

ALESSANDRO MANTELERO

## **BIG DATA: I RISCHI DELLA CONCENTRAZIONE DEL POTERE INFORMATIVO DIGITALE E GLI STRUMENTI DI CONTROLLO**

**SOMMARIO:** 1. I c.d. *Big Data* nell'era dell'accesso. — 2. Il potere correlato ai *Big Data*. — 3. Gli strumenti utili a limitare tale potere: controllo, accesso, competizione.

### 1. I C.D. *BIG DATA* NELL'ERA DELL'ACCESSO.

Si ripete ormai da anni la considerazione secondo cui quella in cui viviamo è la società dell'informazione, tuttavia tale sintesi dell'attuale rilevanza riconosciuta in tutti gli ambiti ai processi di generazione, elaborazione e comunicazione dei dati informativi non costituisce un'acquisizione immutabile, un punto di arrivo definitivo. Al contrario il fenomeno informativo risulta variare nel tempo in ragione dei diversi parametri concernenti la sua distribuzione, dimensione e possibilità di un'analisi efficace dei contenuti.

Con riguardo al primo aspetto, va osservato come le informazioni non siano uniformemente distribuite sotto il profilo della fruibilità delle stesse.

Occorre in primo luogo distinguere fra accessibilità e capacità di sfruttamento dei dati; fattori che concorrono poi entrambi a concentrare il potere informativo in capo a gruppi di soggetti via via più circoscritti, sino a quelli che potremmo definire i « signori dei dati »<sup>1</sup>.

Sotto il profilo dell'accessibilità occorre rilevare come grandi quantità di dati siano disponibili solo in capo ad un numero limitato di soggetti, prevalentemente in virtù del ruolo che quest'ultimi rivestono nella generazione dei flussi informativi correlati. Così la pubblica amministrazione, come le grandi imprese private, in ragione delle proprie attività, quotidianamente raccolgono enormi quantità di dati che vanno poi a formare un

\* Il presente scritto è stato preventivamente sottoposto a referaggio anonimo affidato ad un componente il Comitato Scientifico dei Referenti della Rivista secondo le correnti prassi nella comunità dei giuristi.

<sup>1</sup> In seguito al venire in essere di *database* di dimensioni un tempo impensabili ed all'accresciuta capacità di analizzare con strumenti informatici grandi

quantità di dati, la tradizionale nozione di signoria dell'individuo sulle informazioni che lo riguardano pare indebolita in favore dei detentori delle risorse informative e di calcolo. Quest'ultimi, come si avrà modo di esporre *infra* nel testo, dispongono di un notevole potere informativo derivante dal controllo sulla gestione dei dati, tale da evocare la nozione di signoria.

patrimonio informativo in molti casi gestito in forma chiusa, senza concedere a terzi l'accesso all'intero insieme dei dati grezzi, ovvero limitandone l'accesso solamente ad alcune parti o ad alcuni soggetti, o infine rilasciando dati selezionati o rielaborati all'esterno senza tuttavia permettere di accedere ai dati grezzi in maniera integrale. La stessa posizione di forza si rinvia altresì in capo ad alcuni intermediari del flusso informativo, soggetti che, a differenza dei precedenti, non generano le informazioni, ma, svolgendo un ruolo chiave nella circolazione delle stesse, possono accedervi. Un esempio in tal senso possono essere gli internet *provider* o più in generale gli intermediari dei servizi dell'informazione, attraverso i quali passa un immenso flusso di dati non necessariamente generato dai medesimi, ma da questi monitorabile e fatto oggetto di analisi più o meno sofisticate. In analoga condizione si trovano, nel contesto *off-line*, le società che raccolgono presso terzi informazioni al fine di generare grandi basi di dati per erogare servizi di consulenza (si pensi ai servizi di affidabilità creditizia o a quelli di analisi di mercato). Nuovamente siamo in presenza di un numero circoscritto di operatori nelle cui mani si concentra un grande patrimonio informativo e che hanno la possibilità di scegliere chi possa accedervi, molto spesso in cambio di un'utilità economica.

Esistono poi casi in cui le informazioni sono invece accessibili, sia in forma grezza che elaborata. Si pensi agli *open data* (*data-set* resi disponibili al pubblico accesso da parte delle pubbliche amministrazioni), ovvero ai dati privati resi accessibili in virtù di un obbligo legale di pubblicità o ancora all'insieme degli studi, *report* e comunicazioni generati dai più disparati soggetti, alle opere di libera utilizzazione, sino agli *user-generated-content* che costituiscono parte assai rilevante dell'informazione disponibile *on-line*.

Il concorrente apporto di tutte queste disparate fonti rende solo apparentemente più fruibile l'informazione. In realtà la quantità di dati finisce per dar luogo ad un'overdose informativa che porta all'opposto risultato di indurre una diminuzione della conoscenza. Sussiste infatti un differenziale fra la ricezione o il reperimento delle informazioni e la generazione di conoscenza. Perché quest'ultimo processo avvenga, e perché dunque l'informazione assuma un valore, occorrono chiavi interpretative che consentano di discernere fra le informazioni, di riorganizzarle, di inserirle in un contesto sistematico ed atto a decodificare la complessità. Ove questo articolato processo non si realizzi, il risultato è quello della confusione e dell'attribuzione di valore a fonti scarsamente attendibili. Sono questi meccanismi ben noti al mondo professionale dell'informazione ed al giornalismo, che non a caso ha sviluppato un'attenta politica, anche giuridicamente rilevante, dell'uso delle fonti. La persona comune che attraverso una pluralità di canali multimediali accede oggi ad una quantità di informazioni inimmaginabile solo alcuni anni fa, molte volte non ha la stessa consapevolezza del professionista dei *media*; nel momento in cui dispone di un maggior numero di dati, questi finisce così per conseguire l'opposto risultato di incrementare la propria carenza informativa, attraverso false o incomplete informazioni, piuttosto che ridurla<sup>2</sup>.

<sup>2</sup> Cfr. G. DA EMPOLI, *Overdose. La società dell'informazione eccessiva*, Venezia, 2002, 52.

La situazione da ultimo descritta porta ad attribuire un maggior peso agli esperti, agli *opinion leader*, a coloro che offrono il proprio sapere per fungere da faro fra le nebbie dell'overdose informativa di ciascun settore, spesso micro-settore. A tal proposito non è un caso che i *social media*, luogo tipico dell'eccesso informativo, siano sempre più caratterizzati dall'attribuzione di un ruolo di *driver* ad alcune figure in cui l'utente ripone la propria fiducia (i *blogger*, gli utenti di Twitter con un maggior numero di *followers*, gli amici di Facebook). Da tale tendenza deriva tuttavia un'ulteriore concentrazione del potere informativo in capo ad un numero limitato di mediatori dell'informazione, potere che, come qualunque altro, può essere utilizzato in maniera indipendente o può essere deviato e condizionato.

Quest'ultima categoria di soggetti porta dunque a riflettere sulle capacità di analisi dei dati. In proposito va sottolineato come tali capacità non dipendano unicamente dalle competenze e conoscenze personali, come nei casi poc'anzi considerati, ma anche dalla disponibilità di strumenti tecnologici. Per dominare il mondo dei *big data* si può infatti agire attraverso il cervello umano, o l'aggregazione di più cervelli in *think tank*, oppure attraverso il cervello informatico dei calcolatori in grado di processare in maniera automatica un'immensa mole di dati e, grazie alle ricerche nel campo dell'intelligenza artificiale, sempre più capaci di agire in maniera esperta sulle informazioni, con modalità che progressivamente vanno ad affinarsi nella prospettiva di creare macchine in grado di replicare i processi cognitivi del nostro cervello. Se si escludono coloro che, per narcisismo o per spirito di condivisione, offrono il proprio sapere agli altri, in molti casi poi i migliori cervelli umani e digitali lavorano in favore di pochi soggetti capaci rispettivamente di ripagarli adeguatamente e di sostenere i costi delle macchine e della ricerca.

L'ultimo elemento che completa l'illustrazione del potere detenuto dai « signori dei dati » ha natura psicologica. Recenti studi hanno infatti dimostrato come la consapevolezza della disponibilità *on-line* di molteplici informazioni induca i soggetti a memorizzare meno di quanto appreso ed a focalizzare invece il ricordo su dove si trovino i dati<sup>3</sup>. Da tale adattamento dei meccanismi della memoria umana al nuovo contesto tecnologico deriva dunque un ulteriore accrescimento del potere dei *gatekeeper* delle informazioni. Ad oggi, dunque, sempre più la conoscenza collettiva sta migrando nel mondo *on-line*, ma nel contempo per accedervi occorre passare dal sentiero stretto dei detentori delle informazioni o dei mediatori necessari al fine del reperimento e della comprensione delle stesse.

Dal quadro che emerge si delinea una disponibilità di dati la cui dimensione è tale da non poter essere più dominata con gli strumenti tradizionali, disponibilità che tuttavia è in vari casi concentrata in mano a pochi e che in tali ipotesi, ma anche in quelle in cui le informazioni siano di libero accesso, richiede comunque un processo elaborativo che nuovamente non è alla portata di tutti. Da questo complesso contesto, qui sommariamente delineato, nascono i « signori dei dati » ed il loro immenso potere.

<sup>3</sup> Cfr. B. SPARROW, J. LIU, D.M. WEGNER, *Google Effects on Memory: Cognitive Consequences of Having Informa-*

*tion at Our Fingertips*, 2011, in [www.sciencemag.org](http://www.sciencemag.org).

2. IL POTERE CORRELATO AI *BIG DATA*.

La concentrazione dell'informazione in mano ad alcuni operatori non rappresenta di per sé una novità, è tuttavia la dimensione dei *big data*<sup>4</sup> unita alla natura globale del fenomeno che pone non pochi interrogativi circa la gestione di un potere informativo che accentra la gran parte dei flussi di dati del pianeta in capo ad un limitato numero di soggetti.

Occorre a tal riguardo rilevare la connotazione peculiare di tale potere, che differisce dalla semplice capacità di profilazione o di schedatura di massa, a cui hanno da sempre guardato le norme in materia di *data protection*. Nel caso dei *big data* emerge infatti una nuova ed ulteriore valenza assai rilevante ovvero la capacità predittiva che le analisi condotte con strumenti sofisticati su tali grandi aggregazioni possono conseguire, da qui una notevole valenza strategica, socio-politica e, non da ultimo, patrimoniale dei *big data*.

L'ampiezza dei *data set* consente infatti di trarre inferenze su fatti ignoti a partire da ricorrenze e correlazioni<sup>5</sup>. Se è vero che tali analisi non seguono il modello dell'indagine a campione di tipo statistico, e conseguentemente possono presentare distorsioni nei risultati, purtuttavia consentono di fare emergere inferenze non prevedibili, fenomeni non ricercati, che si delineano nel dinamismo del quadro complessivo, rispetto ai quali (in quanto ignoti *ex ante*) non sarebbe comunque possibile un campionamento preventivo. Affinché tale metodologia di indagine porti dei risultati occorre però acquisire il più ampio numero di dati possibili; non avendo infatti predefinito l'oggetto dell'indagine non è individuabile una strategia di campionamento ed occorre estendere l'analisi nella maniera più estesa. D'altra parte il ricorso ad un campionamento casuale, nell'ignoranza dei fenomeni attesi, non permetterebbe di raggiungere il medesimo risultato poiché ridurrebbe in ogni caso la complessità eliminando dati potenzialmente utili e rilevanti<sup>6</sup>.

Pur con tutti i limiti attinenti la correttezza dell'analisi eziologica, rilevando ricorrenze e correlazioni piuttosto che formulando vere e proprie osservazioni di carattere statistico, i *big data* possono costituire dunque uno strumento per avvertire il manifestarsi e l'evolvere di macro-tendenze, la cui causa — sempre che interessi — potrà poi essere indagata con metodi più classici<sup>7</sup>. Risulta quindi evidente come chi gestisce grandi quantità di

<sup>4</sup> Fra i primi studi ed analisi sul tema dei *big data* si segnalano le seguenti pubblicazioni: THE ASPEN INSTITUTE, *The Promise and Peril of Big Data*, David Bollier *Rapporteur*, Washington, 2010, in <http://www.aspeninstitute.org/publications/promise-peril-big-data>; OECD, JOINT WPISP-WPIE ROUNDTABLE, *The Economics of Personal Data and Privacy: 30 Years after the OECD Privacy Guidelines*, 1 December 2010, in [http://www.oecd.org/document/22/0,3746,en\\_2649\\_34255\\_46565782\\_1\\_1\\_1\\_1,00.html](http://www.oecd.org/document/22/0,3746,en_2649_34255_46565782_1_1_1_1,00.html); MCKINSEY GLOBAL INSTITUTE, *Big data: The next frontier for innovation, competition, and productivity*, May 2011, in [http://www.mckinsey.com/mgi/publications/big\\_data/](http://www.mckinsey.com/mgi/publications/big_data/); BOYD-CRAWFORD, *Six Pro-*

*vocations for Big Data*, in [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1926431](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431).

<sup>5</sup> Un esempio banale è l'analisi sull'evolversi delle epidemie influenzali effettuate da Google sulla base delle quantità e della localizzazione delle *query* ricevute dal proprio motore di ricerca, cfr. <http://www.google.org/flutrends>.

<sup>6</sup> Per maggiori approfondimenti sul dibattito circa la rilevanza predittiva dei *big data* ed il rapporto con le metodologie statistiche tradizionali, cfr. THE ASPEN INSTITUTE, *op. cit.*, 12 ss.

<sup>7</sup> Cfr. C. ANDERSON, *The End of Theory, Will the Data Deluge Makes the Scientific Method Obsolete*, 2008, in <http://>

dati venga così ad acquisire una capacità predittiva sul futuro agli altri preclusa, costituente un indubbio vantaggio, sia in termini competitivi per le imprese, sia in termini di controllo sociale per gli stati e per i gruppi di potere<sup>8</sup>.

Nel momento in cui le grandi aggregazioni di dati non risultano accessibili a tutti ne consegue altresì che tale potere informativo acquisisce anche un valore economico, specie in una società sempre più bisognosa di dati funzionali all'elaborazione dei processi decisionali. In tal maniera i *data set* divengono una risorsa cedibile a terzi. La disponibilità dei dati non è tuttavia il solo elemento che caratterizza i *big data*, ma, come si è visto, ciò che rileva altresì è la capacità elaborativa, ne consegue che anche con riguardo ad informazioni di libero accesso, chi possiede tali tecnologie potrà trarre gli opportuni risultati predittivi, mentre chi ne è sprovvisto dovrà accontentarsi di indagini ben più limitate e non di rado finirà addirittura per essere vittima di un eccesso di informazioni.

Il potere economico deriva dunque tanto dalla disponibilità esclusiva dei dati, quanto dalla disponibilità di idonei mezzi di analisi, mezzi che consentono di acquisire un potere informativo, economicamente rilevante, anche nel caso di dati « aperti », in ragione del valore aggiunto che l'intermediario dell'informazione è in grado di conseguire e rendere eventualmente accessibile a terzi<sup>9</sup>.

Va infine rilevata un'ulteriore caratteristica peculiare del potere detenuto dai « signori dei dati », ovvero il suo essere geo-localizzato, sotto il duplice profilo della nazionalità di coloro che lo detengono e, specie ove correlato ai grandi *database* gestiti in modalità *cloud*, del luogo in cui le strutture deputate a raccogliere e gestire i dati sono ubicati. Deve però rilevarsi come i due luoghi, quello della sede dei titolari dei dati e quello delle strutture di servizio, non necessariamente coincidono, specie in relazione ai detentori privati delle informazioni. Così le grandi multinazionali hanno sede prevalentemente nel mondo occidentale, mentre non di rado i grandi *data center* vengono delocalizzati altrove, in luoghi ove siano minori i costi di

[www.wired.com/science/discoveries/magazine/16-07/pb\\_theory](http://www.wired.com/science/discoveries/magazine/16-07/pb_theory).

<sup>8</sup> Con riferimento alle strategie di controllo sociale in chiave predittiva si vedano i vari progetti sviluppati negli Stati Uniti dopo il 2001, fondati proprio sull'analisi di grandi quantità di dati, quali i programmi *Total Information Awareness program* e *Novel Intelligence from Massive Data*. In merito a quest'ultimi programmi cfr. le seguenti fonti: DARPA, *Total Information Awareness Program (TIA). System Description Document (SDD), Version 1.1*, July 19, 2002, in <http://epic.org/privacy/profiling/tia/tiasystemdescription.pdf>; EPIC, *Total « Terrorism » Information Awareness (TIA)*, in <http://epic.org/privacy/profiling/tia/>; NATIONAL RESEARCH COUNCIL, *Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment*, Washing-

ton, D.C., 2008, Appendix I e Appendix J; Congressional Research Service, CRS Report for Congress. *Data Mining and Homeland Security: An Overview*, April 3, 2008, in [www.fas.org/sgp/crs/homsec/RL31798.pdf](http://www.fas.org/sgp/crs/homsec/RL31798.pdf); S. HARRIS, *TIA lives on*, in *National Journal*, February 23, 2006, in <http://shaneharris.com/magazines/tia-lives-on/>. Cfr. inoltre il più recente *Open Source Indicators (OSI) Program*, in [http://www.iarpa.gov/solicitations\\_osi.html](http://www.iarpa.gov/solicitations_osi.html)[http://www.iarpa.gov/solicitations\\_osi.html](http://www.iarpa.gov/solicitations_osi.html), che mira a sviluppare « methods for continuous, automated analysis of publicly available data in order to anticipate and/or detect significant societal events, such as political crises, humanitarian crises, mass violence, riots, mass migrations, disease outbreaks, economic instability, resource shortages, and responses to natural disasters ».

<sup>9</sup> Cfr. G. DA EMPOLI, *op. cit.*, 53.

realizzazione e di gestione. La geo-localizzazione comporta rilevanti conseguenze, non solo in termini di leggi applicabili e di maggior o minor tutela dei dati, ma anche di condizionamento da parte del potere politico locale. Non di rado infatti sono riconosciuti dalle leggi alle autorità giudiziarie o amministrative ampi poteri di accesso agli archivi informatizzati, che possono mettere a serio rischio non solo la tutela dei dati personali, ma anche quella dei segreti aziendali o, in caso di soggetti pubblici, dello stato.

In un'ottica geopolitica l'aggregazione di grandi quantità di dati in capo a pochi soggetti, specie ove condizionata dalla disponibilità delle risorse tecnologiche (grandi *data center*, strumenti di analisi dei dati sofisticati, sistemi di *cloud computing*, ecc.) riveste una valenza notevole per l'Unione Europea, posto che nel settore ICT i principali operatori dei diversi ambiti sono statunitensi e, dunque, si genera in capo ad essi una posizione di forza in termini di aggregazione e gestione dei dati, con potenziali rischi sotto il profilo della continuità di erogazione del servizio e, soprattutto, del possibile condizionamento ad opera del contesto legislativo e politico statunitense.

Non è dunque un caso che in tale ambito l'Unione europea abbia recentemente prospettato una nuova politica tesa a tutelare i dati europei ovunque si trovino<sup>10</sup>, anche al di fuori dei confini comunitari, e che tale indirizzo abbia trovato conferma nella bozza di regolamento volto a riformare il quadro normativo comunitario in materia di *data protection*<sup>11</sup>. Si tratta dunque di una svolta destinata a condizionare le politiche degli operatori ICT e finalizzata ad indurre gli stati terzi a soluzioni compromissorie in termini di garanzie concesse, facendo leva sul fatto che nell'attuale contesto globale l'interazione e lo scambio di informazioni sono imprescindibili.

### 3. GLI STRUMENTI UTILI A LIMITARE TALE POTERE: CONTROLLO, ACCESSO, COMPETIZIONE.

A fronte del potere derivante dai *big data* e della natura in gran parte occulta dello stesso, occorre interrogarsi sui rimedi che possono essere introdotti al fine di limitare le asimmetrie e le implicazioni in termini di controllo sociale che ne derivano.

In proposito non pare opportuno ricorrere a rimedi drastici quali l'obbligo di cancellazione delle informazioni trascorso un dato tempo. Obbligo

<sup>10</sup> Cfr. V. REDING, *Your data, your rights: Safeguarding your privacy in a connected world*, SPEECH/11/183, Brussels, 16 March 2011, in <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/11/183&format=PDF&aged=0&language=EN&guiLanguage=it>. KROES e ID., *Privacy matters - Why the EU needs new personal data protection rules*, Brussels, 30 November 2010, in <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/10/700>.

<sup>11</sup> Cfr. art. 2 (2) della Proposal for a Regulation of the European Parliament and of the council on the protection of individuals with regard to the processing of

personal data and on the free movement of such data (General Data Protection Regulation), Version 56 del 29 novembre 2011, consultabile in [http://staff.polito.it/alessandro.mantelero/dati\\_personali.html](http://staff.polito.it/alessandro.mantelero/dati_personali.html). Nel lasso di tempo intercorso fra la redazione del presente contributo e la correzione delle bozze è stato presentato il testo definitivo della proposta di Regolamento, consultabile al seguente indirizzo: [http://ec.europa.eu/justice/data-protection/document/review2012/com\\_2012\\_11\\_en.pdf](http://ec.europa.eu/justice/data-protection/document/review2012/com_2012_11_en.pdf). Cfr. ora art. 3 (2) della proposta, ove con riguardo all'efficacia del Regolamento è stata adottata una formulazione in parte più restrittiva rispetto alla bozza del novembre 2011.

che, nel contesto delle informazioni di libero accesso *on-line*, si vorrebbe ricondurre al diritto all'oblio<sup>12</sup>. Senza soffermarsi sulla nota differenza fra diritto alla cancellazione dei dati e diritto all'oblio, va comunque rilevato come una simile soluzione paia scarsamente praticabile in relazione alle comunicazioni *on-line*<sup>13</sup> e poco efficiente rispetto alla creazione di grandi *database*. A ciò si aggiunga che in molti casi le basi di dati su cui vengono poste in essere le elaborazioni qui in esame non contengono necessariamente dati personali, ma non per questo risultano meno rilevanti ai fini predittivi.

Altre sembrano allora le modalità di intervento destinate ad avere maggiore efficacia nel ridimensionare e ridistribuire il potere informativo detenuto da pochi e nel limitarne i possibili abusi.

A tali fini occorrerebbe agire tanto sul mercato, stimolando la competizione e favorendo così il pluralismo di attori, quanto sull'accessibilità all'informazione. Ove infatti i dati detenuti dai soggetti in esame fossero in gran parte accessibili, si aprirebbe a nuovi interessati la possibilità di trarre dagli stessi inferenze (non solo grandi operatori non ancora presenti nel mercato, ma anche vasti gruppi di singoli individui capaci di aggregare oltre la soglia critica le limitate risorse di ciascuno).

La rilevanza assunta in termini di potere informativo dalle grandi concentrazioni di dati, unitamente al valore strategico anche per le nazioni delle stesse, dovrebbe poi indurre a valutare l'adozione di forme di controllo di tali aggregazioni di dati, prevedendo specifiche autorità sovranazionali indipendenti ed introducendo obblighi di notificazione. Non a caso la notificazione della costituzione delle nuove banche dati era uno degli obblighi caratterizzanti le prime legislazioni in materia di trattamento dei dati personali, in un tempo in cui le risorse informatiche erano patrimonio di pochi, accentrate in luoghi specifici e portatrici (per quegli anni) di un potere di controllo prima inimmaginabile. L'analogia fra l'epoca dei *main frame* e quella attuale del *cloud computing* e dei *big data* è significativa, poiché nuovamente (pur restando un potere informatico distribuito) le grandi risorse informatiche si concentrano in mano a pochi soggetti e risultano anche fisicamente aggregate in enormi *data center*. È dunque nuovamente possibile conoscere chi crea tali grandi basi di dati, chi le gestisce e, quindi, porre in essere le attività di controllo necessarie a garantire la sicurezza delle informazioni inerenti i cittadini.

In tal senso la creazione di autorità di controllo sovranazionali dovrebbe da un lato incidere sulla standardizzazione dei servizi in termini soprattutto di sicurezza, ma dovrebbe anche servire come strumento per sorvegliare ed eventualmente contenere sia le pretese invasive dei governi, sia gli eventuali abusi dei detentori/gestori dei *big data*.

<sup>12</sup> Cfr. V. REDING, *Your data, your rights: Safeguarding your privacy in a connected world*, cit. Con riguardo a *right to be forgotten and to erasure* si veda anche l'art. 15 della bozza di General Data Protection Regulation, cit. ed ora art. 17 del testo definitivo, cfr. *supra*, nota 11.

<sup>13</sup> Un primo ostacolo è dato dalla facilità con cui le informazioni possono essere replicate *on-line* nei meandri più disparati

della Rete; a ciò si aggiunga che, come si avrà modo di rilevare *infra* nel testo, la Rete rappresenta ad oggi una sorta di memoria/conoscenza collettiva che un indiscriminato diritto di cancellazione potrebbe impoverire. Ove ciò accadesse si finirebbe inoltre per acuire il potere di coloro i quali hanno risorse sufficienti per creare autonomamente grandi archivi di informazioni, non liberamente accessibili.

Tale soluzione potrebbe apparire in contrasto con la spontanea evoluzione dei processi informativi e di *business*, ed essere vista come un'indebita ingerenza di poteri pubblici nello sviluppo dei sistemi e dei modelli economici, tuttavia va ricordato che proprio la dimensione internazionale di internet ha progressivamente spinto ad una revisione del dogma iniziale incentrato sulla semplice autoregolamentazione degli utenti. In una società pesantemente condizionata dalle informazioni, dalla loro appropriazione e gestione, non pare così illogico che tali risorse, qualora assumano proporzioni assai rilevanti, vengano poste sotto un limitato controllo come avviene per altre risorse di vitale importanza e di valore strategico (dalle centrali nucleari agli spazi aerei). Si tratta ovviamente di un percorso molto lungo e tortuoso poiché richiede sinergie a livello internazionale, tuttavia occorrerebbe iniziarlo quanto prima, magari avvalendosi degli organismi già esistenti, onde evitare di introdurre forme di regolamentazione quando ormai sarà troppo tardi. Ovviamente è opportuno graduare in maniera adeguata l'intervento, rivolgendolo non a qualsiasi *data farm* realizzata in una qualche parte del globo, bensì solamente a quelle che hanno una dimensione assolutamente rilevante o che, in ragione dei dati trattati (es. sicurezza nazionale e militare), hanno un'importanza primaria.

Un effetto a valle di simili decisioni si avrebbe poi non solo in termini di sicurezza dei dati ed uniformità degli *standard*, bensì anche di trasparenza della società dell'informazione, intesa come conoscenza dei soggetti cui è allocato un notevole potere informativo. Questo costituirebbe il presupposto necessario per agire sull'altro importante fronte dell'accessibilità dei dati e della condivisione delle informazioni, in maniera tale da contenere il potere dei gestori dei *big data* e nel contempo permettere all'intera società di accedere ad una risorsa vitale quale sono le informazioni.

Con riguardo a quest'ultimo profilo è evidente il ruolo che può avere il tema degli *open data*<sup>14</sup>, laddove l'apertura ai cittadini delle banche dati pubbliche e la possibilità di fruire di dati grezzi ed elaborabili non solo riduce in maniera diretta il potere dei detentori delle informazioni sotto il profilo dell'acquisizione dei dati, ma anche sotto quello, forse ancor più rilevante, delle capacità tecniche e culturali di indagine. Se infatti gli archivi aperti non permettono certo all'uomo medio di poter trarre

<sup>14</sup> Con il termine di *open data* si intendono quelle informazioni suscettibili di essere liberamente utilizzate, riutilizzate e ridistribuite, nel rispetto di alcune specifiche modalità di diffusione, cfr. per un maggior dettaglio *Open Definition Defining the Open in Open Data, Open Content and Open Services*, in <http://opendefinition.org>. Non essendo possibile approfondire in questa sede un tema così vasto e rilevante, si rinvia per un approccio generale al problema alle attività dell'Open Knowledge Foundation (<http://okfn.org/>), mentre con riferimento al più indagato ambito dell'informazione del settore pubblico, si vedano invece il progetto di ricerca LAPSI-The European Thematic Network on Legal

Aspects of Public Sector Information (<http://www.lapsi-project.eu/about>) ed all'European Public Sector Information Platform (<http://epsiplatform.eu/>), nonché il portale dedicato della Commissione europea ([http://ec.europa.eu/information\\_society/policy/psi/index\\_en.htm](http://ec.europa.eu/information_society/policy/psi/index_en.htm)). Sul tema del riutilizzo dei dati pubblici cfr. anche il progetto di ricerca Extracting Value from Public Sector Information (<http://www.evpsi.org/>). In dottrina si vedano ZENO-ZENCOVICH, *Usa a fini privati dei dati personali in mano pubblica*, in questa *Rivista*, 2003, 197 ss., ed i diversi recenti contributi sul tema raccolti in D. TISCORNIA (a cura di), *Open Data e riutilizzo dei dati pubblici*, in *Informatica e dir.*, 2011.

dagli stessi tutti i benefici possibili in termini di analisi, tuttavia proprio l'accessibilità consente ai singoli di unire i loro sforzi, anche fuori da logiche imprenditoriali (si pensi all'associazionismo di settore), per creare gruppi di studio o per aggregare risorse finanziarie utili all'acquisizione di strumenti informatici di elaborazione, dando vita a poli alternativi di formazione della conoscenza a partire dai *big data*<sup>15</sup>.

Un ultimo profilo di criticità che occorre superare concerne il ruolo dell'Europa nel contesto dei processi di aggregazione delle informazioni presi in considerazione. In proposito assume rilievo il confronto con gli USA ed emerge una posizione di vantaggio in capo a quest'ultimi. Sul piano della gestione dei dati pubblici gli Stati Uniti vantano non solo un'omogeneità strutturale ancora ignota alla giovane Unione Europea, ma hanno anche investito significative risorse nell'ammodernamento della stessa attraverso il ricorso alle tecnologie di *cloud computing*, favorendo l'aggregazione delle basi di dati<sup>16</sup>, presupposto necessario sia per l'incremento e l'esercizio da parte dei soggetti pubblici del proprio potere informativo, sia per l'accessibilità in chiave democratica a tali informazioni. Sul versante privato, poi, seppur è innegabile che le grandi imprese europee rientrino a pieno titolo fra i gestori dei *big data*, tuttavia l'eccellenza delle società statunitensi in alcuni settori strategici dell'ICT (motori di ricerca, servizi di *cloud computing*, piattaforme UGC, *social network*), pongono quest'ultime in una posizione di vantaggio, atteso che proprio negli ambiti più strettamente correlati all'ICT si assiste ai maggiori flussi di dati.

In un'ottica geo-politica e di politica industriale, tale assetto può rivelarsi una debolezza per i Paesi europei, in termini di perdita di controllo sui dati dei cittadini e di affidamento della gestione di informazioni strategiche a soggetti stranieri. Non è forse dunque un caso che recentemente si sia fatta strada l'idea di una più forte tutela dei dati personali incentrata sul concetto di appartenenza all'Unione del soggetto cui le informazioni si riferiscono<sup>17</sup>, come non è un caso che l'industria europea venga sollecitata ad assumere un ruolo maggiormente rilevante nella realizzazione delle nuove architetture informatiche<sup>18</sup>. Trascurare tali sollecitazioni, può essere un rischio perché comporta l'utilizzo di servizi ero-

<sup>15</sup> Va in proposito rilevato come le imprese ed i governi dispongano di risorse informatiche tali, in termini di strumenti di indagine, per cui il tempo di latenza fra analisi e soluzione del problema risulta assai ridotto, con un netto vantaggio sotto il profilo predittivo. Le alternative di tipo civico e comunitario sono dunque destinate a soccombere in termini di rapidità di raggiungimento dei risultati, tuttavia rispetto ai temi di interesse per la società civile la velocità può non necessariamente costituire un obiettivo. In tali ambiti ad essere ricercata è piuttosto la capacità di analisi ed interconnessione fra una moltitudine di dati (si pensi ad esempio all'uso dei *big data* in chiave di contrasto alla corruzione), essendo invece la celerità maggiormente connessa alla competizione di mercato.

<sup>16</sup> Cfr. V. KUNDRA, U.S. CHIEF INFORMATION OFFICER, *Federal Cloud Computing Strategy*, February 8, 2011, in <http://www.cio.gov/documents/federal-cloud-computing-strategy.pdf>, nonché i documenti pubblicati sul sito governativo [www.cio.gov](http://www.cio.gov).

<sup>17</sup> Cfr. *supra* note 10 e 11.

<sup>18</sup> Cfr. N. KROES, *The Digital Agenda: Europe's key driver of growth and innovation*, SPEECH/11/629, Brussels, 4th October 2011, 5, in <http://europa.eu/rapid/pressReleasesAction.do?reference=SPEECH/11/629&type=HTML>. Cfr. altresì EXPERT GROUP REPORT, *The Future Of Cloud Computing*, rapporto redatto per la Commissione europea, 2010, 35 ss., in <http://cordis.europa.eu/fp7/ict/issai/docs/cloud-report-final.pdf>.

gati da imprese legate a poteri e condizionate da regole giuridiche stranieri, che potenzialmente possono perseguire finalità difformi da quelle dei cittadini e degli stati dell'Unione. Occorre dunque sia stimolare la competizione nell'ambito dello sviluppo delle nuove tecnologie del settore ICT, sia rafforzare il quadro normativo di tutela delle informazioni<sup>19</sup>.

<sup>19</sup> In quest'ultimo senso va sicuramente la revisione del quadro regolamentare comunitario in materia di *data protec-*

*tion*, stando a quanto emerge dal testo della bozza di General Data Protection Regulation, *cit. supra* nota 11.