Il Mulino - Rivisteweb

**Davide Quaglione, Cesare Pozzi**

# Economia dei big data: lineamenti del dibattito in corso e alcune riflessioni di policy

**(doi: 10.1430/90435)**

L'industria (ISSN 0019-7416)
Fascicolo 1, gennaio-marzo 2018

**Ente di afferenza:**
*Università Luiss (luiss)*

Copyright © by Società editrice il Mulino, Bologna. Tutti i diritti sono riservati.
Per altre informazioni si veda https://www.rivisteweb.it

**Licenza d'uso**

L'articolo è messo a disposizione dell'utente in licenza per uso esclusivamente privato e personale, senza scopo di lucro e senza fini direttamente o indirettamente commerciali. Salvo quanto espressamente previsto dalla licenza d'uso Rivisteweb, è fatto divieto di riprodurre, trasmettere, distribuire o altrimenti utilizzare l'articolo, per qualsiasi scopo o fine. Tutti i diritti sono riservati.

# Economia dei big data: lineamenti del dibattito in corso e alcune riflessioni di policy

Davide Quaglione e Cesare Pozzi

**Big data economics: The features of the ongoing debate and some policy remarks**

There is increasing awareness that the collection of big data and the new data analytics capabilities will have deep social and economic impacts. Benefits are expected in terms of higher quality and variety of goods and services, higher firms' productivity and, ultimately, more sustainable and robust economic development. At the same time, some concerns are raised in terms of possible negative consequences on the way the competitive processes in the big data and related markets will unfold. This paper presents an overview of the most significant features of the ongoing scientific debate, with the aim of critically deriving some implications on the most relevant policy dimensions, with an eye to the Italian case.

**Keywords**: Big data, Antitrust, Industrial Policy, Research Policy
**Classificazione JEL**: **L100; L400**

## 1. INTRODUZIONE

Che i *big data*, con le tecnologie dell'informazione e della comunicazione ad essi complementari, saranno in futuro uno tra i motori più significativi dello sviluppo economico è, ormai, molto più di una suggestione accademica. Anche prescindendo da alcune letture particolarmente mitologiche del fenomeno – secondo cui per conoscere i legami tra le grandezze economiche, sociali, politiche e ambientali sulla base dei quali disegnare strategie aziendali e politiche economiche sarebbe sufficiente «far parlare i dati» senza neanche dotarsi di ipotesi di ricerca e modelli teorici da verificare (Anderson, 2008) – non vi è dubbio che la disponibilità di enormi moli di informazioni granulari, praticamente in tempo reale, sui comportamenti e sulle abitudini di consumo degli individui rappresenti una circostanza i cui effetti economici e sociali saranno dirompenti.

Le implicazioni dei *big data* sui mercati digitali in termini di miglioramenti della qualità e varietà dell'offerta nonché di innovazioni nei possibili modelli di business sono già evidenti, con grandi player configurati come

Cesare Pozzi, Professore Ordinario di Economia Applicata, Dipartimento di Giurisprudenza, LUISS «Guido Carli», cpozzi@luiss.it
Davide Quaglione, Professore Associato di Economia Applicata, Dipartimento di Economia, Università degli studi «G. d'Annunzio» di Chieti e Pescara, Tel: 085.453.7610, davide.quaglione@unich.it

piattaforme multilaterali che erogano agli utenti servizi, spesso fortemente *customizzati* e a titolo gratuito, in cambio di informazioni personali. Come sottolineato da più parti, anche in questa rivista (Andreoni, 2017; Pontarollo, 2016; 2017), effetti positivi diffusi sono altresì attesi nei settori manifatturieri e del primario, a patto che il percorso evolutivo di tali tecnologie sia adeguatamente accompagnato da politiche pubbliche (industriali, antitrust, della ricerca e della formazione) che assicurino: la realizzazione delle condizioni strutturali di base necessarie a coglierne i possibili benefici (ad esempio l'introduzione delle innovazioni organizzative e architetturali necessarie a livello di impresa, la riduzione del *digital divide*); l'esplicarsi di dinamiche competitive sane; la disponibilità del capitale umano indispensabile per la loro adozione pervasiva.

L'obiettivo di questo articolo è di proporre una ricognizione breve, e per questo certamente non esaustiva, degli elementi che caratterizzano l'economia dei *big data* e il dibattito scientifico intorno ad essa, derivandone alcune preliminari implicazioni sui versanti di *policy* più significativi.

## 2. *BIG DATA* E POLITICHE DI TUTELA DELLA CONCORRENZA

Come ben sottolineato da Pitruzzella (2016), nei mercati in cui la conoscenza derivata dai dati (*data-driven knowledge*) risulta essere più rilevante, la presenza di effetti di rete positivi implica, almeno fino ad una certa misura, rendimenti crescenti nella raccolta dei dati che comporta una tendenza marcata all'aumento della concentrazione e al consolidamento del potere di mercato delle piattaforme multilaterali *incumbent*. Tale circostanza è percepita da alcuni studiosi (ad esempio si veda Stucke, Grunes 2015) come potenzialmente distorsiva dell'esplicarsi del corretto processo competitivo nel mercato, nella misura in cui la maggiore capacità dei *first movers* di accumulare grandi volumi di dati costituirebbe una barriera all'entrata dei rivali, con una conseguente limitazione della contendibilità del mercato stesso e, inoltre, con il rischio che l'*incumbent* sia in condizioni ancora più favorevoli per adottare comportamenti escludenti. Ulteriori frizioni nel processo concorrenziale potrebbero derivare da forti asimmetrie non soltanto nell'accesso ai dati, ma anche nelle capacità di fruizione delle utilità del patrimonio informativo in essi contenuto (Buzzacchi, 2016): i professionisti dei *big data* si trovano in posizione di privilegio, perché esperti e perché dotati di strumentazioni tecnologiche adeguate nell'interpretazione dell'informazione, il che garantisce loro un vantaggio significativo.

Rispetto a tali preoccupazioni, la posizione attualmente prevalente in letteratura è certamente caratterizzata da una maggiore moderazione.

Ferma restando la necessità che l'Autorità di tutela della concorrenza vigili opportunamente e intervenga in caso di comportamenti abusivi da valu-

tare caso per caso, molti sottolineano che l'esistenza di un vantaggio competitivo e di barriere all'entrata connessi alla disponibilità dei dati non necessariamente implica che gli stessi siano un input essenziale [1] senza l'accesso al quale i concorrenti non potrebbero operare profittevolmente. Pertanto, ne consegue che l'eventuale rifiuto a contrarre dell'operatore in posizione dominante non possa considerarsi automaticamente un abuso. Come sostenuto da Lerner (2014), allo stato attuale le *big data companies* non adottano alcun comportamento che abbia per effetto quello di inibire a rivali più piccoli di accedere ai dati degli utenti, né esse controllano tutta l'informazione o porzioni significative di essa. Infatti, non si registrano casi di accordi di esclusiva nella raccolta dei dati e allo stesso tempo i dati sono beni non rivali, cosicché la raccolta e l'utilizzo di informazioni relative agli utenti da parte di un operatore non pregiudica ad altri la raccolta e l'utilizzo delle stesse informazioni. Inoltre, gli utenti tendono a fornire informazioni personali a più piattaforme simultaneamente (il cosiddetto *multi-homing*), il che riflette un elevato grado di differenziazione dei servizi online, in un contesto in cui gli *switching costs* sono molto bassi, se non nulli.

Un ulteriore elemento che contribuisce a stemperare le preoccupazioni di distorsioni del processo concorrenziale risiede nell'esistenza di meccanismi e schemi di incentivi connaturati all'accumulazione di dati che tendono ad assottigliare progressivamente i vantaggi dei grandi operatori. Superata una certa soglia (non necessariamente elevata) quantitativa e qualitativa di dati raccolti, infatti, è da ritenersi che i benefici marginali di ulteriori accumulazioni o qualità degli stessi diventino decrescenti (Sokol, Comerford 2016; Schepp, Wambach 2016; Lerner, 2014; Lambrecht, Tucker 2017). Così, il valore di scale incrementali può tendere allo zero per gli operatori grandi, mentre può risultare significativo per i competitor più piccoli, con l'ulteriore conseguenza che per questi ultimi l'incentivo ad investire evidentemente sarebbe, al margine, più elevato. Inoltre, Sokol e Comerford (2016) evidenziano anche che i dati hanno una vita utile limitata: se la sorgente del vantaggio competitivo risiede nella capacità di disporre di dati sempre nuovi e differenziati, i piccoli operatori rivali non accusano svantaggi se l'*incumbent* dispone di informa-

[1] D'altra parte, argomenta Pitruzzella (2016), il diniego di accesso opposto dall'operatore dominante potrebbe essere oggetto di sanzione da parte dell'Autorità antitrust anche se avesse per oggetto big data ad esclusivo uso interno, qualora per gli stessi fosse riconosciuta la natura di *essential facility* secondo la giurisprudenza comunitaria. Tucker e Wellford (2014) sostengono che in assenza di vendita, non si debba parlare di prodotto né tantomeno della definizione del relativo mercato. Al contrario, Pitruzzella (2016, p. 23) scrive: «In fact, when examining a potential refusal to supply, it is not necessary for the refused input to have been already traded: it is sufficient that there is demand from potential purchasers and that a potential market for the input at stake can be identified. Therefore, the definition of the relevant market and the finding of dominance does not impinge upon whether or not the data is actually traded».

zioni generaliste e datate. In altre parole, i potenziali concorrenti non hanno necessariamente bisogno di accumulare dati fino alla scala dell'*incumbent*, ma piuttosto di accumulare dati rilevanti e aggiornati (Tucker, Wellford 2014).

Insomma, la breve ricognizione proposta finora lascia intendere che esistano elementi per ritenere che i *big data* non siano un contesto che merita particolare allarmismo antitrust.

Allo stesso modo, tuttavia, non sembrano esserci dubbi sulla necessità di ammodernare i criteri e gli strumenti utilizzati nelle valutazioni antitrust in relazione ai mercati dei *big data*, per effetto di diverse peculiarità che li contraddistinguono: tra esse, la già menzionata non rivalità dei dati, la natura di piattaforma multilaterale degli operatori e la connessa non tradizionale struttura dei prezzi praticati sui lati della piattaforma, la possibilità di discriminazioni di prezzo estreme (i cosiddetti prezzi personalizzati).

La prima questione sollevata in letteratura è relativa alla definizione del mercato rilevante. L'individuazione dei prodotti e dei mercati geografici che esercitano vicendevolmente un vincolo competitivo è, come noto, un passaggio indispensabile e preliminare a qualsiasi valutazione del potere di mercato e di possibile condotte anticoncorrenziali. Con particolare riferimento alle operazioni di fusione/acquisizione tra operatori che offrono servizi basati sui dati a imprese che li usano come input nel proprio processo produttivo, nella riflessione scientifica condotta finora appare emergere chiaramente che parlare di mercato rilevante dei *big data* sia del tutto improprio, tenuto conto che i *big data* non possono essere considerati come un insieme di informazioni indistinte; al contrario, essi includono diversi tipi di informazioni che soddisfano imprese differenti con bisogni eterogenei (Pitruzzella, 2016). Per questa ragione, il criterio della sostituibilità dal lato della domanda diventa centrale, senza che invece rilevino la fonte e le modalità con cui i dati sono raccolti (Di Porto, 2016). Bagnoli (2016), con l'obiettivo di identificare con maggiore accuratezza la struttura del mercato dei *big data* e consentire di verificare con maggiore chiarezza l'esistenza di preoccupazioni concorrenziali (ad esempio, potere di mercato, barriere all'entrata, abuso di dominanza), ne propone la segmentazione in tre parti: 1) raccolta; 2) immagazzinamento e 3) analisi.

In secondo luogo, la natura di piattaforme multilaterali di molti operatori dei *big data*, e la circostanza che essi pratichino prezzi nulli sul lato della piattaforma attraverso cui raccolgono i dati, mette in crisi alcuni strumenti tradizionalmente utilizzati nell'analisi antitrust. Ad esempio, l'utilizzo del test dell'ipotetico monopolista (il cosiddetto test SSNIP) è, in questi contesti e nella sua formulazione standard, precluso poiché quando il prezzo di partenza dell'analisi è pari a zero non è evidentemente matematicamente possibile contemplare alcun incremento percentuale dello stesso; analogamente, l'indice di Lerner non può essere utilizzato per quantificare il potere di mercato (Evans, 2011) risultando nella fattispecie una frazione con denominatore nullo. Come

suggerito da Evans (2012), sebbene sia teoricamente possibile predisporre un'estensione del test dell'ipotetico monopolista per i mercati a più versanti, la sua implementazione pratica sarebbe particolarmente difficoltosa.

Un ultimo profilo di criticità legato ai mercati dei *big data* e a comportamenti degli operatori potenzialmente dannosi per i consumatori finali è stato sollevato per l'inedita possibilità che essi hanno di utilizzare schemi di discriminazione di prezzo estremi, conosciuti come «prezzi personalizzati», che potrebbero peggiorare la situazione di alcuni consumatori senza che questi abbiano consapevolezza che tale strumento di prezzo in effetti sia reso possibile dall'aver consentito l'accesso ai propri dati personali. Mentre non pare in discussione l'inopportunità dell'utilizzo di prezzi discriminati quando gli acquirenti sono imprese – poiché quelle che pagano prezzi più alti si troverebbero ad accusare uno svantaggio competitivo che altera l'esplicarsi della concorrenza nel mercato in cui operano, a maggior ragione se l'operatore *incumbent* dei *big data* dovesse essere verticalmente integrato a valle, nel qual caso si ravviserebbe la fattispecie del *margin squeeze* – l'assenza di collegamento tra il pregiudizio di un consumatore sottoposto a prezzi personalizzati e un «danno alla concorrenza» suggerisce che sarebbe sbagliato ragionare in via analogica (per approfondimenti sul tema si vedano Maggiolino, 2016; Bourreau *et al.*, 2017). Tra gli ulteriori motivi addotti a supporto di una lettura essenzialmente non patologica dei prezzi personalizzati, allo stato peraltro talmente poco diffusi da renderli piuttosto un'ipotesi di scuola, vi è innanzitutto la circostanza che l'impatto di prezzi discriminati sul benessere sociale sia ambiguo, potendo in certe condizioni persino risultare in un miglioramento dell'efficienza allocativa. In secondo luogo, gli operatori potrebbero non trovare utile adottare la pratica di prezzi personalizzati per paura che essa si ritorca loro contro. Per dirla con le parole di Maggiolino (2016, p. 98): «Molte ricerche osservano [...] che gli individui, anche perché abituati a confrontarsi con la società dei consumi di massa dove i prezzi praticati dalle imprese paiono uguali per tutti, continuano a percepire la parità dei prezzi come un valore e, al contrario, i prezzi personalizzati come un disvalore tale da addirittura danneggiare la reputazione delle imprese che li praticano».

Infine, nonostante si tratti ancora più di una speculazione sul piano scientifico che di una problematica in essere, un cenno va riservato al tema della cosiddetta «collusione algoritmica» ed al trattamento antitrust della stessa qualora dovesse emergere. Ezrachi e Stucke (2017) prima e ITMEDIA (2018) poi evidenziano che una possibile conseguenza dell'utilizzo di algoritmi applicati ai *big data* per disegnare strategie competitive aziendali possa determinare forme di collusione tra gli operatori: come riportato da ITMEDIA (2018, p. 21), «le imprese possono congegnare [algoritmi] per perseguire i più disparati obiettivi o svolgere le più differenti funzioni. [...] a destare maggiore preoccupazione è l'ipotesi per cui ogni operatore arriva a praticare un prezzo

collusivo, sebbene abbia sviluppato il proprio algoritmo in modo autonomo e indipendente». Diversamente dall'ipotesi in cui gli algoritmi disegnati dalle imprese siano esplicitamente costruiti con l'obiettivo di massimizzare i propri profitti senza escludere comportamenti collusivi, nel qual caso la posizione sostenuta da Margrethe Vestager per l'Europa [2] è quella di ritenere responsabili e quindi sanzionabili per l'intesa coloro che hanno congegnato l'algoritmo, non sussisterebbe violazione della normativa antitrust qualora il coordinamento fosse il risultato involontario di comportamenti autonomi, unilaterali e indipendenti.

## 3. BIG DATA, POLITICHE INDUSTRIALI E DELLA RICERCA

Se mai la crisi economico-finanziaria originata nel 2008 ha avuto lasciti positivi, tra di essi va annoverata una ritrovata attenzione al ruolo e all'importanza delle politiche industriali. Dopo anni di politiche orizzontali e dell'innovazione incapaci di favorire un'evoluzione virtuosa del sistema economico europeo (Sarra *et al.*, 2018), men che meno di quello italiano, sarebbe il momento di predisporre politiche che non perseguano aprioristicamente ricette e obiettivi preconfezionati, ma che assecondino una crescita e uno sviluppo coerenti con le aspettative dei cittadini e le vocazioni delle comunità cui appartengono. Per dirla con le parole di Cassetta e Pozzi (2016, p. 535) «la politica industriale racchiude tutti quegli interventi che consentono di garantire la sostenibilità della traiettoria culturale che la comunità legittimamente sceglie e che si esprime, in un'economia di mercato, nei rapporti di scambio con le altre comunità». Questa visione, tuttavia, può essere messa in grande difficoltà in un contesto dominato dai *big data* – non solo perché non tutti i sistemi produttivi sono pronti per godere appieno dei benefici legati all'economia dei dati, cosicché in assenza di investimenti in specifiche tecnologie e conoscenze il percorso di crescita potrebbe finire per generare o rafforzare divari tra territori e tra individui (Di Tommaso, Barbieri 2017) – ma anche perché è serio il rischio che si confondano le esigenze e le vocazioni delle collettività con ciò che potrà inferirsi, con obiettivi e algoritmi interpretativi propri di chi provvederà, dai dati generati dal sottoinsieme digitale degli individui che ne fanno parte. Su quest'ultimo punto si tornerà nel paragrafo conclusivo.

Per quanto attiene, invece, alle criticità da superare, la situazione italiana è un caso interessante cui fare riferimento. L'alfabetizzazione digitale di fami-

[2] Qui si fa riferimento ad uno *speech* tenuto il 16 marzo 2017 a Berlino, alla Bundeskartellamt 18th Conference on Competition, reperibile al seguente link: https://ec.europa.eu/commission/commissioners/2014-2019/vestager/announcements/bundeskartellamt-18th-conference-competition-berlin-16-march-2017_en.

glie e imprese è notoriamente ancora modesta, pur dopo sforzi con i quali si è tentato di favorire l'accesso alle nuove tecnologie dell'informazione e della comunicazione. Ad esempio, emblematico è il caso della penetrazione della connettività a banda larga tra gli individui: nonostante una copertura praticamente integrale della popolazione, la percentuale di individui che dispongono di banda larga fissa è la più bassa nell'Unione Europea a 28 (Quaglione, 2017). Dal lato delle imprese, l'adozione di una nuova tecnologia ha costi che vanno ben al di là del prezzo per il suo acquisto. Sono spesso richiesti cambiamenti organizzativi significativi, la disponibilità di fattori produttivi complementari, alcuni caratterizzati da indivisibilità che mal si coniugano con dimensioni aziendali piccole o micro (notoriamente prevalenti nel sistema economico italiano). In un contesto strategico in cui i dati sono centrali, la fruizione delle informazioni a tutti i livelli dell'impresa richiede la capacità di gestire il patrimonio informativo, di derivarne indicazioni operative, di modellare conseguentemente i processi, le architetture e le tecnologie disponibili (Buzzacchi, 2016; Traversi, Ruggieri 2013). A questo proposito, uno dei filoni scientifici maggiormente innovativi è quello relativo all'*Enterprise Architecture*, un approccio olistico che, in estrema sintesi, ha la finalità di allineare i processi alle strategie. Come si può intuire da queste poche righe, si tratta di un approccio per lo più rivolto alle grandi imprese, tuttavia recenti sviluppi ne hanno verificato le modalità di applicazione alle piccole e medie imprese (Bernaert *et al.*, 2014). Non è questa la sede per approfondire il tema; ciò che qui si vuole sottolineare è che l'introduzione a livello delle piccole e medie imprese di cambiamenti significativi come quelli implicati dalla rivoluzione dei *big data* richiede non soltanto capacità di acquisto delle tecnologie necessarie, ma anche, e soprattutto, una mutazione radicale dell'approccio all'imprenditorialità e alla gestione aziendale (Cassetta *et al.*, 2016). Sebbene sia ancora presto per valutarne i risultati, ad esempio, le misure di incentivazione incluse nel pacchetto Industria 4.0 avranno certamente favorito una transizione del sistema produttivo alle nuove tecnologie, ma è assai probabile che tali opportunità siano state colte da imprese che erano già predisposte a quel cambiamento e che, prima o poi, l'avrebbero percorso anche senza gli incentivi. La vera sfida è invece quella di alimentare un cambio di mentalità nell'intero tessuto imprenditoriale, prima che il vantaggio competitivo di sistemi più pronti del nostro diventi inaffrontabile (Losurdo *et al.*, 2018). In questo, e chi scrive ne è profondamente convinto, il sistema della ricerca universitaria potrà e dovrà necessariamente fornire il proprio contributo, e perché ciò accada, i canali formali e informali di trasferimento della conoscenza accademica dovranno superare ostacoli che fino a questo momento, salvo qualche rara eccezione, sono risultati fatali. Attività di consulenza conto-terzi, brevetti e generazione di spin-off accademici potranno davvero fertilizzare il sistema economico soltanto se le singole università davvero si orienteranno in questo senso. Uffici del trasferimento tecnologico di facciata, ricavi da conto-

terzi trattenuti dagli Atenei o dai Dipartimenti e utilizzati per finanziare attività che niente hanno a che vedere con la ricerca, schemi di incentivazione razionale e monetari per i ricercatori del tutto inadeguati, eterogeneità tra i regolamenti universitari tali da rendere di fatto impossibile per un'impresa chiedere una consulenza che coinvolga più università sono soltanto alcune delle criticità sollevate in letteratura, con esplicito riferimento alla situazione italiana (Muscio *et al.*, 2015; 2016).

## 4. ALCUNE ULTERIORI RIFLESSIONI GENERALI

Nelle scienze sociali, ed in particolare in economia, si è da sempre ceduto alla tentazione di rappresentare l'interazione tra gli individui come approssimazione di legami e regole naturali, come se i sistemi sociali ed economici fossero caratterizzati da un ordine implicito, e gli scostamenti da esso fossero un «fallimento» dei partecipanti (Pozzi, 2014). Sia chiaro, una qualsiasi teoria scientifica è in effetti una proposta di oggettivizzazione della realtà, alcuni modelli sono più aderenti altri meno, alcuni matematicamente più raffinati altri più semplici; tutti però sono utili, a patto che chiunque li studi e li utilizzi (per stabilire le proprie strategie d'impresa, le policy da adottare, gli investimenti da effettuare) sia consapevole che le indicazioni che se ne possono trarre dipendono strettamente dalle ipotesi di base e dalle forme funzionali scelte per rappresentare le relazioni tra le variabili coinvolte, e vanno dunque adeguatamente contestualizzate e soppesate. Quando si dimenticano principi basilari come questo, accade che il rapporto tra approccio scientifico e conoscenza della realtà si ribalti: se teoria e realtà non si sovrappongono non è perché il modello sia impreciso o parta da ipotesi inverosimili, ma perché è la realtà ad essere fallace. E, quasi senza neanche accorgersene, questa inversione logica ci se la ritrova nel lessico, cosicché si finisce per usare con naturalezza locuzioni che sono non solo controfattuali, ma che finiscono per diventare acriticamente strutture valoriali se non ideologiche di riferimento. Tra gli innumerevoli esempi: il modello di Cournot (1838) [3], in ragione del quale si è definito «perfetto» l'equilibrio di una concorrenza in cui le imprese non hanno alcuna leva per competere, utilizzato ancora oggi come riferi-

[3] Il modello di concorrenza perfetta nasce dalla riflessione di Cournot su cosa accada in equilibrio quando gli effetti della concorrenza raggiungono il loro limite: «The effects of competition have reached their limit, when each of the partial productions Dk [the output of producer k] is inappreciable, not only with reference to the total production D = F(p), but also with reference to the derivative F(p), so that the partial production Dk could be subtracted from D without any appreciable variation resulting in the price of the commodity (Cournot, 1838 (1927 reprint), p. 90)».

mento dell'approccio normativo e antitrust; più recentemente, nella letteratura comportamentista si descrive come «behavioural failure» quella situazione in cui il comportamento reale degli individui non collimi con quello, teorico, che sarebbe adottato da un agente economico perfettamente razionale.

L'economia dei dati è terreno fertilissimo, e pericolosissimo, da questo punto di vista, poiché la suggestione secondo la quale la conoscenza può essere acquisita senza teorie e ipotesi, con algoritmi di intelligenza artificiale che scandagliano i dati per «trovare» nessi probabilistici utili a «predire» i comportamenti sociali «giusti» o «che funzionano», si presta facilmente ad essere fraintesa, cosicché qualcuno arriverà a dire che i dati sono rappresentazione più reale della realtà stessa. È la mitologia dei *big data* cui si faceva cenno in introduzione. Anderson (2008), il direttore di Wired, scrive: «This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves».

Insomma, potremmo essere alle porte dell'ennesima ondata (neo)positivista che rischierebbe di fornire nuova linfa alla retorica dell'oggettività tanto cara ad alcuni. Come sottolinea lucidamente Davies (2013), «[...] data is being icily naturalised, with its institutional and methodological preconditions being marginalised from discussion. Arguably this is a rhetorical ploy on the part of those who simply want to collect more data. The idea of an entirely neutral methodology, which allows data to "speak for itself", will be most appealing for those who don't wish to justify their scientific activities»; e poi prosegue: «apolitical, purposeless knowledge is often one of the greatest assets to the strategies of the powerful».

Sono sempre più numerosi i contributi scientifici di studiosi delle scienze sociali e di epistemologi che, pur riconoscendo la portata innovativa dei *big data* e la necessità di comprenderne l'impatto sulle metodologie delle scienze sociali, ne avversano la mitizzazione. Il comune denominatore di queste posizioni è che i grandi dati *sono* teoria: sono artefatti socio-culturali, tutt'altro che neutrali (Crawford *et al.*, 2014; Lombi, 2015) [4]. Sono poi evidenziati possibili preoccupazioni relative alle attività di ricerca basata su *big data* (Boyd, Crawford 2012; Crawford, 2013; Kitchin, 2013; 2014; Gitelman, 2013; Busch, 2014; Lombi, 2015; Mittelstadt, Floridi 2016) in termini di: rappre-

[4] Come ben sottolineato dalla Lombi (2015, p. 218): «va riconosciuto come i Big Data siano artefatti socio-culturali della società digitale globalizzata, la cui produzione e fruizione subisce l'influenza di processi politici, sociali e culturali. Oltre a non poterli considerare oggettivi, dobbiamo considerare come non siano neppure neutrali».

sentatività – poiché essi non tengono in considerazione coloro la cui partecipazione alla società avvenga senza la registrazione di segnali digitali (e qui tornano i problemi, specialmente italiani, del *digital divide* e della scarsa attitudine ad utilizzare strumenti che lasciano tracce digitali, come le carte di credito ad esempio) – con evidenti problemi di distorsione da selezione campionaria; validità e attendibilità, poiché senza una loro contestualizzazione potrebbero essere soggetti a fraintendimenti (commenti ironici sui *social network* interpretati come seri, profili falsi, eccetera) o comunque non correttamente considerati [5].

Quand'anche tali criticità fossero affrontate e risolte, resta una questione di fondo legata alla trasparenza dei processi con cui gli operatori dei *big data* derivano la conoscenza dai dati, in particolare la non anticipabilità e non intuibilità dei risultati delle loro analisi (ciò che Andrejevic, 2014, chiama *big data divide*). I consumatori, non necessariamente per pigrizia o ignoranza nell'uso delle tecnologie, potrebbero faticare a immaginare che, ad esempio, l'essere tifoso di una specifica squadra di calcio potrebbe aumentare o ridurre la loro appetibilità nel mondo del lavoro o comportare condizioni diverse nella stipula di un contratto di assicurazione (Zeno-Zencovich, 2018): in generale, è impossibile che un individuo concepisca tutti i possibili punti di caduta generati dall'interazione complessa di migliaia di variabili relative a milioni di consumatori. Un mondo in cui gli individui sono sottoposti a sollecitazioni, ad esempio pubblicitarie, concepite a partire da analisi predittive non trasparenti sul comportamento di un consumatore medio derivato dai loro dati personali è un mondo in cui la compatibilità tra crescita economica e rispetto dei valori e delle aspettative delle collettività è tutt'altro che preservato.

L'economia e la società dei *big data* portano con sé molte opportunità e altrettante sfide. L'auspicio è che, tra queste ultime, non vi sia quella di dimostrare di essere all'altezza dei propri dati!

[5] Ancora la Lombi (2015, p. 221): «Secondo Kitchin (2014), la sfida nell'analizzare i Big Data dipende dal fatto che è necessario affrontare la loro abbondanza, esaustività, varietà, permanenza, dinamicità, disordine e incertezza, elevata relazionalità, senza essere guidati da domande specifiche. Con le parole dell'Autore: "L'analisi fondata sui BD consente l'emergere di un approccio epistemologico completamente nuovo attraverso cui dar senso al mondo; anziché testare una teoria attraverso dati ritenuti rilevanti, le analisi dei dati cercano di fondare prospettive 'nate dai dati'" (ibid.: 2). Il rischio di questo approccio è quello di lavorare su dati superficiali, senza consapevolezza alcuna del contesto da cui sono tratti e degli effetti che la piattaforma che li ospita esercita sulla loro natura e struttura, affidandoci soltanto al potere degli algoritmi matematici che, tuttavia, abbiamo visto essere artefatti sociali».

## Riferimenti bibliografici


Anderson C. (2008), *The End of Theory, Will the Data Deluge Makes the Scientific Method Obsolete?*, in «Edge», disponibile al link: http://www.edge.org/3rd_culture/anderson08/anderson08_index.html.

Andrejevic M. (2014), *The Big Data Divide*, in «International Journal of Communication», 8, pp. 1673-1689.

Andreoni A. (2017), *Strategies for Emerging Technologies and Strategic Sectors: Evidence from OECD Countries and Some Critical Reflections on the Italian Case*, in «l'industria», XXXVIII, 1, pp. 3-14.

Bagnoli V. (2016), *The Big Data Relevant Market*, in «Concorrenza e Mercato», 23, numero speciale, pp. 73-94.

Bernaert M., Poels G., Snoeck M., De Backer M. (2014), *Enterprise Architecture for Small and Medium-Sized Enterprises: A Starting Point for Bringing EA to SMES, Based on Adoption Models*, in Devos J., van Landeghem H., Deschoolmeester D. (a cura di), *Information Systems for Small and Medium-Sized Enterprises*, Berlin Heidelberg, Springer-Verlag.

Bourreau M., de Streel A., Graef I. (2017), *Big Data and Competition Policy: Market Power, Personalised Pricing and Advertising*, Bruxelles, Centre on Regulation in Europe.

Boyd D., Crawford K. (2012), *Critical Questions For Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon*, in «Information Communication and Society», 15, 5, pp. 662-679.

Busch L. (2014), *Big Data, Big Questions. A Dozen Ways to Get Lost in Translation: Inherent Challenges in Large Scale Data Sets*, in «International Journal of Communication», 8, pp. 1727-1744.

Buzzacchi C. (2016), *La politica europea per i big data e la logica del single market: prospettive di maggiore concorrenza?*, in «Concorrenza e Mercato», 23, numero speciale, pp. 153-180.

Cassetta E., Meleo L., Pini M. (2016), *Il ruolo della digitalizzazione nel processo di internazionalizzazione delle imprese*, in «L'industria, Rivista di economia e politica industriale», XXXVII, 2, pp. 305-328.

Cassetta E., Pozzi C. (2016), *Politiche industriali, selettività e mercato*, «L'industria, Rivista di economia e politica industriale», XXXVII, 4, pp. 531-536.

Cournot A. (1838), *Researches into the Mathematical Principles of the Theory of Wealth*, New York, Macmillan (1927 reprint).

Crawford K. (2013), *The Hidden Biases in Big Data*, in «HBR Blog Network».

Crawford K., Miltner K., Gray M.L. (2014), *Critiquing Big Data: Politics, Ethics, Epistemology*, in «International Journal of Communication», 8, pp. 1663-1672.

Davies W. (2013), *Data's Double Life*, in «Potlatch», disponibile al link: http://potlatch.typepad.com/weblog/2013/06/datas-double-life.html.

Di Porto F. (2016), *La rivoluzione big data. Un'introduzione*, in «Concorrenza e Mercato», 23, numero speciale, pp. 5-14.

Di Tommaso M., Barbieri E. (2017), *Investimenti, innovazione e politiche per lo sviluppo industriale: quo vadimus?*, in Baravelli M., Bellandi M., Camagni R., Capasso S., Cappellin R., Ciciotti E., Marelli E. (a cura di), *Investimenti, innovazione e nuove strategie di impresa: Quale ruolo per la nuova politica industriale e regionale?*, Milano, EGEA.

Evans D.S. (2011), *Antitrust Economics of Free*, in «Competition Policy International», Spring, pp. 1-26.

Evans D.S. (2012), *Two-Sided Markets*, in *ABA Section of Antitrust Law, Market Definition in Antitrust: Theory and Case Studies*, Chicago, American Bar Association.

Ezrachi A., Stucke M.E. (2017), *Artificial Intelligence & Collusion: When Computers Inhibit Competition*, in «University of Illinois Law Review», 5, pp. 1775-1810.
Gitelman L. (2013), *«Raw Data» Is an Oxymoron*, Cambridge, MIT Press.
Graham M. (2012), *Big Data and the End of Theory?*, in «The Guardian», disponibile al link: http://www.theguardian.com/news/datablog/2012/mar/09/big-data-theory.
ITMEDIA (2018), *L'economia dei dati. Tendenze di mercato e prospettive di policy*, Roma, ITMEDIA Consulting.
Kawsar F., Brush A.J.B. (2013), *Home Computing Unplugged: Why, Where and When People Use Different Connected Devices at Home*, in «Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing», pp. 627-636.
Kitchin R. (2014), *Big Data, New Epistemologies and Paradigm Shifts*, in «Big Data & Society», Aprile-Giugno, pp. 1-12.
Lambrecht A., Tucker C.E. (2017), *Can Big Data Protect a Firm from Competition?*, in «CPI Antitrust Chronicle», gennaio, pp. 1-8.
Lerner A.V. (2014), *The Role of «Big Data» in Online Platform Competition*, SSRN Papers, 26 Agosto, pp. 1-63.
Lombi L. (2015), *La ricerca sociale al tempo dei Big Data: sfide e prospettive*, in «Studi di Sociologia», 2, pp. 215-227.
Losurdo F., Marra A., Cassetta E., Monarca U., Dileo I., Carlei V. (2018), *Emerging Specializations, Competences and Firms' Proximity in Digital Industries: The Case of London*, in «Papers in Regional Science», Early View, 1-17.
Maggiolino M. (2016), *Big data e prezzi personalizzati*, in «Concorrenza e Mercato», 23, numero speciale, pp. 95-138.
Mittelstadt B.D., Floridi L. (2016), *The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts*, in «Science and Engineering Ethics», 22, 2, pp. 303-341.
Muscio A., Quaglione D., Ramaciotti L. (2016), *The Effects of University Rules on Spinoff Creation: The Case of Academia In Italy*, in «Research Policy», 45, 7, pp. 1386-1396.
Muscio A., Quaglione D., Vallanti G. (2015), *University Regulation and University-Industry Interaction: A Performance Analysis of Italian Academic Departments*, in «Industrial and Corporate Change», 24, 5, pp. 1047-1079.
Pitruzzella G. (2016), *Big Data, Competition and Privacy: A Look from the Antitrust Perspective*, in «Concorrenza e Mercato», 23, numero speciale, pp. 15-28.
Pontarollo E. (2016), *Industria 4.0: un nuovo approccio alla politica industriale*, in «l'industria», XXXVII, 3, pp. 375-381.
Pontarollo E. (2017), *L'industria verso il futuro*, in «L'industria, Rivista di economia e politica industriale», XXXVIII, 4, pp. 413-429.
Pozzi C. (2014), *Crisi, teorie e politiche: il coraggio di cambiare*, in «L'industria, Rivista di economia e politica industriale», XXXV, 2, pp. 173-186.
Quaglione D. (2017), *Exploring Additional Determinants of Fixed Broadband Adoption: Policy Implications for Narrowing the Broadband Demand Gap*, in «Economics of Innovation and New Technology», 27, 4, pp. 307-327.
Sarra A., Di Berardino C., Quaglione D. (2018), *Deindustrialization and the Technological Intensity of Manufacturing Subsystems in the European Union*, in «Economia Politica», online first, pp. 1-39.
Schepp N., Wambach A. (2016), *On Big Data and its Relevance for Market Power Assessment*, in « Journal of European Competition Law & Practice», 7, 2, pp. 120-124.
Sokol D., Comerford R. (2016), *Antitrust and Regulating Big Data*, in «George Mason Law Review», 119, pp. 1129-1161.
Stucke M.E., Grunes A.P. (2015), *Debunking the Myths Over Big Data and Antitrust*, in «CPI Antitrust Chronicle», Maggio, pp. 1-10.

Traversi V., Ruggieri R. (2013), *L'Enterprise Architecture e la gestione del patrimonio informativo aziendale*, in «Economia e diritto del terziario», 1, pp. 49-63.
Tucker D.S., Wellford H.B. (2014), *Big Mistakes Regarding Big Data*, The Antitrust Source, Dicembre, pp. 1-12.
Zeno-Zencovich V. (2018), *Dati, grandi dati, dati granulari e la nuova epistemologia del giurista*, in «Rivista di diritto dei media», 2, pp. 1-7.