Privacy e fruizione della conoscenza scientifica

Versione 1.0 – marzo 2009

Paolo Guarda

Privacy e fruizione della conoscenza scientifica

Versione 1.0 marzo 2009

Paolo Guarda[*]

1. Premessa: la ricerca dell'onniscienza e l'occhio di Odino.....2
2. Proprietà intellettuale, Digital Rights Management (DRM) e privacy 3
3. Servizi «gratuiti» di accesso alle informazioni on-line e privacy 8
4. Conclusioni: verso il superamento della contrapposizione privacy v. conoscenza.....12

1. Premessa: la ricerca dell'onniscienza e l'occhio di Odino

La conoscenza rappresenta da sempre il sogno irraggiungibile dell'uomo e l'onniscienza, ovvero la conoscenza totale ed illimitata, viene considerata nella maggior parte delle culture una caratteristica divina. Sapere significa potere: questo nella storia dell'uomo – oltre che nella riflessione filosofica recente - è un punto fermo!

Ma cosa siamo disposti a cedere pur di avere accesso alla conoscenza? I miti e le storie sulla ricerca della conoscenza narrano di lunghi viaggi e di estremi sacrifici. Ed è per questo che, ripescando tra vecchie passioni, vorrei cominciare quest'intervento proponendovi il racconto di una saga nordica il cui significato spero di riuscire a rendervi chiaro nelle conclusioni[1].

Odino, il padre degli Dei nella tradizione scandinava, viene spesso raffigurato con un occhio solo. Una delle tante varianti delle saghe nordiche racconta che, per soddisfare la propria sete di conoscenza, il dio abbia sacrificato un suo occhio a Mimir, il guardiano di Mimirsbrønd, la magica fonte di ogni sapere posta nei pressi di una delle radici di Yggrdrasil[2]. L'occhio divino da allora si trova nelle acque gelide della sorgente, quale prezzo pagato per acquisire lo sguardo del saggio e per scorgere, dietro le apparenze, l'essenza delle cose[3].

---

[*] In corso di pubblicazione in R. CASO (a cura di), *Pubblicazioni scientifiche, diritti d'autore ed open access*, Atti del Convegno tenuto presso la Facoltà di Giurisprudenza di Trento il 20 giugno 2008. Questa versione 1.0 – marzo 2009 © 2009 by Paolo Guarda – è pubblicata con Creative Commons Attribuzione-Non commerciale-Non opere derivate 2.5 Italia License. Tale licenza consente l'uso non commerciale dell'opera, a condizione che ne sia sempre data attribuzione all'autore. Maggiori informazioni all'URL: <http://creativecommons.org/licenses/by-nc-nd/2.5/it/>

[1] Per approfondimenti sulle saghe nordiche v. V. GRØNBECH, *Miti e leggende del nord*, Torino, 1996 (traduzione integrale dell'edizione danese *Nordiska Mita og Sagn*, Copenhagen, 1965).

[2] Vedi *Völuspá* (*La profezia della veggente*), il primo e più famoso poema dell'*Edda poetica*.

[3] Nelle saghe si racconta di altri sacrifici di Odino che testimoniano la sua ansia di conoscenza e di comprensione dei segreti del mistero. Il padre degli Dei «sacrifica se stesso» ad un ramo dell'albero del mondo, di nuovo lo Yggrdrasil, e resta appeso per nove giorni e nove notti penzolando privo di conoscenza. Non soddisfatto, con una lancia si infligge un'altra tortura e si ferisce (la c.d. «ferita di Odino», il segno inciso nella viva carne che molti guerrieri si procuravano nel tentativo di ingraziarsi il dio). La mente del dio, mentre il suo corpo pativa immani sofferenze, era libera di vagare alla ricerca di

Questo è il prezzo che ha pagato Odino. Ma qual è quello che siamo disposti a pagare noi? O meglio: siamo veramente consapevoli di quale costo abbia il nostro accesso alla conoscenza, soprattutto nel contesto digitale e nonostante l'apparente gratuità del servizio offerto? Cercherò di dare risposta a questi interrogativi nelle pagine che seguono.

L'intervento avrà questa struttura: nel secondo paragrafo si tratterà dei problemi legati all'utilizzo dei sistemi di Digital Rights Management (DRM), atti a governare la fruizione di opere dell'ingegno regolate da proprietà intellettuale, con riferimento alla privacy dei fruitori nelle sue diverse dimensioni; si delineeranno, poi, nel terzo paragrafo gli aspetti critici che il libero accesso alla conoscenza (c.d. Open Access, di qui in avanti O.A.) presenta, sempre dalla prospettiva della privacy: si studierà un caso di servizio offerto gratuitamente che presenta enormi rischi dal punto di vista del profiling degli utenti. In chiusura, tirando le fila del discorso e così riprendendo gli spunti offerti dalla saga di Odino, si svolgeranno alcune considerazioni conclusive sul rapporto tra privacy e conoscenza nel contesto digitale.

2. Proprietà intellettuale, Digital Rights Management (DRM) e privacy

Non intendo ripercorrere la mappa sapientemente disegnata nelle relazioni che mi hanno preceduto. E, quindi, non dirò nulla su quella contrapposizione che è stata delineata nell'accesso alla conoscenza tra schemi che seguono il paradigma della proprietà intellettuale (copyright, licenze, ecc.) e le norme informali per il controllo della conoscenza. Né affronterò l'analisi di quella che il prof. Caso ha definito come una nuova speranza: l'O.A.[4].

In questo frangente mi concentrerò, invece, su quali siano le ricadute in termini di privacy che le diverse modalità di fruizione delle opere dell'ingegno hanno[5]. L'avvento delle tecnologie digitali ha, infatti, prodotto, da un lato, un'esasperazione del controllo rigido ed accentrato (come quelle basato su DRM, di cui avremo modo di trattare), basato su regole commerciali, dall'altro l'avvento di nuove modalità di fruizione che si fondano su regole di libero accesso ai contenuti (*rectius*, O.A.).

---

nuovi orizzonti: così Odino vide le rune, le iscrizioni dotate di particolari poteri magico-divinatori, e se ne appropriò, raccogliendole da terra.

[4] Per approfondimenti sulla tematica dell'Open Access, ampiamente trattata nel corso del convegno, v. R. CASO, *Proprietà intellettuale, tecnologie digitali ed accesso alla conoscenza scientifica: Digital Rights Management vs. Open Access*, reperibile all'URL: <www.jus.unitn.it/users/caso/PIACS/Libro/access/home.asp>; ID. (a cura di), *Ricerca scientifica pubblica, trasferimento tecnologico e proprietà intellettuale*, Bologna, 2005, in particolare il contributo di R. CASO, *La commercializzazione della ricerca scientifica pubblica: regole e incentivi*, 9; D. L. BURK, *Intellectual Property in the Context of E-Science*, (August 18, 2006) Minnesota Legal Studies Research Paper No. 06-47, reperibile all'URL: <http://ssrn.com/abstract=929479>.

[5] Nell'era di Internet, privacy e copyright conoscono spesso momenti di contrasto. Vedi, ad esempio, il caso Peppermint commentato in R. CASO, *Il conflitto tra copyright e privacy nelle reti peer to peer: in margine al caso Peppermint. Profili di diritto comparato*, in *Dir. dell'Internet*, 2007, 471 (reperibile anche in formato digitale all'URL: <http://www.jus.unitn.it/users/caso/DRM /Libro/peppermint/home.asp>).

Metterò, dunque, l'accento sul «prezzo» nascosto che l'utente paga per avere accesso alle opere intellettuali digitalizzate, sia nelle piattaforme che incorporano i valori e le regole proprie del controllo rigido ed accentrato, sia, e lo vedremo nel prossimo paragrafo, in quelle che sono ispirate a logiche di apertura e di libero accesso[6].

Su di un piano prettamente tecnologico, il modello di produzione dell'informazione basato sulla chiusura dell'informazione e sul controllo rigido ed accentrato della stessa viene implementato tramite sistemi di Digital Rights Management (DRM)[7]. Questo termine individua il più avanzato tipo di sistema di protezione anti-accesso ed anti-copia presente sul mercato. Il cuore di ogni DRM è costituito da due moduli. Da un lato, il c.d. «content-module», il quale contiene i dati digitalizzati, siano essi testo o file audio, che sono stati resi «sicuri» tramite un processo di criptazione e sono pronti per essere distribuiti; prima che ciò avvenga, però, nel contenuto viene incorporato il nome dell'autore, il titolare del diritto d'autore, la data di creazione, il titolo, il formato, la dimensione o altre informazioni tecniche per identificare il file (quale ad esempio l'ISBN). Dall'altro, un «licensing module», il quale genera la licenza digitale che automaticamente garantisce all'utente finale l'accesso al contenuto alla luce dei «diritti di utilizzo» (usage rights) di regole commerciali (business rules). Questi determinano sia il «cosa», cioè esattamente quale pezzo del contenuto può essere utilizzato, che il «quando», il che significa collegare particolari attributi ad ogni «diritto», quali ad esempio il tipo di utilizzatore che può esercitare il «diritto», l'estensione di ogni diritto (la durata o il numero delle volte) ed il prezzo per esercitare il diritto.

Facciamo un esempio per capire come questo meccanismo funziona[8]. Mettiamo il caso che io voglia avere accesso alle nuove pubblicazioni di una rivista giuridica on-line. Mi registro conferendo i dettagli del mio profilo al portale che contiene la rivista; i miei dati vengono salvati in un database adibito ad immagazzinare le informazioni relative all'identità degli utenti all'interno del licesing module. Nello stesso momento il sistema genera le regole di utilizzo che definiscono il tipo di uso, il costo e le categorie di utenti abilitate e le inserisce all'interno del license module. Ogni qual volta io desideri leggere l'ultima pubblicazione, il sistema di lettura del documento (un software che ho installato nella mia macchina al momento della registrazione) contatta il content module e una copia decriptata mi viene spedita. A questo punto posso visualizzare ciò che cercavo, ma ogni accesso futuro mi sarà proibito. Per avere altri accessi, dovrò sottoscrivere un abbonamento che mi garantisca più visualizzazioni, o maggiori margini di utilizzo. Questo un semplice esempio di come funziona un sistema di DRM.

---

[6] Un saggio in generale sulle nuove frontiere della monitoraggio anche per quanto concerne le opere protette da proprietà intellettuale, v. S.K. KATYAL, *The New Surveillance*, 54 *Cas Western L. Rev*. 297 (2004).

[7] Per approfondimenti v. l'opera di R. CASO, *Digital Rights Management. Il commercio delle informazioni digitali tra contratto e diritto d'autore*, Padova, 2004 (ristampa digitale, Trento, 2006, reperibile all'URL: <http://www.jus.unitn.it/users/caso/pubblicazioni/drm/homeDRM.asp?cod=roberto.caso>).

[8] L'esempio, modificato alla luce delle considerazioni che stiamo svolgendo, è ripreso da P. GANLEY, *Access to the Individual: Digital Rights Management Systems and the Intersection of Informational and Decisional Privacy Interests*, 10 *International Journal of Law and Information Technology* 241 (2002).

Il DRM presenta anche funzionalità che hanno una diretta ricaduta sulla privacy del soggetto che accede alle informazioni[9]. Esso contiene, infatti, almeno una delle seguenti funzioni di base[10]:

a. controllo sull'accesso al contenuto;
b. controllo sugli usi del contenuto;
c. identificazione del contenuto, dei titolari del contenuto e delle condizioni generali per l'utilizzo del contenuto;
d. autenticazione dei dati di identificazione.

Inoltre, per sue specifiche caratteristiche, un sistema di DRM è comunque in grado di monitorare la fruizione del contenuto e, qualora programmato in tal senso, è in grado di «sanzionare» i comportamenti che non rispettano le sue regole, ad esempio disattivando l'accesso al medesimo contenuto.

Scorrendo, anche velocemente, quello che può fare un DRM, per uno studioso che approfondisce tematiche relative alla privacy, subito un termine riecheggia nella mente: profiling[11]. I dati personali dei vari utenti di Internet, dei fruitori delle opere dell'ingegno e più in generale i loro interessi commerciali rappresentano un autentico tesoro per chi opera sulla Rete. La possibilità di monitorare le attività dell'utente non rappresenta solo il modo di controllare-gestire la sua fruizione del contenuto digitale al fine di imporre, e se del caso sanzionare, le attività che gli sono consentite in base alla licenza che regola gli usage rights, bensì soprattutto diventa a sua volta un modo per commercializzare il profilo dell'utente stesso. Questo nuovo «bene» può servire al profiler per porre in essere una strategia di marketing più mirata e, quindi, efficace nei confronti del suo cliente (magari attraverso dei banner pubblicitari rispondenti ai gusti dell'utente), ma può diventare anche merce di scambio allorquando si intenda vendere le informazioni raccolte a terzi.

Questo tipo di considerazioni fa comprendere facilmente quale sia l'impatto che tecnologie come quelle oggetto di analisi abbiano nei confronti della privacy degli individui. Ma chiariamo subito cosa intendiamo per privacy. Quando si affronta la descrizione di questo concetto, soprattutto dal punto di vista del rapporto tra diritto e tecnologia, si racconta come da una privacy intesa come «right to be let alone» derivante dal celeberrimo articolo di Warren e Brendeis e dalle prime sentenze italiane della seconda meta del XX secolo, si sia arrivati, grazie alla diffusione dei computer e

[9] Sul rapporto tra i sistemi DRM e la privacy v. CASO, *Digital Rights Management. Il commercio delle informazioni digitali tra contratto e diritto d'autore*, cit., 98 ss.; A. PALMIERI, *DRM e disciplina europea della protezione dei dati personali*, in R. CASO (a cura di), *Digital rights management: problemi teorici e prospettive applicative.* Atti del convegno tenuto presso la Facoltà di giurisprudenza di Trento il 21 ed il 22 marzo 2007, Trento, 2007, 197 ss.; M.A. EINHORN, *Digital Rights Management, Licensing, and Privacy*, Ottobre 2002, reperibile all'URL: <http://ssrn.com/abstract=332720>; J.E COHEN, *Overcome Property: (Does Copyright Trump Privacy?)*, *University of Illinois Journal of Law, Technology & Policy* 375 (2002).

[10] V. CASO, *Digital Rights Management. Il commercio delle informazioni digitali tra contratto e diritto d'autore*, cit., 100.

[11] Sulle tematiche relative all'attività di profiling v. B. CUSTER, *The Power of Knowledge. Ethical, Legal, and Technological Aspects of Data Mining and Group Profiling Epidemiology*, Nijmegen, The Netherlands, 2004.

quindi delle tecnologie digitali, ad un significato più legato al controllo dei propri dati personali. E, infatti, ogni volta che diciamo «privacy» riferendoci alla normativa che la regolamenta subito ci affrettiamo a ricordare come in realtà si stia parlando di dati personali, di controllo dei dati, di consenso, informative, ecc.

Ma qui vorrei richiamare un concetto più ampio di privacy, il quale riprende gli studi di diversi giuristi d'oltreoceano. Questo tipo di descrizione del concetto, che a breve andremo ad analizzare, permette di comprendere meglio quale sia la reale importanza della privacy nella Rete, quale sia il suo vero ruolo. In un contesto in cui sempre di più sono le informazioni, ed il commercio che si fa su di esse, a balzare alla ribalta. La stessa personalità dei soggetti viene scomposta in tanti piccoli frammenti, che sono i dati personali. Questi frammenti, quando ricomposti, delineano, poi, nel contesto virtuale il profilo di una persona, la sua identità virtuale. Il controllo su questi frammenti di identità ha una ricaduta ben più pericolosa di quanto si è portati a ritenere se si dà al concetto di privacy una lettura ristretta, in quanto coinvolge gli aspetti fondamentali dell'esistenza stessa del soggetto cui tali dati sono riferiti.

Vediamo, allora, quali sono queste ricostruzioni dogmatiche a cui ho fatto riferimento. Il consumo di opere intellettuali, inteso come fruizione di opere dell'ingegno e di altre informazioni, è direttamente collegato a tre fondamentali dimensioni del concetto di privacy: «spaziale», «informazionale» e, ultima ma più importante, «decisionale»[12].

La prima dimensione, quella «spaziale», riguarda lo spazio fisico, ed in particolare l'estensione di territorio all'interno del quale la solitudine di un individuo è difesa dall'invasione esterna: questo tipo di privacy corrisponde a quel concetto che richiamano i sociologi quando parlano di spazio privato contro spazio pubblico e di situazioni di sovraffollamento. La dimensione in oggetto richiama direttamente la libertà dal non essere disturbati dalle immissioni o, più in generale, da fonti di disturbo provenienti dall'esterno. Nel contesto in cui ci troviamo a trattare oggi, si riferisce a quegli spazi all'interno dei quali l'individuo è libero di tenere comportamenti che risultano anche aberranti per le norme sociali dominanti o che, più semplicemente, non sono destinati ad esser tenuti in pubblico.

La seconda dimensione, la privacy «informazionale», riguarda il flusso delle informazioni personali. Più precisamente, essa concerne il controllo dell'individuo nei confronti del trattamento dei propri dati personali. È, dunque, strettamente collegata a quel modo di concepire la privacy come controllo sui propri dati, come regolamentazione del trattamento degli stessi. Dalla nostra visuale, cioè studiando l'impatto di sistemi DRM, i dati «processati» sono quelli relativi al consumo

[12] V. J. KANG, *Information Privacy in Cyberspace Transactions*, 50 *Stan. L. Rev*. 1193 (1998), 1202-1205; GANLEY, *Access to the Individual: Digital Rights Management Systems and the Intersection of Informational and Decisional Privacy Interests*, cit.. Il prof. Caso riprende, invece, da un'autrice statunitense una doppia dimensione del concetto di privacy (informazionale e spaziale): CASO, *Digital Rights Managemen. Il commercio delle informazioni digitali tra contratto e diritto d'autoret*, cit., 103 ss.; J. COHEN, *DRM and Privacy*, 18 *Berkeley Tech. L.J*. 575 (2003). Più in generale sulle implicazioni dei DRM per quanto concerne la privacy e la libertà di espressione, v. I. KERR, J. BAILEY, *The Implications of Digital Rights Management for Privacy and Freedom of Expression*, 2 *Info., Comm. & Ethics in Society* 87 (2004).

intellettuale, informazioni acquisite secondo le modalità che abbiamo visto caratterizzare le misure tecnologiche di protezione.

La terza, e sicuramente più importante, dimensione della privacy è quella «decisionale»: essa ha a che fare con la scelta, con la libertà che deve essere riconosciuta all'individuo di poter prendere decisioni che lo riguardano senza condizionamenti esterni. È su questa dimensione che desidero spendere qualche parola in più.

La privacy «decisionale» riguarda l'essenza stessa dell'uomo. Il libero arbitrio, la libertà di autodeterminazione, in sostanza la libertà di poter essere uomini. Non c'è diritto se non vi è libero arbitrio, se non vi è la possibilità di scegliere, anche di sbagliare se del caso. La violazione di questa dimensione della privacy, quindi, incide direttamente sulla «capacità» di un soggetto di essere uomo. Il monitoraggio, il sapere di essere controllati, la consapevolezza che il contesto intorno a noi si modifica alla luce del profilo che altri ci stanno ritagliando addosso modificano il nostro comportamento. L'idea stessa del Panopticon di Bentham si basava su questo[13]: la forma radiocentrica dell'edificio e gli opportuni accorgimenti architettonici e tecnologici, che permettevano ad un unico guardiano di poter osservare tutti i prigionieri in ogni momento, senza che questi fossero in grado di stabilire di essere osservati o meno (la qual cosa profetizza tremendamente il monitoraggio sulla Rete) avrebbe determinato nei detenuti la percezione di un'invisibile onniscienza, condizionando il loro comportamento ed inducendoli a non violare le regole[14].

Il solo fatto di sapere di essere osservati, dunque, condiziona, e condizionerà, il nostro consumo intellettuale, secondo modalità probabilmente non ancora chiare nella loro tragica portata.

Torniamo all'esempio da cui eravamo partiti per valutare le potenziali implicazioni dei sistemi di DRM con riguardo alla privacy.

Prima che i contenuti pubblicati sulla rivista giuridica on-line siano criptati, all'interno di ogni sezione di tutti gli articoli viene inserito un metadato chiamato «Digital Object Identifier» (DOI). Il livello di «granularità» (cioè di dettaglio) di questi DOI (ogni articolo, ogni sezione, o ogni paragrafo), che sono unici per ogni tipo di contenuto, è deciso dall'editore. Il risultato di questo processo è il seguente: tutti i dati criptati all'interno del content module sono direttamente legati a «prezzi» di metadati identificatori, che rimangono sempre legati al contenuto allorquando questo viene, poi, distribuito. Ogni qual volta io, allora, voglia modificare il contenuto del documento di cui ho chiesto l'accesso al sistema (ad esempio aprirlo, vederlo, stampare una determinata sezione) l'applicazione chiamata «ContentControl» spedisce al licensing module un pacchetto di informazioni, comprensivo del DOI e dei dettagli dell'azione da

[13] Lo stesso Bentham lo descrisse come *«un nuovo modo per ottenere potere mentale sulla mente, in maniera e quantità mai vista prima»*, in J. BENTHAM, *Panopticon*, 1787.

[14] Nel diritto ebraico la dottrina dello hezzek re'iyyah (danno causato dalla semplice vista) costituisce un punto di riferimento per la privacy: la dottrina si basa sull'idea secondo la quale non basta riconoscere protezione nei confronti di sguardi non voluti, ma anche contro la sola possibilità di essere visti. Ciò perché una persona, nell'incertezza di esser vista o meno, è comunque inibita e condizionata nelle sue attività: per approfondimenti, v. J. ROSEN, *The Unwanted Gaze. The Destruction of Privacy in America*, New York, 2001, 18-19.

me posta in essere, il quale verrà, poi, salvato, in un programma che registra le attività svolte («logging program»). Semplicemente accedendo al licensing module, l'editore può compilare delle statistiche aggregate dell'utilizzo dei contenuti o visualizzare il profilo di un determinato utente abbinando le informazioni contenute nel logging program con quelle salvate nella banca dati che, invece, raccoglie le informazioni personali degli utenti[15].

3. Servizi «gratuiti» di accesso alle informazioni on-line e privacy

Trattiamo ora dell'altra faccia della luna: il mondo dell'O.A.. Sarà preso ad esempio paradigmatico il motore di ricerca Google, anche nelle sue varianti di «Google Scholar» e «Google Ricerca Libri»[16]. Ciò perché tale strumento rappresenta uno tra i casi più invasivi presenti sulla Rete di raccolta di dati personali e di profiling degli utenti.

Anche quello che, da un punto di vista prettamente economico, potrebbe apparire un mondo ideale, ovvero quello dell'accesso alla conoscenza in modo del tutto libero e gratuito, in realtà nasconde un lato oscuro. Nonostante l'apparente idillio, anche in questo contesto, spesso inconsapevolmente, un prezzo, assai caro purtroppo, lo stiamo pagando.

Google rappresenta uno dei casi più eclatanti di raccolta di dati personali e, più in generale, di informazioni collegate agli utenti. Ogni giorno milioni di persone lo utilizzano conferendo informazioni sui loro interessi, bisogni, desideri, paure, ecc.

Cominciamo con una notazione: Google registra tutte le ricerche collegandole ad uno specifico indirizzo IP (Internet Protocol address). Questa affermazione probabilmente spaventa, lascia «di sasso», ci fa subito tornare con la mente alle ricerche che abbiamo fatto, timorosi di cosa «abbiamo chiesto a Google».

Ma procediamo con ordine. La privacy, da questa prospettiva, riguarda i dati personali, quelle informazioni che il nostro Codice per la protezione dei dati personali (d. lgs. 30 giugno 2003, n. 196) definisce come relative «a persona fisica, persona

[15] In chiusura, voglio solo ricordare come i sistemi di DRM, qui analizzati in un'ottica sicuramente negativa per quanto riguarda il loro impatto sulla privacy, possono avere anche utilizzi che diremo «privacy oriented». Si può infatti pensare ad una loro utilizzazione «positiva» secondo questa prospettiva: i dati personali che viaggiano sulla Rete potrebbero essere controllati attraverso sistemi di DRM, configurati questa volta dai titolari dei dati stessi, che consentono una corretta gestione dei flussi di informazioni, che sia conferme alle regole in materia di protezione dei dati personali (tale spunto viene ripreso dall'intervento dell'ingegnere Leonardo Chiariglione al Convegno «Digital Rights Management. Problemi teorici e prospettive applicative», tenutosi presso la Facoltà di Giurisprudenza dell'Università degli Studi di Trento il 21-22 marzo 2007).

[16] Per approfondimenti v. O. TENE, *What Google Knows: Privacy and Internet Search Engines*, (Ottobre 2007), reperibile all'URL: <http://ssrn.com/abstract=1021490>. Per un saggio di denuncia dell'attività di profilazione svolta dai motori di ricerca che propone di regolamentare il settore, v. K. LAUDADIO DEVINE, *Searching for Privacy Online: Legislating Against the Retention of Search Histories*, Marzo 2007, reperibile all'URL: <http://ssrn.com/abstract=1111378>; sulla stessa linea, v. lo studio nel contesto giuridico statunitense con riferimento all'applicabilità del Primo Emendamento di F.A. PASQUALE III, O. BRACHA, *Federal Search Commission?: Access, Fairness and Accountability in the Law of Search*, Public Law and Legal Theory Resarch Paper No. 123, Luglio 2007, reperibile all'URL: <http://ssrn.com/abstract=1002453>.

giuridica, ente od associazione, identificati o identificabili, anche indirettamente, mediante riferimento a qualsiasi altra informazione, ivi compreso un numero di identificazione personale» (art. 4, co. 1, lett. b). Quindi, a contrario, le informazioni che non sono ricollegabili ad una determinata persona non rappresentano un problema per la privacy. Il problema, dunque, consiste nella possibilità che questo collegamento venga stabilito.

Di seguito vediamo come i file di log relativi alle nostre ricerche on-line possono essere associati alla nostra persona.

Innanzitutto attraverso la raccolta degli indirizzi IP[17]. Questi sono numeri che identificano univocamente nell'ambito di una singola rete i dispositivi collegati; possono essere statici, quindi fissi, o, più spesso, dinamici, quindi assegnati di volta in volta dall'Internet Service Provider (ISP) ai propri abbonati a seconda della necessità. L'indirizzo IP, che può essere banalmente considerato come l'equivalente di un indirizzo stradale o di un numero di telefono, è considerato un dato personale ai sensi dell'art. 4 Codice Privacy sopra richiamato[18]. Anche se Google non è in grado di associare un indirizzo IP, a cui collega nel file di log la richiesta posta in essere, ad un determinato utente, il fatto che l'ISP abbia tale capacità e che l'autorità statale possa forzarlo a comunicare i dati relativi ai propri abbonati, rende anche i file di log dei dati personali. È, infatti, la possibilità in potenza di poterli collegare a renderli tali, non il fatto che effettivamente ciò avvenga[19].

In secondo luogo, per superare le difficoltà nel profilare gli utenti che l'utilizzo di indirizzi dinamici presenta, un portale che offre servizi liberamente accessibili può utilizzare dei file «cookie», i quali contrassegnano il browser (il software utilizzato per la navigazione in Internet) dell'utilizzatore con numeri univoci di identificazione (unique identifying number)[20]. Questi file, che apparentemente dovrebbero solamente semplificare l'attività dell'utente ricordando i dati relativi al suo login ed alle sue ricerche ad ogni accesso alla pagine Web, in realtà permettono al motore di ricerca di riconoscere l'utente come un «visitatore ricorrente» del sito e di collazionare lo storico

[17] Nelle sue privacy policy Google dichiara che: «Quando visitate il nostro sito web, i server di Google registrano automaticamente informazioni quali l'URL, gli indirizzi IP, il tipo di browser, il linguaggio del browser, la data e l'ora della vostra richiesta», v. <www.google.com/privacy.html>.

[18] V. il Parere 2/2002 del Gruppo europeo dei garanti per la tutela dei dati personali (ex art. 29 Dir. 95/46 CE), sull'uso di identificativi esclusivi negli apparecchi terminali di telecomunicazione: l'esempio dell'IPv6, 10750/02/IT/def.WP 58, nel quale si legge a pag. 3: «[i]l gruppo mette in evidenza che gli indirizzi IP attribuiti agli utenti Internet costituiscono dati personali».

[19] Una soluzione tecnica a questo problema è offerta dall'applicazione chiamata «TrackMeNot», un'estensione per Firefox, creata da Daniel C. Howe e Helen Nissenbaum, la quale lavora in background quando il nostro browser è attivo ed invia periodicamente ricerche casuali: in questa maniera il motore di ricerca riceverà le nostre richieste mescolate a molte altre e non sarà in grado di comprendere con esattezza quali siano i reali argomenti oggetto di ricerca. V. <http://www.mrl.nyu.edu/~dhowe/trackmenot/>.

[20] Sempre nelle privacy policy di Google si legge: "Google si serve di cookie ed altre tecnologie per ampliare la sua esperienza online e per capire come utilizzate i servizi di Google allo scopo di migliorarne la qualità". Per approfondimenti, v. J.J. THILL, *The Cookie Monster: From Sesame Street to Your Hard Drive*, 52 *S.C. L. Rev*. 921 (2001).

delle sue ricerche, anche se si connette attraverso differenti indirizzi IP[21]. Il punto debole di un sistema di profiling che si basa sull'utilizzo dei cookie risiede nel fatto che esso è accessibile solamente dal server che lo ha scaricato sulla nostra macchina. Quindi i cookie scaricati da Google saranno leggibili solo da Google, e non anche da Yahoo o Wikipedia. Questa debolezza è stata recentemente superata da «DoubleClick», una società che sviluppa e fornisce servizi Internet[22]. Attraverso la tecnologia da essa stessa ideata, il c.d. DART (Dynamic, Advertising, Reporting, and Targeting), essa è in grado di impostare i suoi cookie in modo tale da monitorare gli utenti come si spostano di sito in sito e registrare quali pubblicità commerciali vedono e selezionano durante la navigazione[23].

I cookie di per se stessi non possono considerarsi dati personali in quanto identificano uno specifico browser (*rectius*, un computer) piuttosto che un utente. Si immagini, infatti, il cookie come una sorta di etichetta, del tipo «7593ci3209lc12», senza un'apparente ricaduta sulla privacy, ma per così dire appiccicata ad una scatola di informazioni di un'anonima persona. Se si potesse, però, associare un determinato cookie collegato al file di log contenente le ricerche svolte sul motore ad una persona specifica, allora di nuovo si ricadrebbe sotto lo scomodo, per coloro che svolgono attività di profiling, cappello dei dati personali. Ma anche questo ostacolo può essere aggirato; vediamo come. Come tutti ben sappiamo, oltre al motore di ricerca, Google fornisce anche una serie di servizi aggiuntivi (dalla casella postale al portafoglio finanziario) i quali richiedono una previa registrazione tramite credenziali, come il nome reale e l'indirizzo e-mail (l'esempio più ovvio è «Gmail»). A questo punto il gioco è fatto: l'utente si connette ad Internet, carica il portale del motore di ricerca, accede alla sua casella postale identificandosi con i propri dati personali, e magari nel frattempo pone in essere una ricerca. Google utilizza lo stesso cookie con riferimento allo storico delle ricerche dell'utente e per identificarlo quando si connette all'account di posta: l'anonimato del cookie è così definitivamente perso, il nome mancante dei file di log contenenti gli interessi manifestati dall'utente nella sua navigazione hanno ora un nome e un cognome!

Infine, un'ultima possibilità di raccolta di dati personali residua; questa volta il vero colpevole è l'utente stesso, a prescindere da qualsiasi tipo di accorgimento egli

[21] Google ha recentemente annunciato di avere l'intenzione di ridurre la durata dei suoi cookie, che era stata inizialmente prevista sino al 2038, a «soli» due anni dall'ultima ricerca dell'utente sul suo motore di ricerca: v. *Cookies Expiring Sooner to Improbe Privacy*, Official Google Blog, !6 luglio 2007, reperibile all'URL: <http://googleblog.blogspot.com/2007/07/cookies-expiring-sooner-to-improve.html>. Se pensiamo, però, al fatto che ognuno di noi lo utilizza quasi giornalmente, ci rendiamo subito conto di quanto tale impegno abbia un impatto minimo, per non dire assente, sulla protezione della nostra privacy.

[22] V. il sito Web: <http://www.doubleclick.com/>.

[23] Per approfondimenti v. C. TROTTA, *The Google-DoubleClick Merger, the FTC, and the Future of Transactional Privacy Inquires in the United States*, Dicembre 2007, reperibile all'URL: <http://ssrn.com/abstract=1071823>. Per ovviare al problema della profilazione attraverso la raccolta dei cookie, si può installare «Scookies»: questa estensione per Firefox funziona per tutti i motori di ricerca e siti Internet che cercano di tracciare gli utenti tramite i cookie. Scookies cambia i cookie degli utenti mescolandoli gli uni agli altri. L'idea è la stessa del su citato «TrackMeNot»: alterando i coockie legati ai profili degli utenti non è più possibile associare la ricerca posta in essere ad uno specifico soggetto. V. <http://www.autistici.org/bakunin/scookies/>.

possa aver posto in essere per limitare l'uso e l'identificabilità dei suoi cookie. Tutti noi abbiamo provato nella nostra vita a cercare il proprio nome su Google o, nel nostro caso, a cercare i nostri contributi sulla banca dati di contenuti scientifici, magari più volte nell'arco della stessa giornata. Queste ricerche vengono definite «ego searches» o «vanity searches»[24]. Esse permettono ai profiler di ricollegare le informazioni di cui abbiamo sinora dato descrizione ad una determinata identità.

Ora poniamoci un'altra domanda: che fine fanno i file di log che contengono tutte queste informazioni relative alle nostre ricerche?

Sulla privacy policy di Google candidamente viene affermato: «Usiamo i cookie per migliorare la qualità dei nostri servizi memorizzando le preferenze dell'utente e mantenendo una traccia delle sue abitudini, come quella di memorizzare che cosa cerca la gente. La maggior parte dei browser è impostata per accettare i cookie, ma potete resettare il vostro browser di modo che questo rifiuti tutti i cookie o segnali quando vi stanno mandando un cookie. Tuttavia, è possibile che alcune funzioni ed alcuni servizi di Google non funzionino correttamente se avete disabilitato la funzione che accetta i cookie»[25]. Quindi la ragione declamata consiste nella volontà di migliorare la qualità del servizio! Peccato però che nel fare questo Google stia dando vita al più vasto e mai visto nella storia dell'uomo «Database delle intenzioni» (Database of Intentions): «the aggregate results of every search ever entered, every result list ever tendered, and every path taken as a result»[26]. Gli incentivi a mantenere tale enorme quantità di informazioni sono rappresentati da: i relativamente bassi costi di raccolta e mantenimento dei dati, la relativa mancanza di normativa sul tema (questo maggiormente vero per quanto riguarda il sistema statunitense) e un potenziale utilizzo assai remunerativo delle informazioni.

Il database delle intenzioni rappresenta anche un bene estremamente prezioso, una sorta di miniera d'oro se consideriamo il valore economico delle informazioni contenute, le quali fanno sicuramente gola a molti: dagli organismi statali adibiti alla sicurezza nazionale ed all'applicazione della legge fino ad arrivare agli hacker ed ai «ladri d'identità». Al momento, i gestori dei motori di ricerca non possono (o non potrebbero) vendere i dati personali che posseggono a soggetti terzi. Aldilà del fatto che ciò non significa che questo non sarà possibile in futuro, essi comunque scambiano i dati degli utenti con società affiliate o altri partner commerciali «fidati» al fine di trattare queste informazioni e fornire servizi[27].

---

[24] V. C. SOGHOIAN, *The Problem of Anonymous Vanity Searches*, Winter 2007, reperibile all'URL: <http://ssrn.com/abstract=953673>.

[25] Vedi le *Norme sulla privacy di Google*, reperibili all'URL: <http://www.google.com/intl/it/privacypolicy.html>.

[26] Termine ripreso da J. BATTELLE, *The Database of Intentions*, John Battelle's Searchblog, 13 Novembre 2003, reperibile all'URL: <http://batellemedia.com>; cfr. anche ID., *The Search: How Google and its Rivals Rewrote the Rules of Business and Transformed our Culture*, Boston, MA - London, 2005.

[27] Sempre sulle privacy policy di Google si legge: «Offriamo alcuni servizi che sono collegati con altri siti web. I dati personali che fornite a tali siti potrebbero essere inviati a Google per poter erogare i servizi. Trattiamo tali informazioni conformemente a quanto previsto dalle presenti Norme. I siti collegati possono avere prassi sulla privacy differenti e vi invitiamo quindi a leggere le loro norme a tale proposito».

Comunque sia, c'è sicuramente almeno una parte terza rispetto al trattamento che potrà provare ad ottenere la comunicazione dei dati personali dai motori di ricerca, e questo attraverso un procedimento del tutto legittimo. Ci riferiamo ai poteri statali i quali possono avere l'interesse ad utilizzare i file di log per finalità di pubblica sicurezza. Soprattutto dopo i tragici eventi dell'11 settembre, si è rafforzata una tendenza già in atto che porta a modificare il modo in cui le persone considerano lo Stato e le sue responsabilità, rimarcando il suo tradizionale ruolo di custode della sicurezza pubblica. Sempre di più lo Stato, quindi, che inizialmente si era defilato nella regolamentazione di Internet, ritorna a giocare un ruolo da protagonista stringendo un'alleanza strategica con gli ISP e con i fornitori di servizi su Internet, tutti soggetti che possiedono informazioni preziosissime per l'attività di prevenzione e contrasto alla criminalità. Si parla allora di «Invisible Handshake»[28].

Abbiamo portato l'esempio di Google però le stesse considerazioni potrebbero svolgersi per i portali specializzati nella ricerca di materiale scientifico che offrono servizi di questo tipo. A paradigmatico esempio si può richiamare il celebre portale Social Science Research Network (<http://ssrn.com>).

4. Conclusioni: verso il superamento della contrapposizione privacy v. conoscenza

Alla fine del percorso che, seppur al prezzo di alcune scorciatoie, abbiamo compiuto, tiriamo le fila di quanto detto cercando di svelare il perché la saga di Odino descritta all'inizio abbia a che fare col nostro argomento.

Che si tratti della fruizione della conoscenza scientifica secondo gli schemi proprietari trasposti nell'ambiente digitale dai sistemi di DRM o dell'accesso ai contenuti libero su quella linea di tendenza che porta verso l'O. A., noi stiamo perdendo (scambiando, se ne avessimo la consapevolezza) la nostra identità, il nostro profilo, i nostri dati in cambio delle opere dell'ingegno di cui intendiamo usufruire. Come per quanto successo ad Odino, stiamo cedendo un occhio, cioè una parte importantissima del nostro corpo, a Mimir, ovvero a coloro che posseggono quella conoscenza a cui tanto aspiriamo. Rimanendo sullo stesso esempio mitologico, vengono alla mente altre interessanti, e al tempo stesso terrificanti, similitudini. Per poter bere dalla fonte magica un «sorso di conoscenza» cediamo l'occhio, e facendolo è come se dessimo ad altri la possibilità di vederci dentro attraverso quell'occhio ceduto, di conoscere ciò che pensiamo, ciò che desideriamo, in poche parole di impossessarsi della nostra parte più nascosta e al tempo stesso più preziosa: i nostri pensieri.

La privacy rappresenta anche, e soprattutto, un fenomeno culturale: tutti i soggetti devono acquisire una maggiore consapevolezza dei rischi collegati al trattamento illecito dei dati che li riguardano e diventare essi stessi le prime «misure di sicurezza» atte ad evitarne indebiti utilizzi. Fino a che gli utenti di Internet non comprenderanno i rischi collegati al trattamento dei loro dati ed il reale «prezzo» che

[28] V. M.D. BIRNHACK, N. ELKIN-KOREN, *The Invisible Handshake: The Reemergence of the State in the Digital Environment*, 8 *Va. J.L. & Tech.* 6 (2003).

pagano nell'accesso ai vari servizi che la Rete offre, essi saranno sempre terribilmente esposti al controllo di quanti, invece, di questa consapevolezza hanno fatto commercio.

Conoscenza scientifica v. privacy verrebbe da dire, quindi. Però, se veramente si vuole risolvere questo conflitto, occorre utilizzare un approccio non tradizionale.

La privacy (almeno nella sua accezione informazionale) non consiste solo nel creare paletti alla circolazione dei dati; essa deve declinarsi, invece, nella corretta gestione, nel giusto management dei flussi di dati che contraddistingue in maniera ineliminabile l'era delle informazioni. La conoscenza, d'altro canto, si caratterizza dell'ormai famoso fenomeno della dematerializzazione che trasforma in libri, poesie, musiche, disegni, dei file espressi in codice binario. Non ha forse più senso concepire entrambi i termini del nostro discorso in maniera fisica, continuando in una contrapposizione che nel contesto digitale non esiste più. Meglio sarebbe cominciare ad interrogarsi sul fatto che essi possono essere considerati per quello che sono nel mondo digitale, cioè informazioni, e, dunque, ripensare regole, tecnologie e costumi alla luce di questa categoria uniformante che sposta la nostra attenzione sull'aspetto gestionale dei flussi che ne determinano lo scambio, piuttosto che sulla loro diversità intrinseca.