dell'area geografica di appartenenza: anziani, persone meno colte, abitanti in Paesi in via di sviluppo sono i soggetti più vulnerabili dell'era digitale.

[17] Cfr. K.W. Müller et al., *A Hidden Type of Internet Addiction? Intense and Addictive Use of Social Networking Sites in Adolescents*, "Computers in Human Behavior", 2016, 55, pp. 172-177; K. Yung et al., *Internet Addiction Disorder and Problematic Use of Google Glass™ in Patient Treated at a Residential Substance Abuse Treatment Program*, "Addictive Behaviors", 2015, 41, pp. 58–60.

Dati gli elementi di positività della connessione, andrebbe garantito un equo accesso, consentendo a tutti di acquisire strumenti, conoscenze, competenze e motivazione all'uso delle nuove tecnologie informatiche per partecipare democraticamente alla società globale e non essere emarginati dalla rete. Al tempo stesso, va garantito - almeno temporaneamente, in attesa della implementazione delle tecnologie - un accesso alternativo ai servizi (in particolare nell'ambito sanitario) alle persone o gruppi che non possiedono le tecnologie o non possiedono le competenze per utilizzarle. L'obiettivo dovrebbe essere quello di consentire a tutta la partecipazione secondo il principio di uguaglianza, pari opportunità e non discriminazione.

Ciò che va evitato è la c.d. "medicina a due velocità": veloce per chi dispone delle nuove tecnologie e lenta per chi non può partecipare. All'equità di accesso deve anche corrispondere l'uguaglianza nell'educazione all'accesso alle ICT, in modo consapevole e critico rispetto alle problematiche etiche emergenti. Inoltre i risultati positivi dell'uso delle ICT per la salute dovrebbero essere condivisi da tutti, secondo il principio sempre più consolidato nella bioetica globale del "*benefit sharing*".

Andrebbero poi implementati gli strumenti che sollecitano alla partecipazione degli utenti ad iniziative digitali che offrono nuove opportunità di confronto, condivisione, conoscenza. Si pensi alle nuove piattaforme create per consentire a soggetti che partecipano a sperimentazioni di essere costantemente aggiornati sull'avanzamento delle ricerche; a piattaforme elaborate per consentire ai pazienti la condivisione, particolarmente utili e preziose nell'ambito delle malattie rare.

2.6. La governance

Le ICT e i *big data* hanno reso dunque problematici concetti come privacy, autodeterminazione, consenso informato e, in ultima analisi, diritto individuale, così come conosciuti e trattati finora. A fronte di ciò la tutela della sicurezza e della privacy (intesa, come detto, in senso allargato) devono essere una responsabilità condivisa fra le gli utenti, aziende, e le istituzioni. Si discute sulla opportunità che Società informatiche oltre a contendersi l'*hosting* di *genomic data*[18], debbano mettere in atto strumenti efficaci che possano consentire agli utenti di acquisire consapevolezza; si discute sulla rilevanza che istituzioni possano pianificare una governance efficace in questo settore. E spetterà proprio a questi ultimi il compito di sfruttare in modo corretto tali possibilità nella piena consapevolezza che il non utilizzo si traduce in una perdita di privatezza e di sicurezza, a fronte dell'accumularsi dei dati. È per altro noto agli esperti di informatica che il c.d. tasto "cancella" non garantisce la perdita dei file, e-mail e messaggi, che possono essere recuperati con grande facilità (persistenza dei dati). Questa conoscenza digitale di qualsiasi cittadino, che svolga attività pubblica o professionale, può e potrà produrre all'interno della società

[18] J.T. Wilbanks, E.J. Topol, *Stop the Privatization of Health Data. Tech Giants Moving into Health May Widen Inequalities and Harm Research, Unless People can Access and Share their Data*, "Nature", 2016, 535(7612), pp. 21-7.

conseguenze rilevanti. Il passato, documentato, del cittadino potrà tornare ed essere utilizzato per influenzare la sua immagine sul luogo di lavoro e nella vita di ogni giorno.

Il contesto sin qui sommariamente descritto è già di per sé sufficiente a cogliere la ragione per cui la governance e regolamentazione delle modalità di raccolta dei dati e di conservazione degli stessi siano destinate ad incidere in maniera rilevante non solo sulla tutela dei singoli e dei gruppi, ma anche sugli interessi di chi è intento a sfruttare economicamente tali informazioni. Certamente le strutture che organizzano i grandi data center, che implicano una concentrazione nello spazio delle informazioni, saranno protette, pur tuttavia la logica dell'accentramento rende tali data center obiettivi assai appetibili per gli autori dei furti di dati. Infatti se le informazioni hanno un valore commerciale per le imprese, se sono divenute merce di scambio fra le stesse e in parte fra imprese e clienti, è ovvio che esse risultano anche appetibili a coloro che sono mossi dall'intento di conseguire un profitto illecito, ovvero, in un diverso ambito, come quello politico, sono strumenti che consentono vere e proprie guerre informatiche.

Vanno, pertanto, rafforzati i meccanismi di tutela dei dati e la difesa dei propri sistemi da attacchi degli hacker. Ma a fronte di una comune volontà in tutte le nazioni, ove l'economia digitale si è andata affermando, di regolare proprio gli accessi illegittimi ai data, di contro si riscontra un'evidente difficoltà nel definire le regole che governano la materia. Siamo in un territorio di confine dove si intersecano le esigenze di protezione degli individui, gli interessi lobbistici delle imprese di settore, la volontà degli Stati di garantire un ecosistema informativo efficiente e sicuro. E il bilanciamento di queste opposte esigenze, nel panorama internazionale globale, è conseguito con prospettive differenti in considerazione del diverso modo di considerare e rapportare i dati ora come merce ora come parti della personalità e in considerazione delle diverse forme di governo, che hanno una ricaduta sul maggiore o minore controllo e tutela dei dati personali (si pensi ai modelli repressivi presenti in diverse parti del mondo). Tuttavia la condivisione degli interessi in gioco può ben consentire di avviare su queste esigenze un dialogo fra i vari modelli regolatori, al fine di facilitarne la convergenza verso soluzioni condivise. Soluzioni convergenti sempre più necessarie data la globalizzazione dell'informazione e l'inevitabile passaggio "trans-border" delle informazioni.

Comunque, si sta diffondendo nelle aziende tecnologiche maggiore attenzione nei confronti della comunità virtuale, un'attenzione che si traduce nella tendenza che tutti i prodotti e servizi online prevedano che l'utente accetti termini e condizioni rispettando linee guida contrattuali. E dai legislatori è sempre più avvertita la necessità di affiancare all'affermazione della centralità del consenso, l'istituzione di specifiche autorità dotate di capacità di analisi e di controllo che il singolo non può avere e in grado d'incidere sulle politiche adottate dalle imprese dell'ICT. Ma resta il fatto che, malgrado questi sforzi verso le garanzie dell'utente, il cittadino deve essere ben consapevole che la tutela di una persona i cui dati finiscono sul Web è molto difficile, per mancanza di strumenti efficaci e che questa è prevalentemente nelle sue mani, nel momento in cui

considererà la qualità del prodotto, ma anche la facilità con cui le aziende permetteranno di monitorare e garantire la riservatezza.

La difficoltà a regolare la materia non deve portare il legislatore a rinunciare ai suoi compiti: predisporre norme che tutelino non solo la protezione dei dati personali (confidenzialità, riservatezza, privacy)[19], ma anche la libertà personale[20] con adeguate leggi anti-discriminazione, che consentano di esercitare una vigilanza adeguata sui "rischi sociali" delle nuove tecnologie. Occorrono, in specie, strumenti normativi che puniscano l'abuso nell'uso dei dati personali e delle informazioni sulla vita privata degli utenti o la pubblicizzazione non autorizzata di materiali informatici a scopi denigratori; che puniscano l'intenzionale uso di algoritmi che su un piano meramente predittivo (rispetto ad abitudini e comportamenti) producono come conseguenza una marginalizzazione di categorie "minoritarie" con stigmatizzazione sociale.

3. Raccomandazioni

Alla luce delle precedenti considerazioni il CNB intende sottolineare come nell'era dei big data il problema consista non tanto nell'uso dei dati, quanto piuttosto nell'uso "appropriato" dei dati per il bene dell'uomo e della sua salute. A tal fine propone le seguenti raccomandazioni:

- Individuare e definire la responsabilità dei provider, soprattutto in talune circostanze, di cui si rileva socialmente una condizione di particolare rischio: tale responsabilità non deve limitarsi alla dichiarazione relativa all'uso commerciale dei dati personali, bensì estendersi a una verifica della qualità dei dati e a una trasparenza degli algoritmi dei motori di ricerca; impegnare inoltre le autorità per l'informazione e la comunicazione a un adeguato controllo.

- Individuare strumenti efficienti per richiedere il consenso o dissenso al trattamento dei dati all'utente, basato su informazioni leggibili e sintetiche relative al tipo di dati e modalità di raccolta, alle finalità per cui vengono utilizzati nonché alle procedure di trattamento e di analisi dei medesimi. Nella misura in cui tali informazioni non potessero essere fornite, in ogni caso l'utente deve esplicitamente dare prova della sua presa di coscienza dei limiti alla privacy, della possibilità di re-identificazione, dei possibili usi anche commerciali dei suoi dati.

- Dare attuazione a un riconoscimento effettivo del diritto all'oblio, stabilendo con procedure chiare e trasparenti la possibilità per il soggetto di richiedere la cancellazione dei dati personali, di modo che questi non

[19] È questa l'impostazione promossa nel documento F. Caldicott (ed.), *Information: to Share or not to Share? The Information Governance Review*, March 2013. Si sottolinea la necessità, in ambito sanitario, di bilanciare le esigenze di privacy con le esigenze di condivisione di informazioni tra medici e sanitari per l'interesse del paziente e della salute pubblica.

[20] C. Bock, *Preserve Personal Freedom in Networked Societies. Broad Anti-Discrimination Laws and Practices could Compensate for Failing Data Protection and Technology-Linked Loss of Privacy*, "Nature", 2016, 537(7618), p. 9.

siano più accessibili al pubblico sotto qualsiasi forma (copie o riproduzione).

• Promuovere campagne stampa, campagne pubblicitarie (per esempio Pubblicità Progresso), e programmi educativi sul funzionamento delle tecnologie sociali che gestiscono informazioni fra utenti e fra utenti e gestori. L'obiettivo è di coinvolgere attivamente gli utenti e consentire la presa di coscienza critica dei problemi etici emergenti delle nuove tecnologie, soprattutto per le persone particolarmente vulnerabili, e renderle consapevoli dei possibili rischi nella messa in rete di informazioni e nella condivisione di informazioni.

• Sollecitare le istituzioni sanitarie pubbliche, ai vari livelli, a predisporre e a tenere aggiornato un sito informativo, presso il quale il cittadino possa eventualmente trovare conferma o smentita ufficiale sulla veridicità e qualità delle notizie in campo medico[21].

• Sollecitare il MIUR a redigere e diffondere nelle scuole Linee Guida per un corretto uso delle tecnologie sociali. Occorre favorire l'emergere di una nuova "netiquette"[22], che prescriva l'uso socialmente e psicologicamente sostenibile, ed eticamente consapevole dei rischi, di internet e dei socials, con particolare attenzione alle questioni attinenti alla salute. Si dovrebbe fin dai primi anni di vita, soprattutto con l'aiuto dei genitori e dei sistemi scolastici, fare emergere l'esigenza della tutela della riservatezza di talune informazioni e le modalità di difesa della propria privacy virtuale. Una educazione alla cittadinanza attiva, che voglia essere adeguata ai tempi (una sorta di Educazione civica 2.0.) non può affatto sottovalutare la rilevanza di questi temi. Ne consegue la raccomandazione che siano istituiti programmi educativi che permettano agli utenti di sviluppare tale consapevolezza nel contesto di una cultura tecnica sulle modalità di uso di internet[23].

• Sostenere la ricerca per l'innovazione dell'approccio etico al disegno delle tecnologie sociali. Finora, l'etica in ICT si è limitata essenzialmente alla protezione della privacy. Ma come abbiamo visto, questo è solo uno dei problemi etici che le tecnologie sociali presentano. Occorre investire ricerca sui sistemi bioeticamente compatibili. Ad esempio, allo scopo di arginare il fenomeno della dipendenza, si potrebbero immaginare sistemi automatici di alerting, che avvertano l'utente di aver oltrepassato una soglia critica di tempo di connessione. Anche l'offerta automatica di strumenti di *self-assessment*, da somministrarsi prima di stabilire la connessione (una sorta di cookies sanitarie) potrebbero rendere l'utente più consapevole dei rischi che corre.

[21] Va senz'altro in questa direzione l'allestimento, a cura dell'Istituto Superiore di Sanità, del cosiddetto *Portale della conoscenza*, i cui contenuti saranno a breve consultabili.
[22] Il vocabolo si riferisce a *network* (rete) e *étiquette* (buona educazione).
[23] Va da sé che questa sorta di *Educazione civica 2.0.* potrebbe riguardare anche argomenti e temi che hanno una più diretta rilevanza bioetica, in coerenza del resto con quanto era stato auspicato dal Protocollo d'intesa tra MIUR e CNB, sottoscritto il 15.7.2010. Tale protocollo individuava nella formazione bioetica, da promuovere all'interno del sistema scolastico, una delle possibili esemplificazioni della più generale formazione dedicata a *Cittadinanza e Costituzione*.

• Garantire condizioni di accesso a tutti coloro che intendono avvalersi delle nuove tecnologie, e al tempo stesso garantire i diritti di chi non può/non intende connettersi, in modo specifico per quanto attiene l'accesso a servizi per la salute dei cittadini.

Sollecitare una disciplina per la protezione dei dati personali (basate sulla tutela della riservatezza e confidenzialità, sulla minimizzazione dell'uso dei dati sensibili, sulla giustificazione e proporzionalità della raccolta, sulla responsabilità di chi li usa); sollecitare inoltre l'adozione di normative volte a predisporre adeguate leggi anti-discriminazione, che contribuiscano ad evitare o almeno mitigare i rischi sociali dell'abuso dei dati virtuali: l'obiettivo è quello di proteggere, oltre ai dati personali, anche e soprattutto la libertà personale dei cittadini.

APPENDICE: Nota giuridica

In Italia il trattamento dei dati personali trova una prima protezione nella legge 675 del 31 dicembre 1996 "Tutela delle persone e degli altri soggetti rispetto al trattamento dei dati personali", poi abrogata ai sensi dell'art.183, comma 1, lettera a) del successivo decreto legislativo n. 196 del 30 giugno 2003 ("Codice in materia di protezione dei dati personali") attualmente vigente. Secondo il Codice il trattamento dei dati personali deve essere improntato ai principi di correttezza, liceità e trasparenza e di tutela della riservatezza e dei diritti di chi li fornisce.

Significative modifiche a tale normativa sono state apportate dal legislatore con il d.l. 69/2012 in attuazione delle direttive 2009/136/CE in materia di trattamento dei dati personali e tutela della vita privata nel settore delle comunicazioni elettroniche, e del regolamento CE, n. 2006/2004 sulla cooperazione tra la autorità nazionali responsabili dell'esecuzione della normativa a tutela dei consumatori. Un complesso di modifiche che si contestualizzano nel più generale panorama delle dinamiche inerenti la circolazione dei dati personali ed il loro sfruttamento per finalità commerciali, alla luce delle linee di riforma emergenti in tale ambito sia in Europa che in diverse parti del mondo. Nello specifico, i due temi centrali su cui verte il d.l. sono la profilazione degli utenti durante la navigazione online e la gestione degli eventi di *data breach* e ciò in ragione della crescente importanza in chiave predittiva assunta dalle informazioni, ed in particolare da quelle personali, nelle economie più avanzate.

Il Garante per la Protezione dei Dati Personali emanava i provvedimenti: n. 229 dell'8 maggio 2014 relativo alla "individuazione delle modalità semplificate per l'informativa e l'acquisizione del consenso per l'uso dei cookie" e n. 353 del 10 luglio 2014 nei confronti di Google Inc. sulla "conformità al Codice dei trattamenti di dati personali effettuati ai sensi della nuova privacy policy"; il 19 marzo 2015 le "Linee guida in materia di trattamento di dati personali per profilazione online"[24].

[24] Pubblicate in Gazzetta Ufficiale il 6 maggio 2015 (Linee Guida) V. il sito: http://www.portolano.it/pcc_newsletters/profilazione-online-le-linee-guida-del-garante-privacy/.

Il Garante ha stabilito che il provider di servizi in rete (motori di ricerca, posta elettronica, social networks, cloud, pagamenti online, analisi statistica e monitoraggio dei visitatori di siti web, ecc.) non potrà utilizzare dati a fini di profilazione, senza il consenso degli utenti, e dovrà dichiarare esplicitamente di svolgere questa attività a fini commerciali. Si tratta del primo provvedimento del genere in Europa, emerso nell'ambito di un'azione coordinata con le altre Autorità europee e a seguito della pronuncia della Corte di Giustizia europea sul diritto all'oblio. In sintesi, le Linee guida sono finalizzate affinché tutti i fornitori dei servizi della società dell'informazione, nonché tutti i soggetti che comunque offrono ai propri utenti servizi online accessibili al pubblico attraverso reti di comunicazione elettronica, con specifico riguardo ai trattamenti dei dati personali relativi all'utilizzo delle funzionalità offerte tengano conto segnatamente: - dell'l'informativa da dare agli interessati di cui all'art. 13 del Codice e secondo quanto indicato al paragrafo 3 delle presenti Linee guida; - del consenso preventivo degli utenti e delle informazioni che li riguardano, anche derivanti dal trattamento, in modalità automatizzata, dei dati personali degli utenti autenticati in relazione all'utilizzo del servizio per l'inoltro e la ricezione di messaggi di posta elettronica; - del rispetto del diritto di opposizione di cui all'art. 7 del Codice; - dell'adozione di una policy di *data retention* conforme al principio di finalità di cui all'art. 11 del Codice[25].

Nell'ambito del trattamento dei dati sensibili a livello europeo (anche se non specificamente riguardanti i *big data*) vanno ricordate la Direttiva 95/46/CE 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 dei dati, poi abrogata con il Regolamento (UE) 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016[26] (di cui si dirà nel seguito) e la Direttiva 2002/58/CE del Parlamento Europeo e del Consiglio del 12 luglio 2002 relativa al trattamento dei dati personali e della tutela della vita privata nel settore delle comunicazioni elettroniche, poi modificata dalla Direttiva 2009/136.

Tra i documenti più significativi sull'argomento adottati dal Consiglio d'Europa vi sono: *Convention for the protection of individuals with regard to automatic processing of personal data*, Council of Europe, 1981 e protocollo addizionale (*Additional Protocol to Convention ETS No.108 on Supervisory Authorities and Transborder Data Flows*); *Recommendation CM/Rec(2010)13 of the Committee of Ministers to member states on the protection of individuals with regard to automatic processing of personal data in the context of profiling* (23 November 2010); *Recommendation CM/Rec(2010)13 on the protection of individuals with regard to automatic processing of personal data in the context of profiling* (23 November 2010); *Recommendation CM/Rec(2012)4 on the protection of human*

[25] Per una illustrazione delle Linee guida e delle misure richieste si veda http://www.diritto24.ilsole24ore.com/art/guidaAlDiritto/dirittoCivile/2014-07-21/google-garante-trasparenza-profilazione-125008.php?refresh_ce=1.

[26] Parlamento Europeo, Consiglio dell'Unione Europea. Regolamento (UE) 2016/679 del Parlamento Europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la Direttiva 95/46/CE (Regolamento generale sulla protezione dei dati). Gazzetta Ufficiale dell'Unione Europea 4 maggio 2016, L119: 1-88.

rights with regard to social networking services; *Recommendation CM/Rec(2014)6 on human rights for Internet users*; *Recommendation CM/Rec (2016)1 on protecting and promoting the right to freedom of expression and the right to private life with regard to network neutrality*.

Va inoltre segnalato come la Corte di Giustizia dell'Unione europea[27] abbia recentemente invalidato la precedente decisione della Commissione europea (2000/520/CE), che accertava una tutela adeguata della privacy dei dati anche personali trasferiti negli Stati Uniti (sistema c.d. *Safe Harbour*). Ostacolando di fatto le autorità nazionali nel verificare il livello di sicurezza garantito ai dati oltre oceano e non prevedendo alcun meccanismo di aggiornamento dei criteri del *Safe Harbour*, la decisione del 2000 della Commissione è stata giudicata lesiva del contenuto essenziale del diritto al rispetto della vita privata, permettendo, ad esempio, ad autorità pubbliche statunitensi di accedere in maniera generalizzata, anche per finalità di profilazione, al contenuto delle comunicazioni elettroniche lì trasferite. In questo modo, la Corte di Giustizia ha ricondotto alle singole autorità nazionali la competenza a giudicare dell'adeguatezza dei livelli di sicurezza dei dati trasferiti, invitando, per lo meno implicitamente, la Commissione a rinegoziare l'accordo sul trasferimento dei dati. Più recentemente il 12 luglio 2016 la Commissione europea ha adottato un accordo per tutelare "i diritti fondamentali di qualsiasi persona nell'UE i cui dati personali siano trasferiti verso gli Stati Uniti" e apportare "chiarezza giuridica alle imprese che operano con trasferimenti transatlantici di dati". Nello scudo UE-USA per la privacy sono state inserite alcune disposizione specifiche in tema di health data[28].

Infine, il trattamento dei dati personali ha ricevuto l'attenzione del Parlamento Europeo e dalla Corte di Giustizia anche per quel che riguarda la protezione delle persone fisiche dal trattamento dei dati personali nel corso di indagini e accertamenti di reati, nonché dalla loro libera circolazione. In particolare, sia la Direttiva (UE) 2016/680[29] emanata dal Parlamento Europeo e dal Consiglio il 27 aprile 2016, sia il già citato Regolamento (UE) 2016/679[30] emanato alla stessa data, pongono a base della tutela degli interessati, cioè di coloro cui si riferiscono i dati personali, "il diritto di non essere oggetto di una decisione che valuta aspetti personali" che lo/la "concernono basata esclusivamente su un trattamento automatizzato e che (…) incida significativamente sulla sua persona. In ogni caso tale trattamento dovrebbe essere subordinato a garanzie adeguate, compreso il rilascio di specifiche informazioni all'interessato (…)" (art. 38 Direttiva).

---

[27] Sentenza del 6 ottobre 2015, causa C-362/14, Maximillian Schrems/Data Protection Commissioner.

[28] Cfr. http://europa.eu/rapid/press-release_IP-16-2461_it.html.

[29] Parlamento Europeo, Consiglio dell'Unione Europea. Direttiva (UE) 2016/680 del Parlamento Europeo e del Consiglio, del 27 aprile 2016, relativa alla protezione delle persone fisiche con riguardo al trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, nonché alla libera circolazione di tali dati e che abroga la decisione quadro 2008/977/GAI del Consiglio. Gazzetta Ufficiale dell'Unione Europea 4 maggio 2016; L191:89-131.

[30] http://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32016R0679&from=IT.

Mentre la Direttiva 2016/680 verte sull'ambito specifico del trattamento dei dati personali da parte delle autorità competenti a fini di prevenzione, indagine, accertamento e perseguimento di reati o esecuzione di sanzioni penali, il Regolamento 2016/679 è di grande rilievo per le tecnologie dell'informazione e della comunicazione perché, tra l'altro, pone le basi per l'esercizio di nuovi diritti e definisce i limiti al trattamento automatizzato dei dati personali.

Il Regolamento stabilisce che le persone interessate nel trattamento di dati personali siano informate del diritto di revocare il consenso a determinati trattamenti, tra cui, per esempio, i trattamenti a fini di marketing diretto. Prevede, inoltre, che i fornitori di servizi internet e i social media debbano chiedere il consenso ai genitori o a chi esercita la potestà genitoriale per trattare i dati personali dei minori di 16 anni.

Particolarmente significativo nell'ambito delle tecnologie sociali è il "diritto all'oblio" previsto dal Regolamento: gli interessati possono ottenere la cancellazione dei propri dati personali anche online da parte del titolare del trattamento qualora ricorrano alcune condizioni: se i dati sono trattati solo sulla base del consenso; se i dati non sono più necessari per gli scopi rispetto ai quali sono stati raccolti; se i dati sono trattati illecitamente; oppure se l'interessato si oppone legittimamente al loro trattamento. Al "diritto all'oblio" si accompagna l'obbligo per il titolare del trattamento che ha pubblicato i dati di comunicare la richiesta di cancellazione a chiunque li stia trattando, nei limiti di quanto tecnicamente possibile. Secondo il Regolamento, il "diritto all'oblio" può essere limitato solo in alcuni casi specifici: per esempio, per garantire l'esercizio della libertà di espressione o il diritto alla difesa in sede giudiziaria; per tutelare un interesse generale (ad esempio, la salute pubblica); oppure quando i dati, resi anonimi, sono necessari per la ricerca.

Il Regolamento, inoltre, introduce il "diritto alla portabilità" dei propri dati personali per trasferirli da un titolare del trattamento ad un altro. Per esempio, si potrà cambiare il provider di posta elettronica senza perdere i contatti e i messaggi salvati.

Nel Regolamento si conferma il divieto di trasferimento di dati personali verso Paesi situati al di fuori dell'Unione Europea o organizzazioni internazionali che non rispondono agli standard di adeguatezza in materia di tutela dei dati, rispetto ai quali il Regolamento introduce criteri di valutazione più stringenti.

Significativo per le tecnologie dell'informazione della comunicazione è anche il principio "privacy by design", in base al quale è necessario garantire la protezione dei dati fin dalla fase di ideazione e progettazione di un trattamento o di un sistema.