



UNIVERSITA' DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI
"M.FANNO"

CORSO DI LAUREA IN ECONOMIA E MANAGEMENT

PROVA FINALE

"BIG DATA E CONCORRENZA"

RELATORE:

CH.MO PROF. FABIO MANENTI

LAUREANDO: ALESSANDRO DALLE FRATTE

MATRICOLA N. 1022019

ANNO ACCADEMICO 2016 – 2017

SOMMARIO

1 - INTRODUZIONE -----	2
2 - I BIG DATA -----	5
2.1 – Definizione e introduzione ai Big Data -----	5
2.2 – Strutture tecnologiche e tipologie di dati trattati -----	9
2.3 – Gli strumenti strategici: Algoritmi e loro finalità -----	11
3 – ASPETTI ECONOMICI CORRELATI AI BIG DATA -----	14
3.1 – I meccanismi di base: il concetto di reti, esternalità -----	14
3.2 – Analisi delle dinamiche economiche introdotte dai Big Data -----	19
3.3 – Vantaggio competitivo derivante dall'utilizzo dei Big Data -----	21
4 – POTENZIALI EFFETTI NEGATIVI SUL MERCATO -----	27
4.1 – Algoritmi e pratiche collusive -----	27
4.2 – Altri scenari anticompetitivi -----	38
5 - CONCLUSIONI -----	42
RIFERIMENTI BIBLIOGRAFICI -----	45

1 - INTRODUZIONE

La copertina di “The Economist” in data 12 Maggio 2017 titolava a piena pagina:
“La risorsa più preziosa al mondo”.



Figura 1 - Copertina di “The Economist”

Sotto al titolo campeggiavano delle piattaforme petrolifere identificate dai nomi “Google”, “Facebook”, “Amazon”, “Tesla”, “Uber”. Tale copertina condensa in una semplice immagine le eccezionali aspettative correlate alla *data economy* nel prossimo futuro.

L'economia del secolo scorso aveva individuato nel petrolio il catalizzatore ottimale per sostenere una crescita economica rilevante e duratura. Gli effetti negativi sull'ambiente, il calo delle scorte ed il conseguente aumento dei costi di estrazione ne hanno fatto rivalutare il ruolo centrale all'interno delle moderne economie occidentali, incentivando la ricerca di nuovi fattori su cui impennare lo sviluppo dei paesi industrializzati negli anni a venire.

Le energie alternative hanno dato notevoli spunti di crescita nel decennio in corso, ma i loro limiti sono già evidenti. Il miglior sostituto del petrolio - l'energia da fusione termonucleare - si trova ancora in uno stadio precoce di studio e, per il suo pieno impiego, si dovrà pazientare ancora per diversi decenni.

La nascita dell'economia digitale maturata alla fine del secolo scorso, la crescita esponenziale delle capacità di calcolo dei microprocessori, la connettività dei dispositivi elettronici e l'evoluzione delle tecnologie di trasmissione dati hanno posto le condizioni ideali per la candidatura dei Big Data quale migliore risposta attuale alle ricerche in corso.

Gli innumerevoli benefici collegati allo studio di banche dati, in anni recenti sta indirizzando una crescente quantità di grandi aziende (soprattutto private) verso investimenti in soluzioni di *Business Analytics* e Big Data.

Attorno al tema dei Big Data persevera tuttavia un alone di mistero, ancora poche sono le imprese effettivamente in grado di lavorare con essi in modo da trarne un valore aggiunto traducibile in profitto, aspetto particolarmente evidente nel contesto italiano.

Eppure il concetto di Big Data ha una storia molto antica che precede l'esistenza di qualsiasi computer o software. Sembra infatti che già nel 600 a.C. l'antico filosofo greco Talete di Mileto fosse dedito archiviare grosse quantità di dati, anche non strettamente connessi fra loro, per farne un uso predittivo e, soprattutto, per trarre un guadagno da questi.

Superando l'uso di storie e miti legati al mondo delle divinità per interpretare e spiegare i fenomeni naturali, il filosofo greco raccolse negli anni ingenti quantità di dati che utilizzò come base per sue minuziose analisi finalizzate alle previsioni meteorologiche. Talete, che oggi potremmo definire come un "*data scientist a.C.*", elaborò una sorta di database nel quale venivano raccolti negli anni tutti i dati utili relativi alle condizioni climatiche, assieme a tanti altri eventi che apparivano del tutto slegati da queste. Grazie alla profonda conoscenza degli astri, il filosofo riuscì ad analizzare ed interpretare i dati per trarre delle informazioni che gli fruttarono un enorme guadagno, con stupore dei suoi concittadini che lo deridevano per l'inutilità dei suoi studi. Talete fu infatti in grado di prevedere con anticipo, quando era ancora pieno inverno, l'avvento di un'abbondante raccolta di olive. Ciò gli permise di assicurarsi l'affitto dei frantoi di un'intera regione, investendovi una piccolissima somma di denaro. Un investimento dal quale ottenne un grande ritorno poiché, quando effettivamente il momento del grande raccolto arrivò, ottenne il monopolio della molitura e la possibilità di subaffittare i frantoi ad alti prezzi, facendo così la sua fortuna e soprattutto dimostrando che la ricchezza più grande deriva sempre dalla conoscenza.

Tralasciando le lecite considerazioni circa l'attendibilità storica di tale aneddoto, esso rappresenta con estrema precisione i meccanismi ed i benefici peculiari collegati al fenomeno odierno dei Big Data.

In passato il tema dei Big Data era tradizionalmente confinato nell'ambito della ricerca e della scienza che, per evidente loro necessità, dovevano gestire grandi database. In un periodo in cui l'economia è in evidente difficoltà, la possibilità di prendere decisioni basandosi su informazioni e analisi aggiornate è estremamente importante. Le decisioni gestionali basate unicamente su intuito ed esperienza sono sempre più spesso considerate poco affidabili, mentre le decisioni aziendali sono sempre più collegate a concrete informazioni analitiche.

Partendo da tale assunto, il presente elaborato si prefigge di introdurre il lettore allo strumento dei Big Data, fornendo alcune nozioni di base necessarie per comprenderne l'architettura e le finalità. Si accennerà alle straordinarie potenzialità di questa tecnologia moderna, delle sue forme di interazione con il mercato e dei meccanismi economici coinvolti, menzionando i principali benefici ottenibili dal suo utilizzo. Data la recente diffusione, non essendo ancora accertate le eventuali criticità correlate all'impiego di queste macchine, si cercherà di dare menzione circa i rischi ed i possibili effetti collaterali a discapito dei consumatori e degli equilibri di mercato nelle economie interessate. Quanto verrà riportato nei capitoli a seguire non si prefigge di indurre il lettore ad un giudizio pro o contro i Big Data; più modestamente, l'intento è quello di trasmettere alcune nozioni utili per comprendere la rilevanza attuale di tale fenomeno ed agevolare un approccio razionale ai temi correlati.

Il capitolo a seguire affronta in sintesi il concetto di base di Big Data, delle loro peculiarità rispetto alle tradizionali banche dati, accennando alle strutture ed agli strumenti tecnologici che li contraddistinguono.

Senza particolari soluzioni di continuità il terzo capitolo si addentra all'interno dei meccanismi economici attivati (generici ed esclusivi), affrontando le loro implicazioni nel mercato ed i vantaggi competitivi collegati.

Come avviene per qualsiasi strumento di provata utilità, un uso improprio dello stesso può condurre ad effetti indesiderati: il quarto capitolo li individua nei possibili rischi di collusione, cartelli dei prezzi, abuso di posizione dominante facilitati dallo sfruttamento dei Big Data e dei suoi algoritmi.

2 - I BIG DATA

2.1 - Definizione e introduzione ai Big Data

Nel corso del 2014 Facebook perfezionò l'acquisizione di WhatsApp. La Commissione Europea non ne bloccò la fusione, motivando che sebbene entrambe le società stessero gestendo i due più grandi servizi di messaggistica, esistevano molti altri operatori alternativi e la fusione non avrebbe portato a Facebook quel gran tesoro di dati in quanto WhatsApp non raccoglieva molte informazioni sugli utenti. E' vero che Facebook acquistava una potenziale azienda rivale, ma l'esorbitante prezzo pagato lasciava ben dubitare circa le reali finalità dell'operazione. Durante la fase di approvazione della fusione, Facebook assicurò che non avrebbe congiunto le due banche dati relative ai dati degli utenti, ma contrariamente a quanto promesso cominciò a farlo poco tempo dopo portando la Commissione a condannare l'azione con l'applicazione di multe (110 milioni di Euro).

In considerazione della loro popolarità nel mondo, l'operazione di mercato tra Facebook e WhatsApp ebbe grande rilevanza nelle pagine dei giornali, guadagnando l'attenzione anche di quelli meno attenti alle dinamiche economiche. L'eccezionalità del caso risiedeva nell'elevato prezzo corrisposto per l'acquisto di una società (WhatsApp) a fronte del suo esile valore di mercato. Molte persone nell'apprendere la notizia si posero il quesito: cosa sono e quanto valgono i dati?

La comune definizione di dato riportata nei dizionari è quella di "informazione fattuale (da misura o statistica) utilizzata come base per ragionamenti, discussioni o calcoli". I dati possono quindi essere qualitativi o quantitativi e, come la definizione suggerisce, l'importanza dei dati deriva dal loro uso da parte degli individui, aziende e governi per supportare i loro processi decisionali.

Un articolo del prestigioso giornale "The Economist" datato 6 maggio 2017 titolava a piena pagina: *"La risorsa più preziosa al mondo non è più il petrolio ma i dati"*. In effetti, similmente a quanto avvenne nel secolo scorso con il petrolio, l'utilizzo dei dati offre oggi forti potenzialità di cambiamento, crescita e sviluppo dell'economia moderna. I dati però differiscono da qualsiasi altra risorsa del passato: differiscono per le modalità di estrazione e raffinazione, per il loro processo di valorizzazione, per i canali di acquisto e vendita.

La crescita del volume di dati immagazzinati nel mondo è oramai tale da richiedere il conio di nuove unità di misura, risultando il semplice "byte" e suoi multipli inadeguato ai tempi. IDC, reputata società di ricerche di mercato, ha previsto una esponenziale crescita del volume di

dati stoccati nel mondo, prevedendo il passaggio dagli attuali 8 Zbytes a 45 Zbytes nel 2020, fino a raggiungere quota 180 Zbytes nel 2025 (Zbyte=10²¹ Bytes).

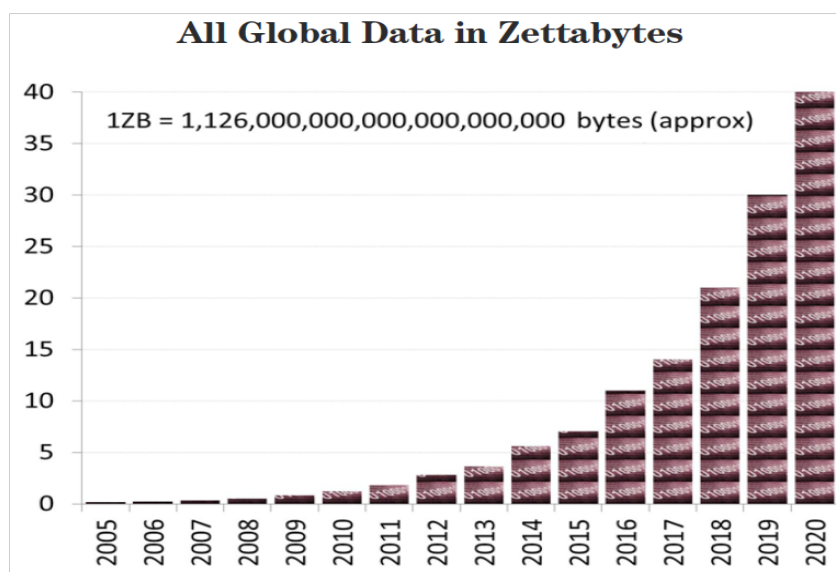


Figura 2 - Volumi di dati immagazzinati nel mondo
(Fonte: Commissione Economica delle Nazioni Unite)

Il grafico sopra fornisce un'immediata comprensione dello straordinario incremento dei volumi di dati e la tendenza prospettica per il futuro.

I numeri menzionati giustificano ampiamente il nome dato a questo recente fenomeno che si avvale di innumerevoli data center dislocati in varie parti del mondo, i quali si configurano come delle vere e proprie “centrali di stoccaggio” di volumi imponderabili di dati: sono i Big Data.

Nel corso degli ultimi anni, la comunità scientifica ha cercato più volte di dare una identità univoca al termine “Big Data”. Di seguito alcune tra le più comuni definizioni (De Mauro, Greco, Grimaldi, 2015. *A formal definition of Big Data based on its essential features*):

- “un pacchetto di dati che non può essere gestito e processato in maniera usuale”
- “volume di dati che eccede le capacità di processo dei convenzionali database”
- “grande quantità di dati, complessa, non strutturata”
- “applicazioni di pacchetti di dati e tecniche analitiche che sono così vaste e complesse da richiedere tecnologie avanzate ed uniche per lo stoccaggio, gestione, analisi e visualizzazione”

Nonostante gli innumerevoli tentativi, non esiste ancora una definizione ufficiale del termine “Big Data”. Tuttavia una particolare definizione teorizzata dal Prof. Andrea De Mauro ha

raccolto il favore della maggioranza degli studiosi del settore, tale da essere comunemente accettata per individuare il fenomeno:

“Big Data rappresenta un pacchetto di informazioni caratterizzato da tale dimensione, velocità e varietà da richiedere specifiche tecnologie e metodi analitici per la trasformazione in valore”.

Le quattro peculiarità distintive dei Big Data sono quindi identificate da:

- *Volume* (dimensione): l'enorme mole di dati è caratteristica primaria che non richiede spiegazioni aggiuntive
- *Velocity* (velocità): la velocità di processo dei dati è diventata di fondamentale importanza per condurre analisi in tempo reale, consentendo risultati previsionali rilevanti per le imprese (es. analisi di mercato) o la società (es. previsioni meteo)
- *Variety* (varietà): qualsiasi tipo di dato (non solo i prezzi) è divenuto rilevante per il raggiungimento degli obiettivi (età, sesso, tipo di impiego, preferenze musicali...)
- *Value* (valore): le caratteristiche sopra sarebbero di poco conto se non vi fosse il modo di ricavare del valore dai dati. La peculiarità dei Big Data è proprio legata alla capacità di estrarre valore dalle informazioni raccolte

comunemente note nell'ambiente come le quattro “V”.

Recentemente, mantenendo lo stesso principio mnemonico, è stata associata una quinta “V” per individuare la veridicità (*Veridicity*) dei dati quale fattore rilevante, fattore che è soggetto a variazioni piuttosto ampie in ragione della tipologia di dati trattati, influenzando sull'attendibilità dei risultati delle elaborazioni.

All'interno della catena del valore dei Big Data si possono distinguere tre fasi principali:

- a) la raccolta dei dati (dal Web, dai dispositivi elettronici...)
- b) lo stoccaggio dei dati (hardware e software che gestiscono i data center)
- c) il processo dei dati (gli algoritmi)

Come ben rappresentato nella figura 3, la raccolta e lo stoccaggio dei dati è comunemente indicata come *“Big Data Management”*, la loro elaborazione tramite algoritmi è nota come *“Big Data Analytics”*.

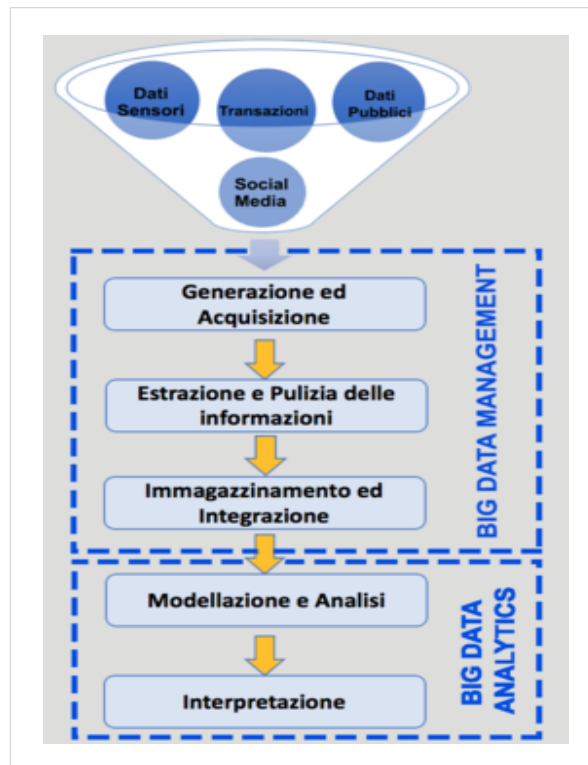


Figura 3 - Ciclo di vita dei Big Data (da Wikipedia)

La crescente maturità del concetto di Big Data consente di evidenziare le differenze con le attività note come *Business Intelligence*, facenti capo a strutture apparentemente simili ma sostanzialmente diverse nei seguenti aspetti (fonte Wikipedia):

- *Business Intelligence* utilizza la statistica descrittiva con dati ad alta densità di informazioni per misurare cose, rilevare tendenze, ossia utilizza *dataset* limitati, dati puliti e modelli semplici
- Big Data utilizza la statistica inferenziale e concetti di identificazione di sistemi non lineari, per dedurre leggi (regressioni, relazioni non lineari ed effetti causali) da grandi insiemi di dati e per rivelare i rapporti, le dipendenze ed effettuare previsioni di risultati e comportamenti, cioè utilizza *dataset* eterogenei (non correlati tra loro), dati grezzi e modelli predittivi complessi

Posto in termini aforistici, la *Business Intelligence* aiuta a trovare le risposte a domande che si conoscono, mentre i Big Data aiutano a trovare le domande che non si sa di voler chiedere (nel senso che rivelano aspetti mai considerati prima).

2.2 - Strutture tecnologiche e tipologie di dati trattati

Una delle principali cause all'origine del fenomeno dei Big Data fu la tendenza alla digitalizzazione di massa, ossia il tentativo di convertire i libri in files con l'ausilio della tecnologia di riconoscimento ottico dei caratteri (OCR). Google Print Library Project iniziò nel 2004 questo ambizioso progetto, con l'obiettivo di archiviare 15 milioni di volumi custoditi nelle biblioteche delle più prestigiose università al mondo.

I risultati positivi riscontrati con la digitalizzazione incoraggiarono il passaggio alla fase successiva, quella dell'archiviazione di massa di dati, ossia la conversione di fenomeni generici in dati di formato idoneo ad essere manipolato dai calcolatori.

L'archiviazione di massa è diventata un fenomeno pervasivo grazie alla capillare diffusione dei molteplici dispositivi elettronici di comune utilizzo, il reperimento dei dati è divenuto un'operazione di estrema semplicità e a basso costo. Le sorgenti di dati più conosciute sono i POS, le operazioni effettuate sul web, i sensori di ultima generazione (installati in casa, nell'auto...), i telefoni cellulari e tutte le altre comuni applicazioni tecnologiche. Cisco stimò nel 2009 che il numero di dispositivi connessi era già superiore al numero di persone viventi; Gartner stima oggi che entro il 2020 ci saranno sulla Terra circa 26 miliardi di dispositivi connessi.

“Dalle metropolitane alle turbine eoliche fino alle toilette ed ai tostapane - ogni sorta di dispositivo sta diventando sorgente di dati. Il mondo si sveglierà con sensori connessi ovunque, le persone lasceranno una traccia digitale ovunque andranno, anche se non connesse a internet. [...] I dati saranno l'esternalità di base: li genereremo qualsiasi cosa facciamo” (The Economist, 6 maggio 2017).

Facebook e Google inizialmente utilizzavano i dati raccolti dagli utenti per indirizzare la pubblicità. In tempi recenti gli stessi hanno scoperto che i dati possono essere convertiti in numeri al servizio di intelligenza artificiale e servizi cognitivi, tramutandosi in nuove fonti di guadagno. Sebbene i segnali della *data economy* siano ovunque, solo ora la sua forma sta diventando manifesta. Tutti tentano ora di sfruttare un potente motore economico chiamato *data network effect*, utilizzando i dati per attrarre altri utenti che genereranno nuovi dati che faranno conseguire un migliore servizio e che quindi attrarranno ulteriori utenti. Più persone utilizzano Google per ricerche, migliore sarà il servizio di ricerca fornito. Uber ha raggiunto il valore stimato di 68 miliardi di dollari avvantaggiandosi del più grande *database* di autisti e passeggeri nell'ambito del trasporto individuale. Similmente, l'enorme volume di dati raccolti dalle auto prodotte, permette a Tesla di ottimizzare l'algoritmo di guida automatica ed il

conseguente aggiornamento del software, il vero valore della sua tecnologia.

Dati i costi d'investimento estremamente elevati, è giusto dire che i Big Data sono uno strumento al solo servizio di poche aziende facoltose? Cosa dovrebbero fare le imprese che intendono avvalersi dei vantaggi della *data economy* ma non sono nella condizione di immagazzinare e processare un traffico dati di petabytes? Alcuni provider di infrastrutture tecnologiche informatiche quali ad esempio Hadoop, IBM e Oracle hanno ben visto di sviluppare una remunerativa attività in risposta a tale esigenza. In tal caso i *provider* non hanno il solo onere di sviluppare il software adeguato ai Big Data, ma, quale fatto più rilevante, forniscono in veste di terza parte il servizio di *cloud computing*, ossia lo stoccaggio dei dati per conto delle imprese che lo richiedono. Date le importanti economie di scala, questo comporta ampi benefici per il fornitore e l'utilizzatore del servizio.

Il *cloud computing* ha il pregio di trasformare un costo fisso, che solo pochi potrebbero affrontare, in un costo variabile accessibile ad una moltitudine di imprese. Società come Amazon, Google o Microsoft forniscono algoritmi intelligenti come parte integrante del loro servizio e le piccole imprese trovano sempre più conveniente far processare i loro dati da queste strutture esterne.

In questo paragrafo si è indistintamente accennato ai dati, evidenziando che, nonostante le loro molteplici origini, tutti sono più o meno idonei ad essere elaborati per ricavarne delle informazioni. Tra le possibili tipologie di dati, una su tutte si distingue dalle altre in quanto le leggi nazionali ne fanno espresso oggetto di tutela: quella dei dati personali.

Data la loro peculiarità, è utile fornire qualche informazione sulla loro natura e gli usi consentiti: un dato si definisce personale quando permette di identificare un individuo attraverso dei parametri definiti. Il loro utilizzo per finalità amministrative o commerciali deve normalmente essere autorizzato dalla persona a cui sono riferiti. La tutela dei dati personali rappresenta ad oggi una delle poche (se non l'unica) limitazioni imposte ai gestori di Big Data verso le forme di dati raccolti. Un'accurata custodia dei dati personali viene percepita dai consumatori come un fattore di qualità e, come tale, esso potrebbe indirizzarli verso quei fornitori di servizi che più ne garantiscono la riservatezza rispetto ad altri, con conseguenti ricadute di carattere economico.

2.3 – Gli strumenti strategici: Algoritmi e loro finalità

Quando si parla di Big Data, si individuano generalmente tre elementi peculiari: i dati (la materia prima), la struttura hardware (la macchina) e gli algoritmi contenuti nel software (l'intelligenza che governa la macchina).

Hal Varian, Chief Economist di Google, partendo dall'assunto che i dati presentano economie di scala decrescenti, ossia ogni dato addizionale possiede un valore minore del precedente, osserva che da un certo punto in poi la raccolta di ulteriori dati non genererà benefici addizionali. *“Quello che più conta” - sostiene lo stesso - “è la qualità degli algoritmi che processano i dati e i talenti che l'azienda ha ingaggiato per svilupparli. Il successo di Google riguarda la ricetta, non gli ingredienti”.*

A tale affermazione fa eco Glen Weyl, economista di Microsoft Research, il quale obietta che la tesi di Varian poteva ritenersi valida negli anni in cui debuttarono i motori di ricerca, ma non è più vera nel nuovo mondo dell'intelligenza artificiale. Weyl osserva che *“gli algoritmi sono sempre più basati sull'autoapprendimento: più sono i nuovi dati trattati, migliore sarà il risultato. Il ritorno marginale dei dati può in effetti aumentare al moltiplicarsi delle applicazioni. Quando una azienda ha raccolto abbastanza dati per poter offrire un servizio di informazioni sul traffico in tempo reale, altri dati non aggiungono valore. Ma se si persevera nel raccogliere dati, ad un certo punto sarà possibile offrire più servizi, come un sistema di navigazione”.*

Qualunque sia l'approccio nel valutare l'utilità di un numero crescente di dati, rimane il fatto che la maggior parte dei dati raccolti non è utile se osservata in sé: l'estrazione del valore intrinseco dei dati richiede strumenti e conoscenze particolari che sono caratteristiche distintive della dotazione delle aziende, frutto delle competenze e degli investimenti in hardware e software, primo tra tutti gli algoritmi di proprietà.

Sebbene il concetto di algoritmo esista da diversi decenni, una definizione univoca e consensuale non esiste ancora. Wilson e Keil (1999) lo definiscono come “una lista precisa e non ambigua di operazioni applicate meccanicamente e sistematicamente a un set di simboli o oggetti (numeri, pedine degli scacchi, ingredienti di un dolce...); lo stato iniziale dei simboli è l'input, lo stato finale è l'output”.

L'evoluzione delle scienze informatiche ha consentito di sviluppare degli algoritmi che possono condurre automaticamente processi di calcolo molto complessi, fino a giungere alla cosiddetta “intelligenza artificiale”. Per Intelligenza Artificiale (AI) si intende un vasto settore delle scienze informatiche che studia e sviluppa algoritmi che dotano le macchine della

capacità di autoapprendimento.

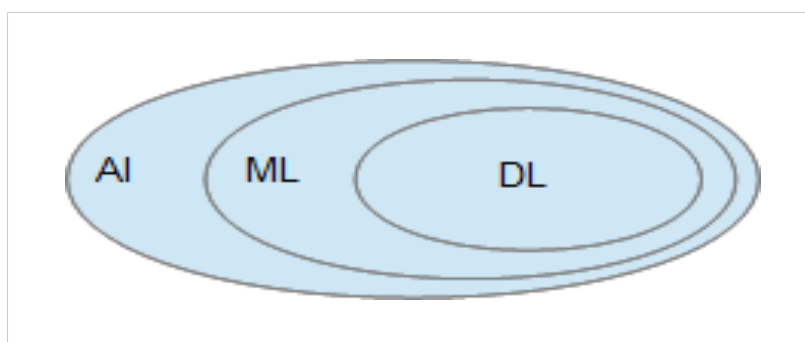


Figura 4 - Intelligenza Artificiale (AI)

Il *Machine Learning* (ML), sottoinsieme dell'AI, progetta macchine capaci di apprendere dalle operazioni svolte andando a correggere o perfezionare le procedure ricorrenti, una sorta di algoritmi “intelligenti” che copiano la caratteristica distintiva della mente animale e umana di apprendere dagli errori precedentemente commessi. Alla base di ciò vi è una minuziosa ed estensiva raccolta ed analisi dei dati che, opportunamente processati da sofisticatissimi algoritmi consentono un costante miglioramento delle prestazioni delle macchine.

L'ultima corrente evoluzione del ML è quella del *Deep Learning* (DL), una scienza che si propone di replicare tramite software le attività neuronali del cervello umano, modellando delle reti neuronali artificiali. Mentre nelle macchine tradizionali gli algoritmi di apprendimento sono lineari, gli algoritmi di *Deep Learning* sono invece strutturati in una gerarchia di crescente complessità ed astrazione, consentendo ai computer di apprendere con precisione e in tempi più rapidi.

Alle enormi potenzialità degli algoritmi di DL di risolvere i più variegati problemi complessi si contrappone una perdita di controllo da parte dell'uomo circa le sequenze logiche e i processi che hanno portato la macchina al risultato fornito. La macchina si sta in un certo modo emancipando dal suo creatore, rispondendo dei risultati forniti ma diventando sempre meno controllabile.

I Big Data sono strettamente correlati alla cosiddetta *Big Analytics* e al fenomeno del *Deep Learning*, tramite i quali i computer apprendono da soli come risolvere complessi problemi avvalendosi di dati e sofisticati algoritmi.

Nell'ambito economico finanziario gli algoritmi svolgono un ruolo fondamentale nell'analisi predittiva per misurare le probabilità di eventi futuri basandosi sui dati presenti. I modelli predittivi stimano la domanda futura, le variazioni dei prezzi, i mutamenti delle preferenze dei clienti, l'evoluzione dei rischi... Nel settore della finanza, reti neuronali ed algoritmi

“genetici” generano segnali di acquisto e di vendita dei titoli in portafoglio e predicono i fallimenti delle società quotate, dimostrando una accuratezza ed una facilità d'uso persino superiore alle tecniche tradizionali.

3 – ASPETTI ECONOMICI CORRELATI AI BIG DATA

3.1 – I meccanismi di base: il concetto di reti, esternalità

Nella usuale architettura dei Big Data, i dati sono raccolti, trasferiti e convertiti in valore economico secondo i meccanismi di un articolato ecosistema (inteso come associazione di elementi in un complesso organico e funzionale) interessato da interconnessioni multiple con diversi mercati. La figura 5 evidenzia la complessità delle strutture di rete e le loro interazioni all'interno dei sistemi dei Big Data.

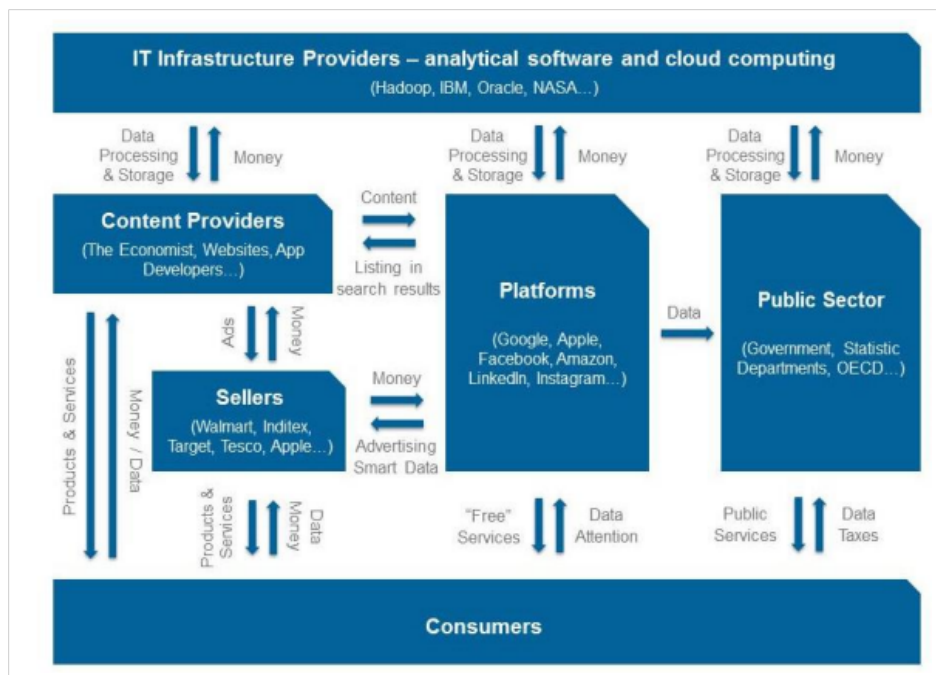


Figura 5 - Ecosistema tipico di Big Data (da OECD, 2016)

Per comprendere i meccanismi economici attivati dalle interazioni tra le parti individuate dall'ecosistema rappresentato sopra, è necessario partire dalla morfologia elementare delle reti, riprendendone le dinamiche correlate.

Il concetto di rete ha origine nell'ambito ingegneristico e definisce una struttura complessa per mezzo della quale più dispositivi interagiscono simultaneamente tra loro sulla base di una determinata tecnologia. L'interazione è resa possibile da canali che permettono di veicolare un particolare fenomeno fisico e da nodi ove suddetti canali si congiungono dando come effetto un fenomeno risultante di variata intensità rispetto a quelli originari. L'interazione ed i suoi effetti sono alla base del concetto di rete.

Mutuando la stessa logica, le scienze economiche hanno esteso l'applicazione del modello di

rete a tutti gli ambiti in cui l'interazione tra persone o cose abbia un impatto sui parametri originari con conseguenze di valenza economica.

Un fenomeno strettamente collegato alle reti sono le esternalità, definite come la condizione nella quale l'azione di un individuo consegue un impatto economico su un altro individuo (generalmente una terza parte) in assenza di una compensazione monetaria (ossia fuori dai meccanismi di mercato). Le esternalità possono avere carattere positivo (es. il livello di educazione, il rispetto delle regole) e negativo (es. l'inquinamento, il traffico urbano).

L'applicazione del concetto di esternalità alle reti porta a definire con un'accezione positiva il termine “esternalità di rete” quale “la situazione nella quale il beneficio individuale derivante dal consumo di un dato bene o servizio cresce all'aumentare del numero di altri utenti che consumano il medesimo bene o servizio”.

In una trattazione più estensiva, si fa riferimento agli “effetti di rete” per indicare il fatto che le decisioni degli utenti che afferiscono ad una medesima rete può avere un impatto diretto - normalmente in termine di benefici - sugli altri utenti. L'esempio tipico che tradizionalmente viene riportato per una immediata comprensione del concetto è quello del telefax: in origine un singolo apparecchio per inviare e ricevere fax non aveva alcuna utilità in sé; mano a mano che la tecnologia cominciò a diffondersi gli utenti che decidevano di acquistare un telefax potevano accedere al beneficio di poter comunicare tramite una nuova tecnologia con una moltitudine di soggetti e questo diveniva possibile per merito di altri utenti che li avevano preceduti nell'acquisto.

La ricerca degli effetti di rete fu un tema centrale già nel 1908 per Theodore Vail, primo presidente di Bell Telephone, il quale ne aveva ben intuito le potenzialità e li aveva inseguiti con tenacia. La sua strategia fu quella di agevolare quanto più possibile la crescita delle utenze telefoniche, fatto che permise a Bell di raggiungere prima di altri concorrenti un numero considerevole di clienti connessi, facendogli così guadagnare il pieno regime di monopolio del mercato.

La teoria economica degli effetti di rete venne ripresa e sviluppata significativamente nell'ultimo ventennio del secolo scorso, nuovi modelli utili ad interpretare le dinamiche causa/effetto presero forma dagli studi di valenti economisti quali Katz, Shapiro, Farrel e Saloner.

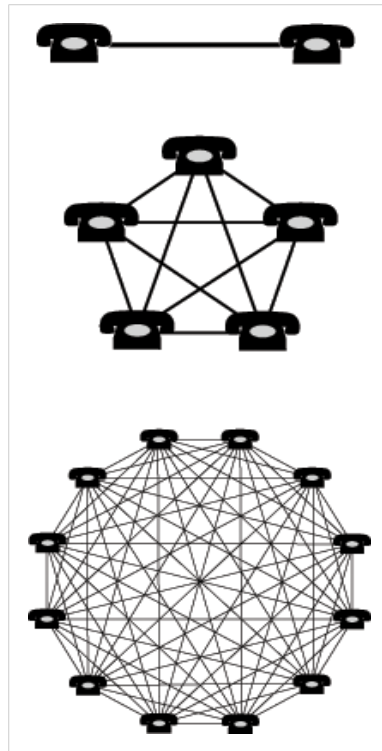


Figura 6 - Rappresentazione intuitiva degli effetti di rete su linee telefoniche (da Wikipedia)

Con l'avvento della *digital economy* gli effetti di rete assumono un ruolo fondamentale, essi si riscontrano ovunque la compatibilità di mezzi o tecnologie impiegati dagli utenti abbia un ruolo strategico all'interno delle dinamiche di comunicazione o scambio.

Quando una moltitudine di soggetti accomunati dall'utilizzo di determinati beni o servizi interagisce, tramite uno strumento intermediario, con un'altra moltitudine di soggetti fornitori di tali beni o servizi, si configura la tipica struttura di rete a due versanti (*two-sided network*). Nell'economia odierna lo strumento intermediario è generalmente una piattaforma digitale, ove il termine piattaforma definisce uno strumento che permette alle parti ad essa connesse di interagire e tale servizio viene prestato in termini onerosi in cambio di beni di varia natura. Nell'ambito delle reti a due versanti le azioni intraprese da un gruppo posto su un lato della piattaforma (es. utilizzatori) manifestano effetti sull'altro gruppo posizionato sul lato opposto (es. fornitori): si parla in tal caso di effetti di rete incrociati.

Le piattaforme a versanti multipli (*multi sided*) si sviluppano sul medesimo principio di quelle a due versanti, esse fungono da anello di congiunzione per una moltitudine di gruppi di soggetti che perseguono proprie finalità (diverse tra loro) all'interno delle loro rispettive attività correlate alla piattaforma. Il modello di piattaforma a versanti multipli rispecchia la struttura tipica dei sistemi che costituiscono i Big Data.

La struttura delle piattaforme è interessata da due distinte tipologie di effetti di rete:

- **effetti di rete diretti:** l'aumento del numero di utenti richiama altri nuovi utenti; è questo il tipico effetto rete che viene generato ed agisce all'interno del singolo lato della piattaforma (es. gli utenti di Facebook aumentano all'aumentare del numero di profili attivi)
- **effetti di rete indiretti:** nuovi utenti da un lato della piattaforma attirano nuovi utenti dall'altro lato della piattaforma; è l'effetto rete incrociato ove il numero di soggetti da un lato influisce sul numero di soggetti posti sul lato opposto (es. gli sviluppatori di videogames scelgono di operare su standard Playstation anziché Nintendo a seconda della loro diffusione)

Gli effetti di rete divengono significativi al raggiungimento di una certa quota di utenti, definita con il termine di “massa critica”. Quando il valore percepito dall'utente marginale aumenta al crescere del numero di utenti già fruitori del prodotto o servizio, in corrispondenza della massa critica il valore ottenuto uguaglia il costo pagato. Al superamento della massa critica, i benefici saranno sempre superiori al costo, provocando una crescita accelerata del numero di utenti (effetto valanga).

La ricerca della massa critica porta i gestori delle piattaforme ad incentivare nella fase iniziale l'incremento di utenti attraverso diverse forme, la più comune quella della prestazione di un servizio gratuito. Il costo di tale prestazione viene ampiamente recuperato da quanto corrisposto da altri utenti che sono disposti a pagare per intrattenere rapporti con tale moltitudine di controparti. L'esempio tipico è quello delle carte di credito, servizio gratuito per chi acquista e oneroso per chi vende: gli esercizi commerciali cercheranno sempre di convenzionarsi con l'emittente della carta (Visa, Mastercard...) che conterà il maggior numero di possessori.

L'ecosistema dei Big Data, del quale ne è stata raffigurata una possibile architettura in figura 5, è un ambiente a due o più versanti rappresentabile come un vettore di interazioni competitive tra i vari attori in gioco, alcuni dei quali con diversi ruoli. Per esempi Apple è simultaneamente una piattaforma (iOS, Apple Store e iTunes), un produttore/venditore di hardware (MacBooks, iPhones, iPads) e un *information technology provider* (iCloud), interagisce con i consumatori, con gli sviluppatori di contenuti, pubblicitari e altre piattaforme.

L'epicentro dell'ecosistema dei Big Data è costituito dalle piattaforme digitali, le quali operano come interfaccia principale tra i consumatori e gli altri operatori nel mercato.

Le piattaforme digitali si possono distinguere secondo le seguenti categorie (OECD, 2016) :

- *piattaforme di attenzione (attention platforms)*: esempio ne sono i motori di ricerca o i social network, i quali tipicamente forniscono una serie di servizi i cui costi sono coperti dalla pubblicità “per-click”. In questo modo, invece di pagare il prezzo monetario del servizio, i consumatori pagano con la loro attenzione verso la pubblicità presentata nelle forme note di banner, video e altro. I consumatori pagano inoltre inviando i loro dati, sia indirettamente attraverso la memorizzazione dei click online per ricerche o shopping, sia direttamente inserendo i propri dati personali all'interno dei moduli dedicati nelle pagine web
- *piattaforme di accostamento (matching platforms)*: forniscono uno spazio virtuale di mercato ove i diversi tipi di operatori possono interagire (es. acquirenti e venditori, chi cerca impiego e chi lo offre, siti di incontri). La piattaforma guadagna caricando un costo fisso di accesso ed un costo variabile per transazione. Frequentemente il gruppo di utenti con maggiore elasticità di domanda è sovvenzionato dall'altro gruppo di utenti sul lato opposto della piattaforma (es. i clienti non pagano per utilizzare i siti per acquisti, chi cerca lavoro non paga per accedere ai siti che offrono impiego)
- *piattaforme di transazione (transaction platforms)*: facilitano le transazioni o gli scambi tra un vasto numero di individui o organizzazioni (es. carte di credito)
- *piattaforme di innovazione (innovation platforms)*: offrono una risorsa tecnologica condivisa sopra la quale delle entità esterne possono sviluppare le loro applicazioni quali servizi complementari o prodotti. Questi innovatori complementari costituiscono il cosiddetto ecosistema attorno alla piattaforma (es. applicazioni per iPhone)

Quale risultato delle vigorose e plurime esternalità di rete dirette ed indirette, le piattaforme a versanti multipli tendono a portare alla concentrazione degli utenti e dei rispettivi dati nelle mani di pochi operatori. Google, ad esempio, potendo contare sulle attività del motore di ricerca (piattaforma di attenzione), di Google Shopping (piattaforma di accostamento), di Google Wallet (piattaforma di transazione), delle applicazioni di Android (piattaforma di innovazione) può avvalersi di un voluminoso e variegato bacino di utenti i quali, per convenienza o per mancanza di alternative, rimarranno fedeli al fornitore di servizi, incrementando così il suo bagaglio di dati e consentendogli di avvantaggiarsi dei rilevanti effetti di rete. L'uso dei Big Data consente alle piattaforme online di guadagnare un sostanziale potere di mercato anche nella diffusione di servizi di informazione (notizie, previsioni meteo, dati finanziari...), ai quali le imprese e i consumatori sempre più si affidano,

stimolando nuovi utenti ad usufruire del servizio. Nonostante questi servizi gratuiti traggano lucro dalla pubblicità e dai dati generati dalle attività degli utenti, il loro scopo primario è quello di fidelizzare questi ultimi alla piattaforma che eroga il servizio, incrementandone la sua popolarità. Questo contribuirà a sua volta a rinforzare gli effetti di rete.

I modelli di business accennati si dimostrano molto profittevoli, permettendo ad alcune piattaforme di guadagnare posizioni tra le prime dieci organizzazioni più capitalizzate al mondo.

Da quanto esposto si intuisce che gli effetti di rete sono una componente rilevante per gli operatori della *data economy*, il paragrafo a seguire ne esamina i meccanismi che li rendono così vigorosi e determinanti.

3.2 - Analisi delle dinamiche economiche introdotte dai Big Data

Le caratteristiche distintive dei dati analizzati sotto il profilo economico sono la loro presenza diffusa (si possono raccogliere ovunque), la non rivalità (il loro possesso da parte di un soggetto non ne preclude il possesso ad altri soggetti) e l'inconsistenza materiale.

I benefici derivanti dall'utilizzo dei Big Data sono individuabili nelle seguenti forme:

- incentivano l'innovazione
- incrementano l'efficienza
- consentono previsioni accurate
- agevolano un ritorno economico rapido (monetizzazione)
- consentono di operare una segmentazione del mercato

I dati hanno un costo molto limitato, all'aumentare del loro volume il costo marginale di produzione e distribuzione tende allo zero.

In aggiunta ai tradizionali fattori di consolidamento del potere di mercato (economie di scala, economie di scopo, effetti di rete), l'applicazione dei Big Data involve un nuovo meccanismo di rilevanza economica: quello dei benefici derivanti dallo sfruttamento delle informazioni estrapolate dai dati generati dagli utenti. Il meccanismo si basa su una logica a retroazione che permette di ottimizzare i prodotti o servizi offerti sulla base delle preferenze esplicitamente o intrinsecamente espresse dagli utenti attraverso comportamenti, scritture o altre forme di espressione registrate attraverso i dati. L'affinamento dei prodotti o servizi ai gusti o necessità degli utenti ne attrae altri che lasceranno a loro volta traccia delle loro aspettative circa i possibili sviluppi futuri.

Secondo le stime attuali l'utilizzo dei Big Data per finalità creative e innovative, processo noto come *Data-Driven Innovation* (DDI), premetterà alle imprese di migliorare la qualità dei loro prodotti e di svilupparli interamente supportati da una migliore comprensione delle esigenze dei loro clienti.

L'innovazione basata sull'utilizzo dei dati (DDI) già oggi ha fatto conseguire alle imprese che l'hanno adottata un'accelerazione della crescita dal 5% al 10% superiore rispetto a quella dei loro concorrenti (OECD, 2015). Quale risultato dei Big Data, l'economia europea crescerà di un addizionale 1,9% entro il 2020 (Buchholtz, 2014).

Diversamente dal modello di vendita al dettaglio “mattoni e cemento” (*brick and mortar*), i modelli di business moderni sono sempre più frequentemente caratterizzati dai *data-driven network effects* che permettono di sostenere una crescita basata sul miglioramento della qualità dei prodotti e dei servizi. Schematicamente i *data-driven network effects* sono il risultato congiunto di due anelli a retroazione (feedback) che interagiscono positivamente secondo le logiche rappresentate nella figura seguente:

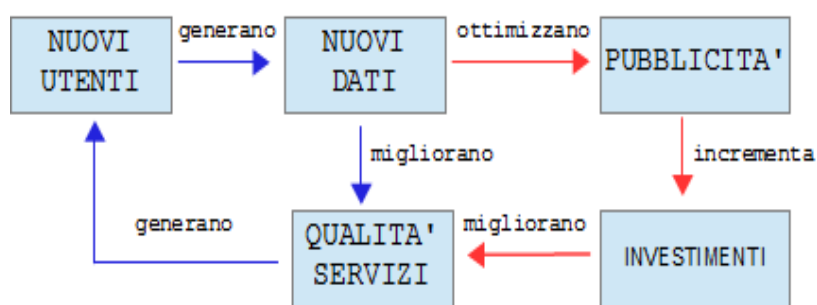


Figura 7 - Data-driven network effects (da OECD, 2016)

Gli effetti di quanto rappresentato sono evidenti: da una parte un'impresa con una larga base di utenti può raccogliere più dati per migliorare la qualità dei servizi prodotti, per esempio creando un migliore algoritmo, guadagnando così nuovi utenti (primo anello colore blu); dall'altra parte l'analisi dei nuovi dati generati dagli utenti consente alle imprese di migliorare l'efficienza della loro pubblicità, i maggiori guadagni genereranno nuovi investimenti che a loro volta faranno conseguire un migliore livello dei servizi offerti, attirando così nuovi utenti nel loro circuito (secondo anello colore rosso). I nodi focali sono perciò la quantità di dati ed il miglioramento della qualità dei servizi, parametri che consentono il sostenimento di una crescita assai rapida e duratura. Gli unici limiti sono rappresentati dalla saturazione delle informazioni disponibili (correlata ad un numero finito di potenziali utenti) ed dal raggiungimento di livelli di qualità non ulteriormente migliorabili.

Uno studio dell'OECD condotto nel 2013 sosteneva che l'utilizzo massivo dei Big Data porterà a dei vantaggi sociali di indubbia valenza. Il settore dei trasporti, attraverso il tracciamento dei dispositivi mobili potrà conoscere in anticipo la presenza di congestioni o nodi critici facendo così conseguire un potenziale risparmio mondiale di 500 miliardi di dollari entro il 2020. Nel settore elettrico, l'adozione di applicazioni “smart grid” per il controllo delle operazioni delle applicazioni domestiche, permetterà di ridurre le emissioni di CO2 con un risparmio netto di 79 miliardi di dollari entro il 2020. Nel settore ospedaliero statunitense, la creazione di specifici database ad uso medico consentirà di ridurre gli errori del personale, di migliorare le diagnosi e di incrementare l'efficienza nella gestione, supportare la ricerca e sviluppo e raggiungere degli obiettivi quantificabili in un risparmio pari a 300 miliardi di dollari entro il 2020.

Dal lato dell'offerta, l'utilizzo dei Big Data abilita un meccanismo virtuoso che sottopone le imprese ad una costante pressione al contenimento dei costi ed all'innovazione, promuovendo così efficienze dinamiche di mercato, quindi prezzi più bassi ai consumatori.

Lo sfruttamento dei meccanismi di *Deep Learning* consente alle imprese di ottimizzare le loro strategie commerciali pressoché istantaneamente, seguendo i giudizi e gli altri feedback.

Gli algoritmi di *dynamic pricing*, consentendo alle aziende di reagire istantaneamente ai cambiamenti adeguando contemporaneamente prezzo e produzione ai vincoli interni ed alle fluttuazioni del mercato, permettono di conseguire una maggiore efficienza globale del mercato stesso. L'implementazione degli algoritmi nei mercati digitali permette alle imprese di avvicinarsi alla perfetta discriminazione dei prezzi, decidendo cioè il prezzo in base alla posizione del cliente, allo storico dei suoi acquisti ed alle altre informazioni rilevanti per stabilire la sua disponibilità a pagare (*willingness to pay*) per il prodotto o servizio offerto.

3.3 - Vantaggio competitivo derivante dall'utilizzo dei Big Data

Una basilare differenza riscontrabile tra le moderne applicazioni dei Big Data e il modello economico tradizionale è la mancanza di limiti fisici alla quantità e varietà dei dati che possono essere raccolti in un mondo digitale e le illimitate informazioni che possono essere generate dagli algoritmi su una molteplice varietà di dati o dalla fusione di essi. In termini grafici il fenomeno è rappresentabile con una curva (figura 8) che evidenzia dopo una prima normale fase iniziale una rapida impennata dell'apprendimento al crescere del volume di dati.

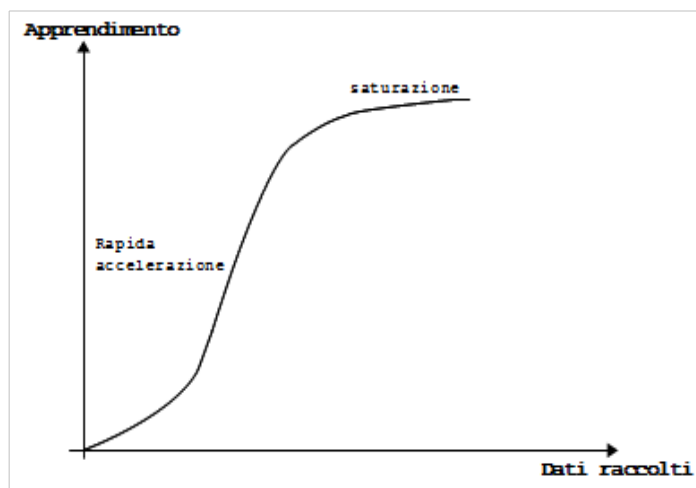


Figura 8 - Curva di apprendimento

I Big Data amplificato l'inclinazione della curva (accelerazione dell'apprendimento) facendo conseguire abbondanti e duraturi benefici. Quando finalmente l'impresa raggiunge la fase di saturazione, le sue dimensioni saranno talmente grandi da ostacolare l'esercizio efficace di qualsiasi pressione competitiva da parte di potenziali rivali, realizzando di fatto una situazione del tipo “il vincitore prende tutto” (monopolio).

Sebbene la piattaforma dominante non faccia nulla che possa essere qualificato come comportamento anticompetitivo, il meccanismo della reatroazione a doppio anello spiegato nel paragrafo precedente le consente di rinsaldare la sua posizione dominante ed esclude le piattaforme rivali all'accesso di nuovi clienti.

In un articolo di “The Economist” in data 13 giugno 2017, l'autore titolava il seguente quesito: “*I Big Data creeranno una intoccabile business elite?*”. Nello stesso si leggeva ancora: “*Sembra che chi metterà le mani per primo sui dati migliori, questo sarà capace di innovare ad una velocità mai vista prima, a spese dei propri rivali [...] Maggiore è il numero di dati, migliori saranno gli algoritmi. E migliori saranno gli algoritmi, migliore sarà la qualità dei servizi offerti. E' questo l'effetto del feedback positivo dell'autoapprendimento. Nel passato le fonti di potere di mercato erano i brevetti (es. Xerox), le esternalità di rete (es. IBM), i regolamenti del governo e le concessioni in franchising (es. AT&T). Tutti questi fattori sono naturalmente ancora rilevanti anche nell'era di internet. La capacità di autoapprendimento delle macchine si aggiunge però quale fattore competitivo tecnologico determinante, per mezzo del quale colui che raggiunge per primo la leadership nel mercato, per qualsivoglia ragione iniziale, guadagnerà in dote una amplificata capacità di migliorare con continuità la propria posizione di mercato. [...] Uno studio dell'OECD dimostra come il primo 5% delle*

aziende mondiali stia già lasciando i rivali nella polvere“.

Per dare una spiegazione a queste affermazioni è necessario riprendere i meccanismi relativi ai *data-driven network effects* descritti nel precedente paragrafo (figura 7). In termini relativi, ogni utente in più che una impresa riesce ad attirare rispetto alle imprese concorrenti può portare ad un differenziale qualitativo positivo dei servizi offerti. Se la differenza qualitativa dei servizi diventa visibile agli occhi degli utenti, i meccanismi del doppio anello a retroazione precedentemente descritti possono portare ad una accelerazione del fenomeno, attirando così sia nuovi utenti del settore (che prima non erano interessati al servizio ma che ora lo diventano, dato l'aumentato livello di qualità) che utenti già fruitori del medesimo servizio offerto dalle imprese concorrenti. Tipico esempio di mercato caratterizzato da *data-driven network effects* è quello dei motori di ricerca, quello dei *social networks*, delle *community-sourced navigation apps*: quando l'utente clicca sulla pubblicità sponsorizzata, il fornitore del servizio non solo incrementa i suoi ricavi, ma raccoglie nuovi dati utili a migliorare la qualità del servizio offerto, incentivando l'attenzione di nuovi utenti e nuovi inserzionisti a fini pubblicitari.

Il risultato di questo fenomeno è che gli utenti possono fidelizzarsi alla piattaforma dominante anche se preferirebbero un differente modello di piattaforma. Per esempio, mentre gli utenti del web potrebbero preferire maggiori garanzie di privacy offerte da alcuni motori di ricerca minori, si affideranno al motore di ricerca leader in quanto fornisce risultati meglio indirizzati. Prova evidente ne è il successo di Google, motore di ricerca piuttosto invadente nei confronti dei suoi utilizzatori (tanto da richiedere l'espressa autorizzazione alla registrazione di informazioni circa indirizzi IP e identificativi dei terminali dai quali vengono effettuate le ricerche) ma scelta quasi obbligata per accedere al più vasto panorama di soluzioni all'interno del web.

Altro esempio può essere quello della navigazione stradale, settore nel quale applicazioni meno conosciute possono offrire migliori funzioni, ma ove gli utenti continuano a preferire l'applicazione dominante (es. TomTom) proprio perché, grazie all'elevato numero di utenti e delle informazioni da questi rilasciate, fornisce migliori aggiornamenti sul traffico.

Una differenza essenziale tra le banche dati tradizionali e i Big Data è la capacità di utilizzare le informazioni in tempo reale, fenomeno noto come “now-casting”. Il *now-casting* è definito come “la previsione del presente, del futuro prossimo e del passato recente” (Marta Banbura) e consiste nell'uso di dati nuovi, aggiornati e rilevati con frequenza elevata per produrre previsioni accurate in tempo reale di variabili che normalmente vengono monitorate con ritardo. L'utilizzo del *now-casting* può consentire di informare in tempo reale circa la

diffusione di una epidemia influenzale partendo dalla semplice rilevazione di un picco di ricerche legate ai rimedi per l'influenza. Nel caso delle imprese, il *now-casting* può portare ad allertare in tempo reale circa la potenziale minaccia di un concorrente, per esempio semplicemente evidenziando l'elevato numero di download di una determinata applicazione e facendo il controllo incrociato con le ricerche preferite dagli utenti. Come già accennato nel primo capitolo, l'istantaneità di accesso ai dati è una prima caratteristica distintiva dei Big Data rispetto alla tradizionale *Business Intelligence*. L'accessibilità in tempo reale alle informazioni arricchisce quindi i dati di un maggior valore intrinseco completamente scollegato dai criteri di valorizzazione tradizionali.

La seconda importante differenza rispetto alla *Business Intelligence* è la possibilità di processare dati grezzi senza richiedere l'intervento umano per indirizzare i risultati dell'elaborazione. Gli algoritmi dei Big Data consentono l'integrazione automatizzata di più fonti di dati ed hanno la capacità di trasformare dati eterogenei in informazioni preziose per il processo decisionale. Questa caratteristica si traduce in un potenziale vantaggio competitivo: la capacità di analizzare contemporaneamente una molteplicità di pacchetti di dati apparentemente non collegati da alcun nesso logico può generare una serie di informazioni inedite tali da guidare una impresa ad intraprendere soluzioni inesplorate per conseguire migliori risultati nel mercato. Per dare un'idea, supponiamo che un database contenente un nome fittizio, un indirizzo e-mail, il sesso e l'età ottenuto per fini statistici da una associazione di alcolisti anonimi venga fuso con un altro pacchetto (ottenuto da altra fonte) ove l'indirizzo email è associato all'indirizzo di residenza: ciò consentirebbe l'individuazione in maniera univoca dell'identità delle persone e la completa conoscenza dei loro dati sensibili e delle loro abitudini (inclinazioni all'abuso di alcol). Una società di assicurazioni che venisse in possesso di tali dati, con un approccio cinico e pragmatico alla realtà dei fatti la farebbe chiaramente desistere dalla stipula di qualsivoglia contratto con i soggetti identificati. Questo aspetto è stato un tema molto dibattuto nel processo di acquisizione di WhatsApp da parte di Facebook, ove la tentata condivisione dei contenuti degli utenti afferenti ai due social network avrebbe potuto mettere a rischio il mantenimento delle condizioni di libera concorrenza nel mercato (attraverso l'incrementato potere) nonché la privacy degli utenti stessi. Quale relazione possa legare l'abuso di posizione dominante e privacy verrà spiegato nel capitolo 4.

In considerazione dei molteplici benefici ottenibili dai Big Data e del vantaggio competitivo che ne deriva, nel caso di una impresa monopolista che abbia guadagnato potere di mercato grazie ad una dotazione importante di dati, è lecito invocare la dottrina della “*Essential Facility*”?

L'*Essential Facility Doctrine* (EFD) è stata da tempo riconosciuta come una delle più controverse tematiche antitrust. La dottrina fondamentale si basa sul principio che una azienda monopolista che si avvale di strumenti unici, in virtù della sua posizione di privilegio ha l'obbligo di condividere le sue dotazioni con chiunque ne richieda l'accesso, incluse aziende potenziali concorrenti. La EFD si delinea come una sottile eccezione alla regola generale che stabilisce che le aziende, anche quelle in regime di monopolio, godono della libertà di scegliere se e con chi contrattare la vendita di un bene o servizio. La nascita della EFD è strettamente legata alla necessità di assicurare l'efficacia del processo di liberalizzazione e l'apertura dei mercati alla competizione. In questo scenario la legge antitrust impone ai proprietari di reti o infrastrutture l'obbligo di concedere, a determinate condizioni, l'accesso alle loro strutture essenziali a condizioni ragionevoli e non discriminatorie.

Nel valutare la liceità dell'applicazione della EFD alla *data economy*, le autorità per la concorrenza statunitense ed europea stanno attualmente affrontando il tema in due modi distinti.

La posizione dell'Autorità Statunitense è piuttosto semplice e chiara: sulla base della sezione 2 dello Sherman Act (che regola la condotta esclusiva verso i concorrenti) e dei successivi casi di legge correlati, le imprese, anche quelle in posizione dominante, non hanno alcun obbligo di negoziare o aiutare le loro concorrenti. Rare eccezioni sono state riconosciute dalla Corte Suprema a tale principio.

La posizione dell'Autorità per la Concorrenza Europea, richiamando l'articolo 102 del TFEU (Trattato sul Funzionamento dell'Unione Europea) stabilisce che la EFD è applicabile qualora sussistano tre condizioni: 1) il rifiuto a negoziare è correlato ad un prodotto o servizio che è oggettivamente necessario per competere in maniera efficace in un mercato a valle rispetto a quello dell'azienda dominante; 2) è probabile che il rifiuto comprometta una efficace competizione nel mercato a valle; 3) è probabile che il rifiuto arrechi un danno al consumatore. Come riportato in un articolo redatto dall'Ordine degli Avvocati Statunitensi "The era of Big Data and EU/U.S. divergence for refusal to deal" (Antitrust, 2017), la diffusione della tecnologia dei Big Data condurrà ad un allontanamento delle posizioni di Stati Uniti e Europa circa l'applicazione della EFD.

L'approccio americano è giustificato dai timori che una impropria imposizione di licenze d'uso potrebbe soffocare l'innovazione e richiederebbe ai tribunali antitrust di agire come centro di pianificazione per identificare quantità, prezzi e altre condizioni di negoziazione, ruolo che va oltre il loro mandato. Inoltre, l'obbligo a negoziare tra concorrenti potrebbe facilitare il peggior nemico dell'Antitrust, ossia la collusione.

La Commissione Europea invece ha recentemente stabilito che la legge generale sulla concorrenza è applicabile al contesto della *data-driven economy* e quindi può essere invocata per rivendicare un maggiore accesso ai dati di proprietà di un operatore economico.

Colangelo e Maggiolino, in contrasto a quanto stabilito dalla Commissione Europea, sostengono nel loro documento di lavoro “Big Data as a misleading facility” pubblicato sull’*“European Competition Journal”* (5 ottobre 2017) che l’applicazione ai dati della EFD risulta una pratica impropria ed ingannevole.

Il loro assunto di base è che il possesso esclusivo di dati non implica essenzialità o monopolio delle informazioni: nel caso di un costruttore di frigoriferi *smart*, i dati esclusivi (essendo l’unico abilitato a leggerli) che esso può ottenere dai suoi apparecchi, utili per dedurre le abitudini alimentari dei suoi clienti, non configura una condizione di essenzialità o monopolio. Le stesse informazioni potrebbero infatti essere derivate dalla lista di acquisti al supermercato o dalle tracce dei pagamenti operati con la carta di credito. Quindi nessun obbligo di condivisione potrebbe essere imposto al costruttore di frigoriferi. In generale, i dati accessibili o acquistabili non dovrebbero mai essere considerati indispensabili.

Una menzione a parte merita il caso dei dati personali: in alcuni casi essi potrebbero configurarsi come “essenziali”, ma essendo questi soggetti alla tutela del Garante per la Privacy, nessuna imposizione di condivisione può essere imposta sugli stessi poiché ne conseguirebbe una chiara violazione della legge sulla protezione dei dati sensibili.

Poiché i dati vengono raccolti disordinatamente e stoccati senza una logica precisa, essi non possono nella fase di raccolta dare risposte a specifiche esigenze di ricerca. Come potrebbe un rivale concorrente rivendicare la necessità di accedere a tali dati se lo stesso non è nella condizione di sapere cosa essi possono rivelare? Come è possibile dimostrare che i dati posseduti da un monopolista sono essenziali se le stesse informazioni possono essere derivate da altre sorgenti? Colangelo e Maggiolino concludono: *“i Big Data possono essere equiparati ai documenti raccolti da un avvocato che approccia un nuovo caso, mentre le informazioni corrispondono all’evidenza che lo stesso può evincere dagli stessi documenti. L’avvocato non vincerà la causa per merito dei documenti, la vincerà solo grazie alla convincente evidenza che egli sarà capace di dedurre dagli stessi. [...] se qualcuno cerca di applicare la EFD ai Big Data, la risposta più semplice è che questi sono l’obiettivo sbagliato. Se esiste qualcosa di essenziale, questo al massimo possono solo essere le informazioni”*.

4 – POTENZIALI EFFETTI NEGATIVI SUL MERCATO

4.1 – Algoritmi e pratiche collusive

Nella primavera del 2015 un commerciante di poster residente negli Stati Uniti che si avvaleva della piattaforma di Amazon per le vendite online, tale David Topkins, guadagnò l'onore delle cronache poiché fu il primo operatore di e-commerce nella storia ad essere perseguito per pratiche anticompetitive contrarie alle leggi antitrust del suo paese. Topkins era il fondatore di Poster Revolution, società rinominata poi nel 2012 con il nome Art.com. Il dipartimento di Giustizia di San Francisco lo dichiarò colpevole di aver manipolato il mercato tramite l'utilizzo di un software da lui sviluppato e successivamente condiviso con altre imprese concorrenti, il quale consentiva di modificare in maniera sincrona con gli altri venditori il prezzo dei poster venduti sulla piattaforma di Amazon in funzione della loro popolarità. In tale occasione il Procuratore Generale Bill Baer, rappresentante della Divisione Antitrust, nel suo discorso ispirato a detta sentenza pronunciò il seguente monito: *“Non tollereremo alcuna condotta anticompetitiva, prenda essa forma all'interno di una stanza piena di fumo o su Internet per opera di complessi algoritmi [...] i consumatori americani hanno il diritto di operare in un leale mercato online, alla pari di quello reale”*.

Dal caso descritto trae spunto il paragrafo introduttivo del documento di lavoro “Intelligenza artificiale e collusione: quando i computer inibiscono la competizione” di Eizrachi e Stucke (Università di Oxford). All'interno della stessa pubblicazione viene citato il caso emblematico relativo alla proposta di vendita di un libro di Peter Lawrence *“The making of a fly”* che, ad opera degli algoritmi automatici di ottimizzazione dei prezzi, venne offerto alla clientela su Amazon allo stratosferico prezzo di 23.698.655,93 \$!

La *Data-Driven Innovation*, pur rimanendo uno strumento assai utile al servizio del mercato, può assumere connotazioni difficilmente interpretabili (ad eccezione di alcuni casi limite tra i quali quello appena menzionato) a mero vantaggio degli attori attivi, ossia della parte più forte che conosce e gestisce i meccanismi della “negoziiazione” virtuale.

Nei capitoli precedenti si è accennato alle molteplici funzioni dei Big Data ed ai potenziali risultati che permettono di conseguire. In ragione dell'assoluta novità che lo strumento rappresenta, un grosso punto interrogativo si pone di fronte ai suoi effetti sulle dinamiche di mercato. Collusione, cartelli (forma avanzata di collusione tra imprese dominanti), abuso di posizione dominante (normalmente nei casi di oligopolio o monopolio) sono dei comportamenti scorretti che vanno a danneggiare le leve del libero mercato.

Nella letteratura economica il termine collusione si riferisce comunemente a qualsiasi forma di coordinamento o accordo tra aziende concorrenti con l'obiettivo di incrementare i propri profitti ad un livello superiore rispetto al regime di normale concorrenza, generando in tal modo una inefficienza.

Al fine di raggiungere e sostenere un equilibrio collusivo duraturo, le imprese concorrenti devono stabilire una struttura di governo delle interazioni che consenta loro di:

- (a) concordare una strategia collusiva
- (b) verificare il rispetto della strategia collusiva da parte delle imprese aderenti all'accordo
- (c) rafforzare la strategia attraverso la punizione di qualsiasi deviazione delle imprese concorrenti

La pratica collusiva viene distinta dagli economisti secondo due tipologie:

- **collusione esplicita:** riferita a condotte anticompetitive che sono mantenute a mezzo di espliciti accordi, siano essi scritti o verbali; il modo più diretto per raggiungere un accordo collusivo diretto è quello di interagire direttamente con le altre parti e concordare il livello ottimale di prezzo o produzione
- **collusione tacita:** si riferisce alle forme di coordinamento anticompetitivo che possono essere raggiunte senza alcun esplicito accordo, ma che le imprese concorrenti riescono a mantenere conoscendo la loro mutua interdipendenza nel mercato

Il basso numero delle imprese e la presenza di barriere all'ingresso sono tradizionalmente identificati come due dei più importanti fattori strutturali che incidono in maniera incrementale sul potenziale rischio di comportamenti collusivi. Un elevato numero di imprese non solo rende difficile l'identificazione di un punto focale di coordinamento, ma riduce anche gli incentivi alla collusione poiché ogni impresa riceverà una porzione minore dei guadagni addizionali (rispetto ad una situazione di concorrenza) derivanti dal regime collusivo. Similmente, in assenza di barriere all'ingresso il regime collusivo è difficilmente sostenibile in maniera duratura, poiché la presenza di prezzi elevati attirerà nuovi entranti facendo deteriorare rapidamente l'equilibrio raggiunto.

L'intelligenza artificiale che caratterizza i Big Data prende forma per mezzo dei suoi algoritmi, ai quali possono essere ricondotti gli effetti delle operazioni svolte dal sistema. Nella valutazione del loro impatto sui due fattori principali che agevolano la collusione (numero delle imprese e presenza di barriere), si possono fare le seguenti considerazioni:

- tipicamente gli algoritmi vengono utilizzati (per fissare dinamicamente i prezzi,

praticare una segmentazione dei consumatori, migliorare la qualità del prodotto) in settori industriali caratterizzati da un limitato numero di imprese di grandi dimensioni in concorrenza tra loro. Per esempio: motori di ricerca, piattaforme di vendita online, *booking agencies*, *airlines*... Tuttavia questi settori sono normalmente caratterizzati da considerevoli barriere all'ingresso legate a elevate economie di scala, economie di scopo e ad effetti di rete. Questi dotano di un vantaggio competitivo le aziende che per prime hanno iniziato ad operare, sostenendo la crescita e la raccolta di grandi quantità di dati per sviluppare algoritmi migliori, i quali a loro volta, avvalendosi degli effetti di rete, faranno conseguire nuove economie di scala e di scopo. Non è agevole stabilire se gli algoritmi siano la causa principale o l'effetto delle barriere all'ingresso, il tipico caso dell'uovo e della gallina

- i Big Data, avvalendosi in particolare dei meccanismi del *now casting*, possono essere utilizzati per identificare rapidamente qualsiasi minaccia nel mercato, consentendo alle imprese dominanti di acquisire preventivamente qualsiasi potenziale concorrente o di reagire in maniera aggressiva al suo ingresso. Dall'altro lato, la sempre maggiore quantità di dati disponibili consente ai potenziali entranti di migliorare la conoscenza del mercato e di ridurre i rischi, quindi i costi di ingresso
- data la celerità di raccolta ed analisi dei dati, gli algoritmi permettono il coordinamento, il controllo e la punizione delle imprese concorrenti anche nei mercati meno concentrati, rendendo il numero delle imprese un fattore di collusione meno rilevante

Quanto detto si può riassumere osservando che le eventuali barriere all'ingresso non sono strettamente correlate all'uso degli algoritmi ed il basso numero di imprese risulta un fattore meno rilevante di collusione.

Le altre due importanti caratteristiche strutturali che influiscono positivamente sui comportamenti collusivi sono la trasparenza del mercato e la frequenza delle interazioni. La trasparenza del mercato consente alle imprese di eseguire un monitoraggio reciproco dei comportamenti: deviazioni dall'accordo sono più facilmente osservabili e ciò rafforza l'accordo stesso. La frequenza delle interazioni permette il ritiro dall'accordo e la punizione aggressiva di ogni deviazione. A differenza dei primi due fattori citati (il numero di imprese concorrenti e le barriere all'ingresso), gli algoritmi incidono in maniera molto più determinante su questi secondi due fattori, ponendo un problema di tutela della libera concorrenza. Si possono fare nel merito le seguenti considerazioni:

- la *data-driven economy* necessita di investimenti in tecnologia al fine di beneficiare dei

suoi vantaggi; non appena alcuni operatori di mercato muovono in tale direzione, le rimanenti aziende del medesimo settore avranno forti incentivi ad operare gli stessi investimenti, rischiando altrimenti l'esclusione dal mercato. Il risultato è una industria ove tutti i partecipanti costantemente raccolgono e osservano in tempo reale le azioni dei consumatori e dei rivali, creando un ambiente trasparente favorevole alla collusione

- l'avvento della *digital economy* ha rivoluzionato la velocità con la quale le imprese valutano ed applicano le loro decisioni. Gli algoritmi consentono di modificare autonomamente il prezzo in tempo reale in funzione delle condizioni rilevate nel mercato, consentendo una immediata offensiva ad eventuali deviazioni dei concorrenti dai livelli collusivi. I meccanismi del *Deep Learning* associati ai dati di mercato consentono una accurata previsione delle azioni dei rivali anticipando ogni deviazione prima che la stessa prenda forma

Seguendo il modello classico della letteratura economica è possibile dimostrare che in un mercato perfettamente trasparente ove le imprese interagiscono ripetutamente e il tempo di reazione tende a zero, la collusione può sempre essere sostenuta come una strategia di equilibrio. Ovviamente il modello si basa su ipotesi estreme, dalle quali la realtà spesso si discosta. Ciononostante gli effetti legati all'utilizzo dei Big Data (disponibilità delle informazioni, velocità) portano ad avvicinarsi molto a tale modello teorico.

Ad avvalorare tale rischio, le Autorità Antitrust Francese e Tedesca (2016. Competition Law and Data) dichiarano :

“ ... processando tutte le informazioni disponibili e quindi monitorando ed analizzando o anticipando la risposta dei loro concorrenti ai prezzi correnti e futuri, le imprese concorrenti possono più facilmente trovare e condividere un equilibrio di prezzo sostenibile sopra il livello di libera concorrenza”.

Un ultimo fattore che favorisce le pratiche collusive delle imprese è quello della simmetria delle loro strutture: più le imprese sono simili, più facile risulterà la condivisione di una strategia vantaggiosa per tutte. Al contrario, l'asimmetria dei costi delle imprese ostacola le pratiche collusive, in quanto ogni impresa giungerà ad un prezzo obiettivo diverso finalizzato a massimizzare i suoi profitti.

Il ricorso agli algoritmi incentiva l'innovazione ed il perfezionamento delle strutture di costo, favorendo uno sviluppo differenziato delle imprese. L'emersione di asimmetrie sortirà l'effetto di disincentivare eventuali pratiche collusive.

La tabella 1 riassume quanto motivato sopra.

Fattori rilevanti di collusione		Effetto degli algoritmi su possibilità di collusione
Caratteristiche strutturali	Numero di imprese	incerto
	Barriere all'ingresso	incerto
	Trasparenza di mercato	aumenta
	Frequenza delle interazioni	aumenta
Variabili lato domanda	Crescita	indifferente
	Fluttuazioni	indifferente
Variabili lato offerta	Innovazione	diminuisce
	Asimmetria dei costi	diminuisce

Tabella 1 - Fattori che incidono su probabilità di collusione (da OECD, 2017)

Il risultato netto degli effetti prodotti dai fattori elencati dipende ovviamente dal loro peso specifico. In base a quanto riportato dagli studi condotti fino ad ora sul tema, la trasparenza del mercato e la frequenza delle interazioni giocherebbero un ruolo dominante rispetto alle asimmetrie dei costi generate, a supporto della tesi che vede i Big Data quale fattore incentivante per le pratiche collusive.

Partendo da questo assunto, Ezrachi e Stucke (nel seguito indicati E&S) distinguono quattro possibili impieghi delle tecnologie informatiche a supporto di pratiche collusive: “*Messenger*” (messaggero), “*Hub & Spoke*” (mozzo e raggio), “*Predictable Agent*” (agente prevedibile), “*Autonomous Machine*” (macchina autonoma).

La prima categoria (*Messenger*), anche se riferita ad un uso della tecnologia oramai obsoleto, merita la citazione di un caso esemplare all'inizio degli anni '90 che divenne la pietra miliare delle pratiche collusive con l'ausilio dei computer, quello di “*Airline tariff publishing*”. Le autorità statunitensi scoprirono che tramite il canale ufficiale della società pubblica delegata alla divulgazione dei prezzi alle agenzie di viaggi, le compagnie aeree comunicavano alle concorrenti la volontà di aumentare le tariffe dei voli attraverso una codifica incentrata sulla prima e ultima data valida per l'emissione dei biglietti ad un determinato prezzo su una determinata rotta. Poiché la volontà di colludere era palese (i computer assumevano le vesti di semplice mezzo di comunicazione), i giudici americani sanzionarono tale comportamento secondo le leggi in vigore.

Da tale momento in poi, con la diffusione di internet e l'evoluzione delle tecnologie di comunicazione, metodi più sofisticati subentrarono nelle comuni pratiche collusive tra aziende.

Una delle maggiori difficoltà ad implementare un cartello tra imprese in un mercato molto dinamico è che i continui cambiamenti relativi a domanda e offerta richiedono frequenti modifiche dei prezzi e della produzione, con il risultato che coloro che prendono parte al cartello devono in qualche modo rinegoziare frequentemente i loro accordi collusivi rischiando di essere scoperti dalle autorità di vigilanza.

L'automatizzazione delle decisioni potrebbe venire in soccorso a tale necessità di allontanare le persone fisiche dalle dirette responsabilità correlate ai comportamenti collusivi, le macchine delle imprese che intendono colludere dovrebbero quindi riprodurre decisioni concordi a delle comuni regole mediante l'utilizzo di algoritmi paralleli. Il modo più facile per le imprese che aderiscono ad un cartello di condurre una comune strategia dei prezzi sarebbe quello di utilizzare uno stesso identico algoritmo di elaborazione dei prezzi. E' questa la categoria definita “*Hub & Spoke*”, ove un “input verticale conduce ad un allineamento orizzontale” (E&S, 2015). E' questa la condizione di Uber, ove un comune algoritmo potrebbe decidere i prezzi delle corse dei taxi inseguendo un prezzo competitivo percepito anziché quello reale di mercato. All'aumentare del numero di taxi convenzionati a Uber, il livello dei prezzi stabilito dallo stesso potrebbe divenire il nuovo riferimento di mercato per le tariffe dei taxi, conferendogli il potere di coordinare l'incremento dei prezzi a sua discrezione.

In genere, la pratica di uniformarsi ad un comune metodo per stabilire i prezzi corre il rischio concreto di essere condannata dalle autorità quale comportamento anticoncorrenziale.

Una soluzione che potrebbe sortire gli stessi effetti di quanto appena descritto ma più difficile da riscontrare da parte delle autorità di controllo sarebbe che “le diverse imprese affidassero lo sviluppo degli algoritmi di elaborazione dei prezzi alla stessa società ed agli stessi programmatori, in modo da utilizzare software apparentemente diversi che però forniscono risultati fondamentalmente allineati ad uno stesso criterio” (OECD, 2017).

Il ruolo più semplice che gli algoritmi possono svolgere a supporto delle pratiche collusive è quello di monitorare le azioni delle imprese concorrenti, verificare la loro aderenza ad una determinata condotta e, nel caso di scostamenti da detta condotta dare seguito con una predeterminata azione punitiva. E' questa la categoria “*Predictabel Agent*”, ove “*il conscio parallelismo prende forma [...] dalla configurazione delle macchine, senza collusione, quando il programmatore sa che una strategia dominante è quella di seguire il prezzo degli altri, quindi [...] sa che se le altre imprese decidono per lo stesso comportamento, un equilibrio*”

può essere stabilito sopra il livello di concorrenza. [...] il computer è quindi programmato per monitorare ed esplorare l'attendibilità di implementare un'azione di interdipendenza, senza avventurarsi in pratiche concertate o accordi illegittimi. Il computer è anche programmato per punire le deviazioni da un possibile tacito accordo e per identificare le imprese non allineate che si discostano dall'equilibrio” (E&S, 2015).

In termini schematici semplificati, l'algoritmo si potrebbe presentare come segue:

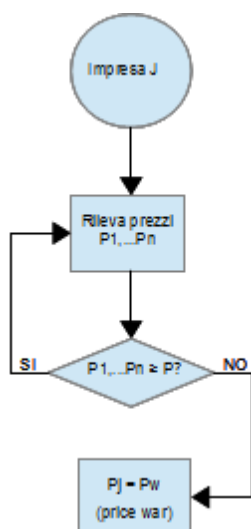


Figura 9 - Algoritmo di monitoraggio (da OECD, 2017)

Data la trasparenza del mercato online, l'algoritmo di monitoraggio di una generica impresa “J” opera una rilevazione costante dei prezzi dei prodotti degli altri concorrenti: qualora il loro prezzo fosse inferiore a quello del regime collusivo “P”, l'impresa “J” risponderebbe con una guerra dei prezzi (Pw).

Altra soluzione ancor più sofisticata per mantenere un rapporto collusivo è quella di utilizzare algoritmi che inseguono in tempo reale il prezzo fissato dall'impresa leader (strategia tit-for-tat) la quale ha la responsabilità di programmare l'algoritmo dinamico che fissa i prezzi sopra il livello di concorrenza.

Come rappresentato nel diagramma in figura 10, l'algoritmo dell'impresa leader “J” stabilisce il prezzo ottimale “P” al quale devono conformarsi le altre imprese che colludono; qualora una di queste tentasse di attirare nuovi clienti abbassando i prezzi, l'algoritmo di monitoraggio rilevarebbe lo scostamento dal regime collusivo ed andrebbe a scatenare una guerra dei prezzi “Pw”.

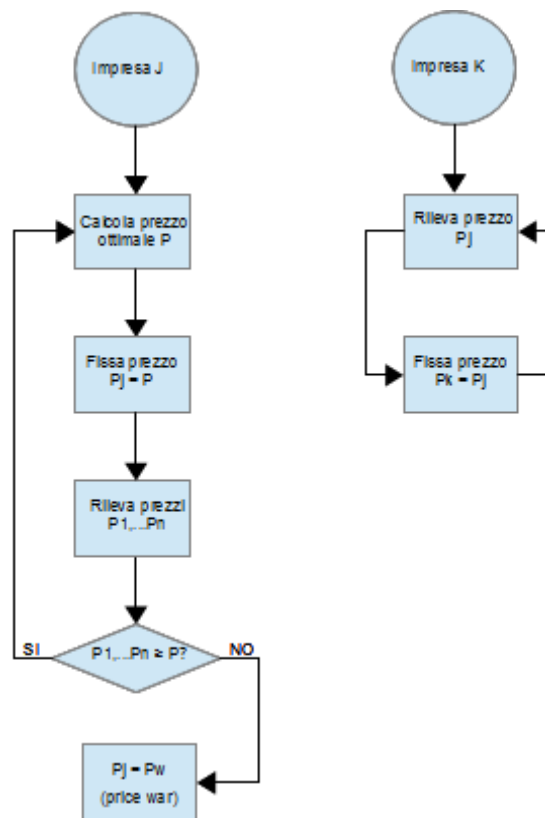


Figura 10 - Algoritmi paralleli (da OECD, 2017)

Nei mercati più dinamici ove imprese di diverse dimensioni vendono diversi prodotti e applicano strategie di business eterogenee, la ricerca di una tacita collusione può essere una pratica molto difficile da condurre, data l'assenza di un punto focale naturale comune. Al fine di evitare la comunicazione esplicita, sempre rischiosa e facilmente scopribile, le imprese che intendono colludere potrebbero concordare un comune coordinamento attraverso segnali dati al mercato e annunci unilaterali di modifica dei prezzi. Sebbene la pratica di lanciare segnali sia facilmente riscontrabile in ogni mercato, essa può causare perdite considerevoli a chi la intraprende. Poniamo il caso di una impresa che incrementi i prezzi con l'intento di comunicare la propria volontà di colludere: qualora la maggior parte dei rivali non recepisse intenzionalmente o meno tale segnale, l'impresa segnalante ne conseguirebbe una perdita di vendite e profitti. Questo rischio potrebbe incoraggiare le imprese ad attendere le azioni degli altri concorrenti, con la possibilità di ritardi o fallimento del coordinamento. Con l'ausilio degli algoritmi questo rischio viene ridotto, consentendo alle imprese di lanciare dei segnali molto rapidi alle imprese rivali e difficilmente interpretabili dai consumatori non essendo

questi muniti di strumenti informatici di pari livello.

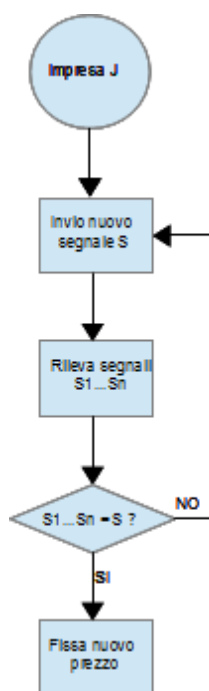


Figura 11 - Algoritmo di segnalazione (da OECD, 2017)

Dal diagramma rappresentato, l'algoritmo dell'impresa leader "J" genera un segnale "S" per verificare la disponibilità delle altre imprese a colludere: se le imprese rispondono con lo stesso segnale allora "J" stabilisce il prezzo ottimale "P" al quale si conformeranno le altre imprese.

Ci possono essere diversi modi di raggiungere l'obiettivo di segnalazione. Un modo potrebbe essere quello di programmare una variazione istantanea dei prezzi durante le ore notturne, la quale non sortirebbe alcun impatto sulle vendite ma verrebbe identificata come un segnale dagli algoritmi delle imprese concorrenti. In maniera più generale, le imprese potrebbero usare gli algoritmi per condividere una moltitudine di dati che verrebbero usati come codici per proporre e negoziare incrementi di prezzo.

Nella quarta categoria descritta da E&S, quella definita "*Autonomous Machine*", "*il computer è programmato per conseguire la massimizzazione del profitto, [...] l'algoritmo opera autonomamente per raggiungere l'obiettivo, [...] consentendo autoapprendimento e sperimentazione*".

Il metodo più sofisticato per raggiungere comportamenti collusivi senza lasciare evidenza di alcun accordo è quello di affidarsi alla tecnologia del *Deep Learning*. La complessità di tali meccanismi è tale che anche gli autori degli algoritmi riescono con difficoltà a stabilire quali

percorsi logici porterà il software ai risultati richiesti. Le esperienze passate indicano che qualora le macchine alle quali le imprese si affidano per stabilire i loro prezzi di offerta si trovassero ad operare in un mercato favorevole ai comportamenti collusivi, le stesse percorrerebbero con estrema rapidità i passaggi necessari a raggiungere l'equilibrio collusivo che consentirebbe di massimizzare i profitti delle rispettive imprese. Tutto questo senza il benché minimo intervento dell'uomo. Anche se un infinito numero di prezzi anticompetitivi potrebbe essere sostenibile per le imprese che colludono, gli algoritmi di *Deep Learning* determineranno con facilità il prezzo che massimizza i profitti delle imprese ed arreca il massimo danno ai consumatori.

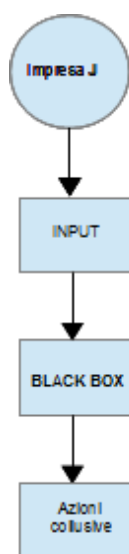


Figura 12 - Collusione quale risultato dell'algoritmo di *Deep Learning* (da OECD, 2017)

L'uso dei meccanismi di *Deep Learning* conduce alla paradossale situazione nella quale le imprese potrebbero manifestarsi incoscienti di porre in essere comportamenti collusivi, essendo questi il risultato delle scelte di una macchina. La collusione tacita ad opera delle macchine potrebbe diventare sostenibile persino in quei mercati dove originariamente non sono soddisfatte le condizioni alla base di un oligopolio secondo la teoria economica tradizionale.

Scrivono E&S: “*se gli umani programmano un computer per ottimizzare i profitti e sanno che attraverso l'autoapprendimento lo stesso reagirà alle mutate condizioni di mercato trovando la collusione quale strategia dominante, saranno essi colpevoli? Forse...*”.

Per giustificare la risposta dubitativa, si deve fare menzione alle regole in essere. Le leggi rivolte alla tutela della libera concorrenza degli stati occidentali, sebbene indirizzate ad economie avanzate, deducono il dolo di un comportamento collusivo sull'evidenza di un

accordo pregresso tra le parti oggetto dell'indagine. L'accezione di “accordo” generalmente presa in considerazione dalla giurisprudenza prevede simultaneamente l'esistenza di una comune volontà delle parti ed una sua manifestazione, sia essa esplicita od implicita. È evidente che tale manifestazione risulta difficilmente riscontrabile qualora i prezzi fossero decisi da algoritmi automatici. Nel caso di algoritmi paralleli o di segnalazione risulta assai difficile imputare a delle persone la responsabilità di un'azione rivelatasi collusiva ad opera delle macchine, sarà assai difficile dimostrare la volontà umana di raggiungere tali risultati.

Il messaggio di Margrethe Vestager, Commissario Europeo per la concorrenza è perentorio: *“Le sfide che i sistemi automatici creano sono reali. Se essi aiutano le imprese a decidere i prezzi, allora possono peggiorare il funzionamento delle nostre economie [...] quindi come sostenitrice della concorrenza penso che dobbiamo dichiarare che le imprese non potranno scappare dalla responsabilità di collusione nascondendosi dietro il programma di un computer”* (2017).

Nell'ambito della *data economy* la linea di demarcazione che individua un comportamento giudicato lecito da uno illecito è piuttosto sottile ed il caso di Athena Capital è esemplare per rappresentare l'inadeguatezza dei parametri di valutazione stabiliti dalle leggi attuali nel caso degli algoritmi. Nel 2014 la SEC sanzionò Athena Capital Research per aver manipolato il mercato, motivando che *“Athena sfruttava complessi algoritmi per manipolare i prezzi di chiusura di migliaia di azioni del NASDAQ [...] Il sofisticato algoritmo il cui nome in codice era “sugo”, si basava su una pratica conosciuta come “marca la chiusura” secondo la quale le azioni venivano comprate o vendute in prossimità del termine delle contrattazioni al fine di incidere sul prezzo di chiusura [...] I massivi volumi trattati da Athena negli ultimi secondi di contrattazione permettevano di opprimere la liquidità disponibile nel mercato e spingere artificialmente i prezzi [...] in favore di Athena. Gli impiegati di Athena [...] erano coscienti dell'impatto sui prezzi del loro algoritmo, al quale nelle loro e-mail interne si riferivano con il termine “possiedi il gioco”*. Le prove che consentirono di dimostrare la condotta dolosa furono banalmente il nome dato all'algoritmo (sugo) ed una lunga serie di e-mail indiziali. Avessero meglio mascherato le loro reali intenzioni, probabilmente i vertici di Athena ne sarebbero usciti senza condanna alcuna (che tuttavia fu limitata ad una semplice multa).

A conclusione di quanto esposto nel presente paragrafo, risulta evidente che gli algoritmi possono fungere da elemento catalizzatore per le pratiche collusive tra le imprese, con l'ulteriore vantaggio di schermare le stesse dalle responsabilità dirette che divengono più difficili da dimostrare. Per scoraggiare il ricorso a tali pratiche, le leggi nazionali dovranno rapidamente evolvere per contrastare efficacemente le moderne tecniche per aggirare il

mercato.

4.2 – Altri scenari anticompetitivi

Nel maggio 2016 le autorità per la concorrenza francese (Autorité de la Concurrence) e tedesca (Bundeskartellamt), nel seguito indicate con AdC&BKA, pubblicarono un report congiunto sul tema “*Competition law and data*” ove si analizzavano le possibili implicazioni della nuova economia dei dati sulle pratiche anticompetitive. Data l'autorevolezza del documento in ambito europeo, risulta utile richiamare in sintesi alcuni punti fondamentali prendendo spunto da essi per accennare ad aspetti non già trattati nei paragrafi precedenti.

Le possibili condotte anticompetitive correlate allo sfruttamento dei dati vengono individuate da AdC&BKA in quattro forme principali:

- (a) fusioni e acquisizioni (*Merger & Acquisitions*)
- (b) condotte finalizzate all'esclusione
- (c) dati quale mezzo per discriminazione dei prezzi
- (d) potere di mercato e tutela della privacy

Fusioni e acquisizioni: per ottenere un migliore accesso ai dati, le imprese potrebbero optare per l'acquisizione diretta di altre imprese che già ne possiedono rilevanti quantità. Qualora l'acquisizione presentasse costi non sostenibili, la fusione concordata delle due imprese porterebbe comunque al conseguimento del medesimo scopo.

In un mercato tradizionale, la fusione tra una azienda radicata nel territorio ed una entrante non porterebbe grandi sconvolgimenti, poiché quella entrante normalmente non possiede quote rilevanti in quel mercato. Nel caso dei mercati sensibili alle attività correlate ai dati, una fusione potrebbe far accedere ad informazioni non note all'azienda entrante, la condivisione dei nuovi dati con quelli ricavati da altri mercati incrementerebbe notevolmente la concentrazione degli stessi in quello specifico nuovo mercato, generando un vantaggio competitivo rispetto agli altri rivali che si avvalgono di database minori. In virtù degli effetti di rete, è verosimile che la fusione di due imprese di successo posizionate su livelli diversi possa precludere totalmente l'ingresso ad altri concorrenti. Ad esempio, un fornitore di servizi online che estrae valore dagli enormi volumi di dati personali potrebbe essere interessato ad acquisire aziende produttrici di computer, smartphone o software al fine di garantirsi l'accesso a nuovi volumi di dati generati dai nuovi utenti del servizio.

Questo ben spiega la tendenza generale: nell'ambito delle società che fanno uso dei Big Data,

il numero delle fusioni ed acquisizioni è in continua crescita, poiché questo è il modo più rapido di accedere a nuovi dati ed espandere le proprie attività vincendo la competizione contro i rivali.

La frustrazione conseguente al mancato rispetto degli impegni presi da parte di Facebook nell'acquisizione di WhatsApp aiuta a spiegare come alcuni paesi in Europa abbiano già iniziato ad aggiornare le rispettive leggi sulla concorrenza. In Germania è in fase di approvazione una legge che consentirà al Bundeskartellamt (ufficio federale contro i cartelli) di intervenire sulle richieste di fusione negli ambiti nei quali i *network effects* e la mole di dati ricoprono un ruolo rilevante nella competizione.

Condotte finalizzate all'esclusione: AdC&BKA sostiene che la negazione all'accesso ai dati (*refusal to access*) da parte di una azienda con elevato potere di mercato potrebbe portare all'esclusione di alcuni rivali e ad un indebolimento della concorrenza. Questo punto si collega a quanto già discusso nel paragrafo 3.3 ove ci si chiedeva se fosse lecito o meno equiparare i dati ad una *essential facility* e le conclusioni permangono quelle già esposte. In rari casi di società ex monopoliste pubbliche divenute poi società private (es. Gaz de France), le autorità nazionali hanno ritenuto lecito imporre la condivisione dei dati alla nuova realtà privata. Per scongiurare lunghi e costosi contenziosi giudiziari, i governi potrebbero essi stessi cominciare a distribuire i dati da loro raccolti creando opportunità per le aziende più piccole: la Germania impone oggi alle società di assicurazioni di tenere un registro di dati statistici, incluso quelli degli incidenti stradali, accessibile anche alle società più piccole le quali non sarebbero in grado di stilarne uno per proprio conto.

Dati quale mezzo per discriminazione dei prezzi: la raccolta di dati può facilitare la discriminazione dei prezzi. Con le informazioni rese disponibili dai Big Data, le imprese sono in grado di ricostruire il profilo economico (preferenze, abitudini, disponibilità) dei loro clienti e di operare quindi una discriminazione quasi perfetta dei prezzi (discriminazione di primo grado), segmentando la clientela in funzione della loro reale disponibilità a pagare i prodotti o i servizi (*willingness to pay*). Qualora una sola impresa fosse in possesso di informazioni esclusive circa le dinamiche del mercato e le preferenze dei consumatori, questo rappresenterebbe un vantaggio competitivo notevole, capace di far guadagnare un elevato potere di mercato che condurrebbe poi verso una condizione di monopolio. Nel caso dei potenziali concorrenti provassero a entrare e competere, l'impresa dominante sarebbe in grado di bloccarne l'accesso proponendo dei prezzi adeguatamente bassi ai clienti tentati dal cambio, mantenendo così saldo il rapporto in essere.

Potere di mercato e tutela della privacy: nonostante la tutela della privacy sia un tema che non ricade nelle competenze delle autorità per la concorrenza, in alcuni casi esso rappresenta un fattore strettamente correlato agli abusi di posizione dominante. Prendiamo ad esempio il caso di una azienda monopolista che necessita di un flusso considerevole di nuovi dati per alimentare il suo business: se questa li raccoglie seguendo metodi non rispettosi delle leggi sulla privacy, da tale comportamento scorretto la stessa riesce a ricavare un vantaggio competitivo che le permette di rinsaldare la sua posizione di potere. In questo caso la tutela dei dati personali non rimane un mero fatto formale, poiché la mancata osservanza delle regole da parte dell'azienda dominante si ripercuote in forma di concorrenza sleale contro le aziende rivali. In un mercato fortemente concentrato le aziende dominanti che non competono per la clientela, in considerazione della loro posizione, potrebbero non avere alcun incentivo a mantenere un alto livello di riservatezza dei dati sensibili nello svolgimento dei loro servizi. Quale effetto collaterale, l'abuso di posizione dominante potrebbe quindi portare al degrado delle garanzie di tutela della privacy. Sulla base di tale assunto, le autorità italiane Agcom, Antitrust e Garante della privacy hanno avviato nel mese di giugno di quest'anno una indagine conoscitiva sui Big Data con la motivazione che *“Le potenzialità dei Big Data, anche rispetto a dati anonimi o aggregati, può tradursi in profilazioni sempre più puntuali ed analitiche, con il rischio di nuove forme di discriminazione per le persone e, più in generale, in possibili restrizioni delle libertà”*.

I possibili scenari descritti confermano che un utilizzo doloso dei Big Data può portare alla distorsioni degli equilibri di mercato, in primis quello di ostacolare l'accesso di nuovi attori volenterosi di competere. L'abnorme investimento richiesto per accedere ai Big Data può rappresentare in sé una barriera all'ingresso di potenziali nuovi entranti. Chi intraprende la strada dei Big Data (*cloud computing* a parte) deve inizialmente investire in costosi data center, server, software per processare i dati, connessioni internet alla massima velocità di trasmissione, firewall di protezione e assai costose figure professionali quali scienziati informatici e programmatori. Una volta che il sistema è completamente operativo, i dati incrementali potranno “istruire” e migliorare gli algoritmi con costi molto contenuti. Questa struttura di costo, caratterizzata da elevate economie di scala e di scopo, avvantaggia chi ha già raggiunto grandi dimensioni e può quindi portare alla concentrazione del mercato nelle mani di pochi operatori.

Scrivono Eyrachi (2016. Virtual Competition): *“Le super piattaforme fanno accedere a troppa potenza [...] con molti più dati freschi di altri, esse possono permettere di rilevare velocemente minacce competitive. Le loro tasche capienti consentono di acquistare le startup*

che potrebbero un giorno divenire rivali. Esse possono anche manipolare il mercato ove operano, [...] i loro algoritmi che reagiscono celermente non consentono ai loro rivali di guadagnare nuovi clienti abbassando i prezzi. La mano invisibile sta diventando una mano digitale“. Difficile trovare sintesi migliore a conclusione di quanto esposto.

5 - CONCLUSIONI

La disponibilità di prezzi online consultabili in qualsiasi momento ha incrementato la trasparenza del mercato sia per i consumatori che per le imprese. Dal lato della domanda, gli strumenti informatici di comparazione dei prezzi consentono ai consumatori di verificare il miglior prezzo di un prodotto tra un'ampia quantità di venditori e facilitando la sua scelta, forzando le aziende a competere più aggressivamente sul fronte dei prezzi. Allo stesso tempo, le imprese razionali sfrutteranno lo stesso grado di trasparenza per identificare il miglior posizionamento delle loro offerte nel mercato.

Nella crescente economia digitale, le imprese si stanno conformando all'utilizzo di strumenti sempre più sofisticati per analizzare i dati in loro possesso al fine di studiare il comportamento dei consumatori e dei loro rispettivi concorrenti. I moderni software possono monitorare i livelli di prezzo di un concorrente su migliaia di prodotti e reagire di conseguenza aggiustando il prezzo di un simile paniere nell'arco di qualche secondo. Le imprese si affidano sempre di più agli algoritmi automatici per stabilire i prezzi di vendita, abbandonando le tradizionali pratiche manuali di supervisione e modifica da parte di un operatore. Un caso emblematico che mise in discussione l'attendibilità dei calcoli effettuati dagli algoritmi fu quello di una “guerra dei prezzi tra robot” che sfociò nell'offerta di un libro (di nuova fattura) sul portale di Amazon ad un valore superiore ai ventitré milioni di dollari americani! L'inspiegabile deviazione dai livelli di prezzo “terreni” fu imputabile a due venditori che avevano “istruito” gli algoritmi ad inseguire i prezzi dei concorrenti, facendo loro trovare nella collusione al rialzo la migliore strategia per ottimizzare i profitti. Il caso di Amazon è un caso estremo, ma dimostra che gli algoritmi possono avere effetti reali nei mercati virtuali ove ad ogni cambio dei prezzi non si pone il problema di re-etichettare fisicamente i prodotti o di ristampare i cataloghi.

I Big Data hanno creato un ambiente dove le imprese possono “vedere” il mercato, attraverso lo schermo di un computer possono apprendere il passato, osservarne il presente e prevederne il futuro, rispondere in tempo reale ai cambiamenti del comportamento dei consumatori e dei concorrenti. L'utilizzo dei Big Data sta consentendo l'apertura di nuovi mercati, l'accesso di nuove imprese ai mercati esistenti, incrementando il potere di acquisto dei consumatori.

A fronte degli innegabili benefici menzionati, permangono dei leciti dubbi riguardo i cambiamenti indotti sulla società e sui mercati quale effetto di un massivo utilizzo di questi moderni strumenti.

Dal punto di vista sociale non vi è dubbio che si tenderà ad una maggiore esposizione dei dati

personali (privacy), i grandi portali di informazione attraverso le notizie diffuse potranno influenzare la percezione della realtà dei lettori, le persone in qualità di utilizzatori di beni e servizi potranno essere “economicamente discriminate” in ragione di loro dati peculiari quali ad esempi età, razza, religione, luogo di residenza.

Dal punto di vista degli effetti sui mercati, va rilevato che una maggiore trasparenza rende il mercato più vulnerabile al coordinamento delle azioni tra imprese. La possibilità di monitorare i prezzi dei concorrenti in tempo reale e di rispondere rapidamente agli attacchi riduce ampiamente il vantaggio del “primo a muovere” al ribasso dei prezzi. Le deboli prospettive di guadagno legate alle azioni per sottrarre clienti alle aziende rivali comportano un indebolimento della concorrenza.

La stabilità necessaria per sostenere la tacita collusione trae giovamento dal fatto che le macchine non riproducono le paure o le intemperanze umane. Se adeguatamente programmati, gli algoritmi possono esprimere tendenze virtuose quali l'avversione al rischio, la coerenza delle strategie, la costanza delle azioni. I computer non temono le multe e la reclusione, non rispondono con l'ira a situazioni di conflitto. In un mondo ove i dati saranno comuni e la massimizzazione del profitto l'obiettivo a cui tendere, risulta verosimile che se le macchine verranno lasciate nella condizione di decidere i modi, la collusione potrà sempre essere la risposta migliore.

La linea di demarcazione introdotta dai Big Data tra l'operato degli algoritmi e l'azione umana solleva una questione riguardo l'attribuzione delle responsabilità. Le leggi si avvalgono dei concetti di volontà e manifestazione fattuale della stessa: come si dovrà allora giudicare la responsabilità di un operatore che non fosse nella condizione di prevedere il comportamento della macchina? Sarà eticamente accettabile affidarsi ad una modalità di giudizio fondata sulla attendibilità di un risultato? Le probabilità quale prova della rea condotta? Per evitare che tale discussione finisca con il compromettere i fondamenti della giurisprudenza, è necessario interrogarsi fin da subito circa le possibili soluzioni per integrare l'etica e la legalità all'interno di questa nuova tecnologia.

Data la natura multidimensionale dei Big Data, le autorità legislative dovranno sviluppare nuove regole coinvolgendo nella stesura la più ampia base possibile: autorità per la concorrenza, associazioni dei consumatori, garanti per la privacy, enti normativi settoriali ed esperti di scienze informatiche. Carenza di interventi normativi o sovra-regolamentazione potranno parimenti sortire effetti deleteri per le parti in gioco. Norme chiare e bilanciate, verifiche sugli algoritmi di determinazione dei prezzi, monitoraggio dei prezzi, stretta vigilanza sulle acquisizioni anticompetitive sono gli ingredienti base per prevenire le

distorsioni del mercato; qualsiasi azione risolutiva venga intrapresa nel futuro, dovrà scaturire da un approccio cauto e da una approfondita analisi del fenomeno.

RIFERIMENTI BIBLIOGRAFICI

AUTORITÉ DE LA CONCURRENCE, BUNDESKARTELLAMT (2016). *Competition law and data*. Disponibile su

<www.autoritedelaconcurrence.fr/doc/reportcompetitionlawanddatafinal.pdf>

BARRETT, N. (2017). Will Big Data create a new untouchable business elite? *The Economist*. Disponibile su: <<https://medium.economist.com/will-big-data-create-a-new-untouchable-business-elite-8dc23bcaa7cb>> [Data di accesso 10/09/2017]

BOURREAU, M., DE STREEL, A. e GRAEF, I. (2017). *Big Data and competition policy: market power, personalised pricing and advertising*. CERRE project report.

COLANGELO, G. e MAGGIOLINO, M. (2017). Big Data as a misleading facility. *European Competition Journal*, 13, 1-33.

COMINO, S. e MANENTI, F. (2014). *Industrial organisation of high-tech markets*. Cheltenham: Edward Elgar

EUROPE ECONOMICS (2017). *Big Data: what does it really mean for competition policy?* Disponibile su <www.europe-economics.com/publications/mar_-_big_data.pdf>

EZRACHI, A. e STUKE, M. E. (2015). *Artificial intelligence and collusion: when computers inhibit competition*. Working paper CCLP (L) 40.

GARANTE PER LA PROTEZIONE DEI DATI PERSONALI (2017). Big Data: Agcom, Antitrust e Garante privacy avviano indagine conoscitiva. Disponibile su <<http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/6441412>> [Data di accesso 10/09/2017].

LUGARD, P. e ROACH, L. (2017). The era of Big Data and EU/U.S. divergence for refusal to deal. *Antitrust Vol. 31*

OECD (2016). *Big Data: bringing competition policy to the digital era*. Note by Annabelle

Gawer.

OECD (2016). *Big Data: bringing competition policy to the digital era*. Background note by the secretariat.

OECD (2016). *Hearing on Big Data*. Note by BIAC

OECD (2017). *Algorithms and collusion*. Background note by the secretariat.