Big Data: istruzioni per un corretto uso.
Responsabilità pubblica e ricadute per l'industria dei media
e per il servizio pubblico nell'era crossmediale
Sede istituzionale: Spazio Europa
via IV Novembre 149 I-00187 Roma
mercoledì 20 settembre 2017 ore 9.30. 13.00

Big Data: governance e responsabilità pubblica per la protezione dei dati personali
Infocivica, 20 settembre 2017

Manlio Cammarata, direttore di InterLex e coordinatore del Forum "La cittadinanza digitale" nel ventennale della prima legge italiana sulla protezione dei dati e del primo numero della rivista

Non posso parlare di dati personali e responsabilità pubbliche senza andare col pensiero a Stefano Rodotà, che ci ha lasciato di recente dopo una vita dedicata allo studio e all'insegnamento di questi temi. Rodotà poneva al centro del suo discorso *la persona*. Non semplicemente il *soggetto* titolare di diritti e doveri, caro ai codici e ai giuristi, ma *la persona* che si proietta nel mondo e costruisce stessa *anche* attraverso il controllo dei propri dati.
È in questa luce che intendo trattare il tema che mi è stato assegnato: con un ricordo affettuoso al Maestro scomparso e cercando di non tradire il suo insegnamento con dettagli che possono sembrare poco importanti, ma che mi sembrano necessari per comprendere meglio il quadro generale dei diritti del *cittadino digitale* e dei problemi che li minacciano.

La questione della protezione della vita privata, e quindi del trattamento dei dati personali, è sorta negli anni '80 del secolo scorso, con la diffusione delle banche dati elettroniche e le aumentate capacità di archiviazione ed elaborazione delle informazioni. L'internet di massa non c'era, era uno strumento di comunicazione per pochi specialisti.
Oggi, con l'esplosione delle connessioni e delle capacità di archiviazione ed elaborazione offerte dai Big Data, esplode anche il problema della protezione della sfera privata, perché i dati personali hanno assunto un rilevante valore economico e sono sfruttati come *merce di scambio* per l'accesso ai servizi della *società della conoscenza*. E come *prodotto da vendere sul mercato*.
Questo è il primo punto che ciascuno di noi dovrebbe considerare ogni volta che alza le spalle di fronte agli ammonimenti sulla protezione dei propri dati: qualcuno guadagna molti soldi vendendo le informazioni che lui stesso gli dà gratis.

La raccolta e l'elaborazione dei Big Data, che richiedono gigantesche risorse di archiviazione e di calcolo, sono alla portata di poche aziende di dimensioni globali, i cosiddetti *Over The Top* (OTT). Quante siano queste aziende è materia di discussioni che qui non ci interessano (per i francesi sono quattro, per altri sono sei – le "sei sorelle" eccetera). Comunque sono tutti d'accordo nell'elencare almeno Google, Facebook, Microsoft e Apple, ma anche Netflix, Yahoo!, Alibaba…). Grazie ai Big Data queste aziende detengono un potere globale che sotto molti aspetti è più forte di quello delle

grandi compagnie petrolifere e delle aziende chimiche e farmaceutiche che controllano il mercato dell'alimentazione e della salute.
Altre entità che trattano grandi quantità di dati sono le strutture interbancarie e le organizzazioni messe in piedi dalle agenzie di assicurazione: i dati elaborati da queste entità, soprattutto se non abbastanza precisi, possono provocarci gravi problemi, come la difficoltà di ottenere un finanziamento o il maggior costo di una polizza di assicurazione.

Dei Big Data ormai sappiamo, o crediamo di sapere tutto. Parliamo di aziende *Over The Top*, di banche dati, di intelligenza artificiale, di *machine learning*... e ci dimentichiamo del primo anello della catena, il fornitore di informazioni, ovvero la persona connessa alla Rete, quello che dovrebbe essere il *cittadino digitale*.
Da una parte il cittadino digitale fornisce direttamente una quantità di informazioni sulla sua vita privata e lavorativa attraverso i social network. Ma di solito non si rende conto di quale sia il valore economico di queste informazioni e di quanto possano essere usate a suo danno.
Dall'altra parte una quantità ancora più grande di dati viene dalla raccolta dalle "tracce" che ciascuno lascia inevitabilmente sulla Rete navigando, compiendo ricerche, acquistando online o fisicamente con le diverse forme di denaro elettronico. Dati che rendono sempre più preciso e dettagliato il suo *profilo* nell'ambito dei Big Data.

La diffusione dei dispositivi cosiddetti "intelligenti" complica e rende più penetrante il controllo della vita privata. Sono apparecchi di uso quotidiano, di cui tutti si servono, che sono connessi in rete e quindi consentono a chi li controlla di raccogliere informazioni sempre più intime e dettagliate sui loro proprietari. L'esempio clamoroso dei televisori che trasmettono a "qualcuno" immagini e suoni dell'ambiente in cui sono installati, anche quando appaiono spenti, per adesso è solo il più eclatante.

In genere noi interagiamo con i dispositivi "intelligenti" attraverso il telefonino, che in questo modo diventa sempre più il "centro di comando e controllo" della nostra vita. E quindi, essendo sempre connesso alla Rete, è la prima fonte di informazioni su chi lo usa.
Essere sempre connessi, scaricare "app", parlare con la lavatrice via telefonino, conservare i propri dati nel cloud: sono solo alcuni degli stimoli che vengono proposti, con un'insistenza che dovrebbe essere sospetta. In qualche caso usare il cloud è *imposto* (anche se non tecnicamente necessario) per compiere certe operazioni, come copiare la rubrica da un telefonino Android a un altro, sempre Android. O per prendere appunti su un telefonino Microsoft. Scaricare e usare una app comporta necessariamente l'autorizzazione della stessa app alla cattura di tutto quello che c'è nell'apparecchio, compresi i contenuti della messaggistica e delle email.

Tutto questo in un mondo dove un numero enorme di individui – non solo giovani – hanno lo smartphone come strumento principale, se non unico, di comunicazione, di informazione e di organizzazione della propria vita. In questo modo forniscono sempre più dati e ricevono sempre più informazioni e stimoli, somministrati proprio in funzione dei dati che loro stessi hanno fornito e continuano a fornire. Si crea così un circolo vizioso da quale è difficile uscire.

Anche la navigazione da personal computer mette a rischio la riservatezza. I siti apparentemente più rispettabili sono spesso anche i più invasivi della privacy. Quando un sito mi chiede il consenso per i cookie, i cookie li ha già depositati nel mio computer e stanno lavorando per lui. Nella maggior parte di casi ci sono almeno i cookie "statistici" di Google.
Ma, ribatte qualcuno, i dati che vengono raccolti con i cookie statistici di Google sono anonimi. Certo, non recano attaccato un cartellino con nome e cognome dell'interessato, ma contengono anche alcune righe di codice che permettono di riconoscere almeno il computer di provenienza. E, per i sistemi di Big Data di Google, bastano frazioni di secondo per collegare una macchina a un nome.

L'invasività delle app non ha confini. Quando comperi un telefonino Samsung nuovo, ci sono già installate delle app che – dicono – ti aiutano a tenere sotto controllo il tuo stato di salute. Perché non attivarle? Può essere utile controllare il battito cardiaco o altri parametri vitali. E così c'è qualcuno che raccoglie i tuoi dati sanitari, dati sensibili, senza che sia stata rispettata la procedura di legge: consenso informato reso per iscritto e autorizzazione del Garante. Non mi risulta, infatti, che il Garante italiano per la protezione dei dati personali abbia autorizzato Samsung e Google a trattare i miei dati sanitari. E, se lo ha fatto, certo era in un momento di distrazione.

E qui arriviamo al problema centrale di questo discorso: quale *governance* sia necessaria per la protezione della privacy nell'era del Big Data, quali siano le responsabilità pubbliche.
Ma subito balza all'occhio una situazione paradossale: anche i siti della pubblica amministrazione italiana presentano l'informativa sui cookie "statistici e anonimi" di Google. Di fatto, poiché almeno in molti casi per Google è facile de-anonimizzare i dati, la pubblica amministrazione fornisce a "Big G" una grande quantità di dati personali, anche sensibili, dei cittadini italiani.

Ma non basta: anche sui siti pubblici ci sono i link a siti esterni e si declina qualsiasi responsabilità (come sembra ovvio) per i trattamenti di dati svolti da tali siti. I quali possono quindi acquisire i dati della precedente navigazione dell'utente, grazie ai cookie "tecnici" (che si considerano inoffensivi) depositati dai siti visitati in precedenza e che permangono fino alla chiusura del browser (ma, anche in questo caso, ci sono molti dubbi che vengano realmente cancellati, come dimostrerebbero alcune recenti analisi).

Le norme che regolano il trattamento dei dati personali esistono da più di vent'anni, in Europa e in Italia. Ora un nuovo regolamento europeo, oltre ad armonizzare la normativa negli Stati membri, dovrebbe rendere più incisiva la protezione della vita privata alla luce delle tecnologie più avanzate e dell'invadenza dei Big Data.
Funzionerà? Ho qualche dubbio, ma non è questa la sede per discuterne.
Piuttosto voglio rilevare una specie di paradosso, che mi sembra illuminante per capire quali difficoltà si oppongono alla *governance* dei dati personali.

Il futuro regolamento europeo "e-privacy", che dovrebbe essere varato presto, prevede che l'interessato possa rifiutare "a priori", per default e senza eccezioni, di essere controllato e tracciato. In sostanza di non cedere a chicchessia i propri dati personali attraverso il browser. Ora il problema è questo: visto che molti servizi (Gmail, Facebook, per fare solo due esempi) sono resi in cambio

dei dati personali, il rifiuto “a priori” significa che chi nega il consenso non potrà più usare un servizio che presuppone lo scambio tra l’accesso e i dati personali? Non è un caso che a Bruxelles nel corso della discussione del testo, su questo un punto abbiano puntato le loro armi più forti le lobby degli OTT.

In Italia il Garante – frenato anche da una cronica insufficienza dell’organico – e il legislatore non hanno ancora fatto chiarezza sui molti interrogativi del passaggio dal regime del Codice del 2003 al nuovo regolamento.

In questo quadro non è facile immaginare strumenti pubblici efficaci per la protezione della vita privata. Ma, tanto per incominciare, si potrebbero fare pulizie sui siti della pubblica amministrazione e sanzionare le violazioni più clamorose, come le informative reticenti o menzognere, che appaiono all’apertura di quasi tutti i siti, pubblici e privati.
Poi si dovrebbero realizzare campagne di informazione su larga scala, per rendere i cittadini consapevoli dei rischi che corrono ogni volta che accendono un PC o usano il telefonino. E suggerire le soluzioni più elementari per una prima auto-protezione.

Ma è realistico immaginare che azioni di questo tipo possano essere intraprese da un governo che ha tra i suoi principali consulenti il numero due di una delle “sei sorelle”, uno dei più potenti padroni dei Big Data su scala globale?