



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

DIPARTIMENTO DI SCIENZE STATISTICHE

CORSO DI LAUREA IN ECONOMIA E MANAGEMENT

PROVA FINALE

"BIG DATA: NUOVE POTENZIALITÀ PER LE AZIENDE"

RELATORE:

CH.MO PROF. TOMMASO DI FONZO

LAUREANDO/A: SARA DOTTO

MATRICOLA N. 1065156

ANNO ACCADEMICO 2015-2016

Sommario

INTRODUZIONE.....	4
Capitolo 1. I Big Data: fenomeno in espansione.....	6
1.1. Sul concetto di big data.....	6
1.2. Dimensioni crescenti.....	7
1.3. Più dati ma meno esattezza: da causalità a correlazione.....	8
1.4. Big data: qual è il loro valore?.....	11
1.5. Big data: quali utilizzi possibili?.....	13
1.6. Rischi e problemi.....	13
Capitolo 2: Big Data e Management.....	15
2.1. Big data: nuova risorsa per le aziende.....	15
2.2. Cosa cambia con i big data?.....	15
2.3. La rivoluzione dei big data: varie reazioni.....	18
2.4. Come sfruttare i big data in modo profittevole?.....	19
2.5. Big data e supply chain management.....	20
2.6. Big data e marketing.....	26
Capitolo 3: I Big Data e le conseguenze per le statistiche ufficiali.....	35
3.1. Big data e statistiche ufficiali, novità e tradizione: scontro o fusione?.....	35
3.2. Big data nelle statistiche ufficiali: nuove sfide e opportunità.....	37
Capitolo 4. Big data e Internet of Things (IOT).....	40
4.1. Internet of things: l'origine dei dati.....	40
4.2. Applicazioni.....	41
4.3. Connessione tra big data e IOT.....	41
Capitolo 5. Conclusioni.....	43
Bibliografia.....	44

INTRODUZIONE

L'argomento che si considera in questo lavoro è quello dei Big Data, del loro impatto in vari ambiti pratici, soffermandoci in modo particolare sulle applicazioni e sulle conseguenze in campo aziendale.

Le tecniche statistiche per analizzare i Big Data costituiscono uno sviluppo originale, un'estensione e un ampliamento della tecnica di regressione, l'unica che ho affrontato nel mio percorso di studi. I metodi impiegati, infatti, sono molto complessi e fanno parte di un mondo in continua e veloce evoluzione: un'analisi dettagliata è quindi fuori dalla mia portata, perché richiede conoscenze approfondite in campo statistico e informatico, aree non facenti parte del percorso di studi che ho seguito.

Alcune classi di metodi sono tecnologie definite "large P large N" (a indicare il gran numero di variabili, P, e l'enorme mole di dati, N), carotaggi e *machine learning*; l'approccio che utilizzano, e che costituisce la grande rivoluzione portata dai Big Data, è che le strategie di modellizzazione dei dati non sono più *model-based*, bensì *data-based*.

Questo significa passare da un mondo in cui ci si basa su dei modelli "imposti dall'alto" per fare previsioni, a prescindere dalla natura dei dati, e talvolta persino del problema con cui si ha a che fare, e le analisi si effettuano tramite un campionamento sulla popolazione, cioè analizzando, in base a determinati criteri di selezione, solo una parte (un campione) degli elementi che la compongono, a un mondo in cui non si "calano dall'alto" dei modelli che comprendono delle assunzioni di base, ma si considera la totalità dei dati disponibili e ci si affida alle informazioni da essi fornite. Sono proprio i dati a indicare quali siano gli aspetti salienti, i trend che si manifestano nei *dataset*. L'approccio è per questo motivo basato su di essi, non si fanno ipotesi a priori, né ci si basa su un modello predefinito. Non si cerca più, e sarebbe impossibile farlo data la mole di dati a disposizione, di comprendere la causalità tra i fenomeni, ma ci si accontenta di scoprire la *correlazione* tra essi: è sufficiente controllare come le variabili si muovono congiuntamente per avere la soluzione ai problemi posti, senza indagare la relazione di causa-effetto. Questo tipo di indagine non garantisce la stessa profondità di analisi e accuratezza fornita dai metodi tradizionali come la regressione: la correlazione costituisce una sorta di semplificazione di tale concetto, ma è sufficiente, grazie anche all'aumento dei dati presi in esame, a estrarre valore dal confuso ammasso di dati raccolti.

L'ambito di indagine è un campo di recente sviluppo, in continua e veloce evoluzione, pertanto si può far riferimento a una letteratura relativamente ridotta, ma a una fiorente pubblicistica, dal momento che la ricerca e le analisi procedono in maniera rapida e interessante.

Si procederà nella trattazione a specificare come si è sviluppato il fenomeno dei Big Data, quali sono le principali caratteristiche distintive, le potenzialità e i rischi che si incontrano. In seguito si analizzeranno gli impieghi nel campo aziendale e le conseguenze per il management di questo cambio di prospettiva radicale, con alcuni esempi concreti e focalizzandosi in particolare sulle funzioni del marketing e della supply chain management. Infine, si farà cenno a due ambiti collegati a quelli esaminati: le conseguenze per le statistiche ufficiali in seguito allo sviluppo del fenomeno dei Big Data e all'Internet of Things, realtà anch'essa strettamente correlata a quella dei Big Data.

Capitolo 1. I Big Data: fenomeno in espansione

1.1. SUL CONCETTO DI BIG DATA

Non esiste una definizione univoca di Big Data: questo termine indica un fenomeno complesso, nato a partire dagli anni Duemila, con l'esplosione della quantità di informazioni e *dataset* disponibili, e tuttora in espansione, grazie ai progressi tecnologici. I campi inizialmente interessati dall'enorme aumento di informazioni sono stati la genomica e l'astronomia, ma la diffusione è stata rapida anche in altri settori, dalla medicina al business. Trattando di Big Data, ci si riferisce dunque alla raccolta e all'interpretazione dell'enorme quantitativo di informazioni e *dataset* di diverso genere, resi disponibili dai miglioramenti informatici e tecnologici, e che diventano difficili da trattare con il tradizionale approccio statistico, data la loro eterogeneità e diversità: non necessariamente questi dati possono essere organizzati in database o tabelle.

Per meglio definire il fenomeno, ci si può soffermare sull'analisi delle caratteristiche che gli sono attribuite in letteratura: le cosiddette "3V", che stanno per Volume, Varietà, Velocità.

In primo luogo, "Volume" sta ad indicare la crescita esponenziale della quantità di informazioni a nostra disposizione: esse provengono da fonti diverse e sempre più numerose, quali Internet, pc, smartphone, sensori e altri dispositivi tecnologici.

In secondo luogo, "Varietà" indica che l'agglomerato di dati che abbiamo a disposizione è fortemente eterogeneo e non organizzato: si tratta di informazioni spesso confuse e disaggregate, portatrici sì di un grande valore potenziale, ma che necessitano di rielaborazioni e interpretazioni per fornire informazioni utili e corrette. Inoltre, i dati sono generati da vari tipi di fonti e sono in formati diversi: si rilevano movimenti, telefonate, foto, messaggi, flussi di persone, eccetera; la loro eterogeneità rende impossibile un utilizzo immediato, ma richiede di elaborarli per estrarne il valore implicito.

In terzo luogo, "Velocità" si riferisce all'impressionante velocità con cui il fenomeno è nato e si è sviluppato, ma anche al modo in cui le informazioni sono trattate e diventano disponibili per l'utilizzo: è fondamentale in molti ambiti avere informazioni in tempo reale o quasi, sia per agire in maniera adatta allo sviluppo di un fenomeno sia per adattare le previsioni.

A volte, a completamento della definizione, si aggiunge alle tre precedenti una "quarta V": può essere "Valore", "Variabilità", o "Virtuale". Tutte e tre le alternative sono altrettante caratteristiche dei Big Data: dalle informazioni raccolte ed elaborate si ricava un enorme valore che può essere impiegato nei campi più disparati; i dati scontano una grande

variabilità; si tratta di informazioni virtuali, digitali, raccolte grazie all'espandersi dell'informatica.

Una differente definizione di questo fenomeno si riferisce alla possibilità di prevedere nuovi trend e nuove indicazioni muovendosi solo avendo a disposizione enormi quantitativi di dati. Ciò che cambia rispetto all'usuale approccio nell'interpretazione dei dati per individuare tendenze nuove è la domanda da porsi: non più *perché*, ma *che cosa*. I big data permettono di indagare relazioni di correlazione, non di causalità: si possono prevedere trend e tendenze, fatti e aspettative, ma senza conoscere il motivo per cui certi eventi accadono. Tuttavia, questo approccio è più che sufficiente per estrarre dai dati le informazioni che occorrono per prendere decisioni ed effettuare previsioni su vari fenomeni: conoscere come le variabili si muovono congiuntamente è diventato più interessante di valutarne i nessi causali.

1.2. DIMENSIONI CRESCENTI

Il concetto di Big Data si riferisce a “data set di dimensioni troppo grandi perché i dati possano essere analizzati, immagazzinati, analizzati e gestiti con le tradizionali tecniche cui si rifanno i normali database”. (Manyka et al., 2011). Questa definizione mette a fuoco una problematica importante nel discutere di big data: le dimensioni e le quantità dei dati rilevati e immagazzinati a nostra disposizione.

La crescita esponenziale della quantità di dati che sono generati, immagazzinati e gestiti è all'origine del fenomeno dei Big Data e costituisce una criticità notevole: infatti, la possibilità di conservare i dati cresce più lentamente rispetto alla quantità di dati che si rilevano, aprendo questioni su nuove modalità per gestire l'aumento di dati a nostra disposizione.

Secondo il già citato report del Mc Kinsey Global Institute (Manyka et al., 2011), la quantità di dati immagazzinata nel 2010 dalle imprese è di più di 7 exabytes di informazioni, e che quella immagazzinata in PC e altri dispositivi simili dai consumatori, sempre nel 2010, è di 6 exabytes (1 exabyte corrisponde a 10^{18} bytes). Inoltre, si stima che nel 2009 la media dei dati immagazzinati sia stata di 200 terabyte (1 terabyte= 10^{12} bytes) per ogni impresa con più di 1000 dipendenti negli USA e che in alcuni settori si siano ottenuti più di 1 petabyte (10^{15} bytes) di dati per ogni azienda. Nel 2010, in Europa i dati generati corrispondono a circa il 70% di quelli ottenuti negli USA: 11 exabytes contro i 16 exabytes americani.

Queste stime suggeriscono che, almeno nel breve periodo, la possibilità di generare valore dai dati sia appannaggio dei Paesi più sviluppati, tuttavia la quantità di dati generati sta crescendo in modo esponenziale e non vi è alcun dubbio che anche altre economie in brevissimo tempo potranno averne a disposizione un ingente ammontare, di pari passo con la diffusione della tecnologia sia tra le imprese che nella popolazione.

Per avere un'idea più concreta dell'enormità di informazioni che si creano ogni giorno, basti pensare che, nel 2012, presso la Borsa degli Stati Uniti, passavano di mano circa 7 miliardi di azioni al giorno, due terzi delle quali elaborate attraverso algoritmi basati su modelli matematici. Inoltre, nel 2012 Google trattava circa 24 petabytes di dati al giorno, mentre si caricavano su Facebook oltre dieci milioni di nuove foto ogni ora (Cukier, Mayer-Schönberger, 2013). Ancora, gli stessi autori riportano che su YouTube gli utenti caricano più di un'ora di video al secondo e il numero di *tweet* postati sul social network Twitter è di oltre 400 milioni al giorno.

1.3. PIÙ DATI MA MENO ESATTEZZA: DA CAUSALITÀ A CORRELAZIONE

Questa immensa mole di dati è archiviata e rielaborata dalle grandi Internet companies ed è la chiave per innumerevoli nuovi utilizzi delle informazioni, nei campi più svariati, in primo luogo per effettuare previsioni.

L'approccio che caratterizza i big data è quello di applicare tecniche matematico-statistiche alle enormi quantità di dati in nostro possesso per ricavarne trend e previsioni: si tratta di un modo diverso e innovativo rispetto a quello della statistica tradizionale. In effetti, siamo passati in brevissimo tempo da un mondo in cui la difficoltà di reperire informazioni e dati rendeva impossibile conoscere l'intera popolazione di interesse e si era costretti a restringere il campo di analisi a un campione sufficientemente rappresentativo della stessa per poi estendere i risultati ottenuti, a un mondo in cui la sovrabbondanza di dati può consentire di conoscerne molti di più, rendendo le previsioni molto più affidabili. Questo processo richiede però un cambio di approccio e costituisce una svolta dal punto di vista delle tecniche statistiche utilizzate (Cukier, Mayer-Schönberger, 2013).

Uno dei problemi della statistica, fin dagli albori di questa disciplina, è sempre stato quello dell'accesso ai dati e alle informazioni che descrivono la popolazione, intesa come l'insieme di elementi di interesse da analizzare. Scegliere di utilizzare tutti i dati a disposizione, quando presenti, può essere troppo oneroso, sia in termini di tempo che di costi, perciò si procede al

campionamento, cioè alla selezione, mediante opportuni criteri, di una parte della popolazione in modo tale da agevolare l'analisi ed estendere successivamente i risultati a tutti gli elementi. Il metodo del campionamento intrinsecamente dà luogo a degli errori, che possono essere ridotti attraverso la casualità del campione, più che ampliandone la numerosità (Cukier , Mayer-Schönberger, 2013): il campionamento richiede delle assunzioni alla base, soprattutto sul metodo di selezione applicato.

Il miglioramento tecnologico che si è riscontrato in questi anni ha reso molto meno oneroso la gestione e rielaborazione di enormi quantità di dati: è pertanto possibile, in svariati campi, superare il metodo del campionamento e trattare tutti i dati o quasi, in modo da ottenere una maggior profondità di analisi, senza dover effettuare approssimazioni, ma potendone cogliere appieno i diversi aspetti. Spesso, infatti, alcuni trend rimangono nascosti nell'applicare il campionamento: si tratta pur sempre di una semplificazione della realtà, di una modellizzazione, e questo non permette di esaminare con precisione sufficiente dei fenomeni di nicchia, riscontrabili talvolta in corrispondenza delle code della distribuzione.

Una conseguenza immediata dell'utilizzo di un quantitativo sempre maggiore di dati è la necessità di convivere con l'errore: i dati inevitabilmente contengono degli errori e risulta impossibile eliminarli del tutto. Pertanto, nel fare previsioni riguardo a trend generali, si è costretti ad accettare delle imprecisioni nei dati, pur di averne a disposizione un quantitativo così elevato: si preferisce la quantità alla qualità dei dati, ritenendo (ma questo non è sempre vero) che le stime ottenute siano più affidabili se ne aumenta la mole. Ciò comporta un ulteriore cambio di prospettiva: non si cerca più l'assoluta esattezza dei dati disponibili, ma si preferiscono dei dati meno precisi che però consentono, grazie alla loro quantità, di ovviare al problema della qualità per ottenere stime precise. "I dati disponibili sono spesso non strutturati e ingombranti, ma c'è un enorme aumento di segnale nel rumore, che aspetta di essere diffuso." (McAfee, Brynjolfsson, 2012). Appare perciò significativa, e per qualche verso illuminante, l'affermazione del direttore di ricerca di Google, Peter Norvig.: "Non abbiamo algoritmi migliori. Noi abbiamo soltanto più dati"

Un aspetto di fondamentale importanza da considerare nel momento in cui si parla di stime effettuate utilizzando i Big Data è quello della correlazione. La correlazione permette di stabilire se e in che modo due variabili statistiche sono collegate tra loro, basandosi sulle variazioni di una variabile a fronte delle variazioni dell'altra.

Tuttavia, il concetto di causalità è estraneo a quello di correlazione: quest'ultima si limita a descrivere, attraverso opportuni indici, l'interrelazione lineare tra coppie di variabili, senza stabilire nessi di causalità. Si è in presenza di un'elevata correlazione positiva se, a fronte del cambiamento di una variabile, l'altra varia nello stesso senso e l'indice di correlazione è vicino al valore 1; si parla di correlazione negativa se l'indice di correlazione è vicino a -1 e al cambiamento di una variabile corrisponde una variazione in senso opposto dell'altra. D'altro canto, si riscontra una mancanza di correlazione se il cambiamento in una variabile non comporta alcuna variazione nell'altra variabile considerata, l'indice di correlazione è pari a zero o vicino ad esso.

Ciò che i Big Data permettono di evidenziare nell'effettuare previsioni è soprattutto la correlazione tra le variabili considerate: se essa è elevata, la probabilità che siano collegate è alta (Cukier, Mayer-Schönberger, 2013) permettendoci di “capire il presente e prevedere il futuro”. Così, se sappiamo che due fenomeni sono fortemente correlati, sarà sufficiente monitorarne uno per tenere sotto controllo i cambiamenti dell'altro.

Le correlazioni, però, non esprimono la causalità: mostrano cosa accade, ma non ne spiegano la ragione. Questo aspetto risulta sufficiente ai fini della comprensione dei trend e della previsione di fenomeni, tuttavia si richiede un cambiamento nella *forma mentis* dell'indagine statistica: siamo abituati a indagare le relazioni di causa-effetto e a chiederci il perché dei fenomeni che accadono, mentre con i Big Data si possono ottenere sì previsioni efficienti ed efficaci, ma è molto più difficile comprendere le relazioni di causalità.

Inoltre, l'approccio su cui si basa la statistica tradizionale, cioè la verifica di ipotesi e la ricerca di relazioni lineari tra le variabili, richiede delle assunzioni alla base che con l'utilizzo dei Big Data non saranno più necessarie. In particolare, il procedimento della verifica di ipotesi postula, in primo luogo, la formulazione delle ipotesi da verificare attraverso l'utilizzo di una funzione test applicata al campione, ma si corre il rischio che le intuizioni sulle quali si basano le ipotesi di partenza siano fuorvianti e rischino di falsare l'intero risultato.

Le tecniche di regressione lineare, che permettono di verificare l'esistenza di una relazione lineare tra due o più variabili qualitative e/o quantitative, non prendono in considerazione la possibilità che vi sia sì una relazione tra le variabili, ma che questa sia più complessa di una relazione lineare. I Big Data, invece, permettono di analizzare anche relazioni non lineari tra le variabili, anche se in questo caso servono strumenti adeguati per gestire la complessità del problema.

In sostanza, dunque, con i Big Data cambia il modo di interpretare la realtà: non si indagano più le relazioni causa-effetto, ma le correlazioni, essendo sufficiente conoscere come le variabili si muovono congiuntamente.

1.4. BIG DATA: QUAL È IL LORO VALORE?

È difficile stimare in maniera precisa il valore apportato dai Big Data ai vari ambiti applicativi nei quali si possono impiegare, ma essi certamente permettono un miglioramento della qualità delle previsioni, dunque forniscono strumenti utili a prendere decisioni più oculate e supportate da evidenze empiriche più consistenti.

Il valore dei dati nasce, secondo Cukier e Mayer-Schönberger, oltre che da un utilizzo primario per estrarre informazioni e fare previsioni, anche dai possibili utilizzi secondari ricavabili se si ricombinano i dati. In sostanza, il valore secondario deriva dal possibile riutilizzo futuro: dalla fusione di più *dataset*, che permettono di estrarre informazioni combinate e arricchite, dall'identificazione delle possibilità di estensione e generalizzazione delle informazioni ricavabili.

Tuttavia, il valore dei dati è soggetto a una rapida obsolescenza: essi perdono affidabilità con il passare del tempo, a causa della rapidità con cui mutano i trend. Inoltre, la presenza nei *dataset* di dati obsoleti fa perdere valore all'intero set di dati: è perciò necessario utilizzare dei modelli sofisticati che permettano di ripulirli dai dati vecchi.

I Big Data schiudono enormi possibilità in svariati ambiti applicativi e permettono di creare valore in diversi settori economici. Secondo il già citato report del McKinsey Global Institute, siamo ad un punto di svolta nell'ondata di innovazione, crescita e produttività che ha coinvolto imprese, consumatori, interi settori economici guidati dall'utilizzo dei Big Data. Le economie di scala e di scopo nei cambiamenti che i Big Data hanno innescato nell'ambiente economico stanno aumentando in fretta, grazie all'accelerazione e alla convergenza di alcuni trend tecnologici.

I dati stanno diventando uno dei fattori produttivi, dei materiali grezzi basilari per l'impresa, sostiene Kenneth Cukier in "*Data, data everywhere*", (The Economist, 2010). In effetti, i modi in cui i dati apportano valore sono i più diversi: il report del McKinsey Global Institute (2011) ne individua parecchi.

In primo luogo, i Big Data permettono di creare maggior trasparenza, rendendo i dati disponibili ad un maggior numero di *stakeholder* in tempi più rapidi, riducendo sprechi di tempo per ricercare informazioni ed elaborarle.

In secondo luogo, si potrà segmentare la popolazione e rendere le azioni rivolte ai singoli più specifiche e di tipo *one-to-one*: aspetto, questo, particolarmente utile sia nell'ambito del marketing, per offrire ad ogni consumatore ciò che desidera, sia nel settore pubblico, che potrà fornire servizi ad hoc ad ognuno, anziché standardizzare, impiegando al meglio le risorse.

Inoltre, i Big Data costituiscono un ottimo strumento nel campo del *decision-making*: il supporto dei dati rende le scelte più oculate minimizzando i rischi e riducendo gli errori legati all'affidarsi *in toto* all'intuizione. Altre volte, i dati potranno fornire migliori previsioni, date dal semplice analizzare un quantitativo molto maggiore di dati.

I Big Data permettono anche la nascita di nuovi modelli di business basati sull'analisi dei dati: è il caso di numerose *start-up* che forniscono alle aziende il *know-how* per analizzare i propri dati ed estrarne il valore. Inoltre, prevedendo trend e tendenze, i Big Data permettono, grazie ad adeguate analisi, di sviluppare nuovi prodotti e servizi o di migliorare quelli già in circolazione, attraverso, ad esempio, servizi post-vendita più completi. È possibile anche per varie aziende sviluppare un business accessorio legato all'analisi e all'utilizzo dei big data, che affianchi il *core business* dell'impresa, creando un'ulteriore fonte di ricavi.

Una questione connessa con l'utilizzo e il possesso dei dati da parte delle imprese riguarda il valore da attribuire agli *asset* intangibili, come i *dataset*, il *know-how* per analizzarli ed estrarne valore, la capacità di analisi dell'impresa. Come determinare tale valore? Come valutarlo e iscriverlo nel bilancio di un'azienda? Si tratta di questioni tuttora aperte, che richiedono un'analisi approfondita del fenomeno e indicazioni su come stimare tale valore.

Il report del Mc Kinsey Global Institute (2011) ha elaborato alcune stime sul valore da attribuire ai big data: si valuta in 350 miliardi di euro l'anno il valore potenziale apportato al settore pubblico in Europa dall'utilizzo dei big data, 300 miliardi di dollari quello apportato all'ambito medico negli USA. Invece, il potenziale surplus per i consumatori se si utilizzassero in tutto il mondo i dati sulla posizione è pari a 600 miliardi. Inoltre, i *retailer* potrebbero aumentare il proprio reddito operativo del 60% utilizzando i big data.

1.5. BIG DATA: QUALI UTILIZZI POSSIBILI?

I campi in cui i Big Data possono trovare applicazione sono molteplici: essi portano benefici alle aree del business, come si approfondirà in seguito, ma anche al settore pubblico, alla medicina (in particolare alla genetica e all'area delle prevenzioni e screening), all'astronomia e alla fisica (grazie all'enorme quantità di dati raccolti sarà possibile avere conoscenze più precise) e a molti altri.

In particolare, i Big Data contribuiscono a grandi cambiamenti nel campo del management: come sottolineano McAfee e Brynjolfsson (2012), grazie ai big data i manager possono “misurare, e dunque conoscere, molto di più riguardo al loro business, e trasferire direttamente questa conoscenza in migliori processi decisionali e performance”.

È tuttavia difficile stabilire se effettivamente l'utilizzo dei Big Data permetta alle aziende di registrare performance migliori: McAfee e Brynjolfsson (2012) hanno notato che, in vari campi, le imprese che rientrano nel miglior 33% nell'utilizzo dei dati per il *decision-making*, ottengono risultati più soddisfacenti in termini sia operativi che finanziari: in particolare risultano essere del 5% più produttive e del 6% più profittevoli dal punto di vista degli investimenti.

1.6. RISCHI E PROBLEMI

I Big Data, dato il rilievo dell'innovazione apportata, possono comportare rischi e delicati problemi da affrontare con serietà e prontezza per governare un fenomeno di tali dimensioni.

Una prima questione da affrontare riguarda la possibilità di conservare tutte le informazioni generate: esse infatti “eccedono già lo spazio disponibile per l'archiviazione” (Cukier, 2010). Ciò significa che le informazioni disponibili crescono con una velocità maggiore rispetto a quella con cui aumenta lo spazio per archivarle: per questo motivo una delle sfide che la tecnologia dovrà affrontare nel breve e medio periodo riguarda la necessità di aumentare la capacità di immagazzinare informazioni dei dispositivi tecnologici.

In secondo luogo, un altro aspetto ricco di questioni è quello della tutela della privacy: i dati che si raccolgono spesso contengono informazioni riservate e dati personali e permettono di ricostruire molti dettagli della vita privata di ciascuno di noi, ad esempio soltanto sulla base delle ricerche effettuate tramite Google.

È pertanto necessario approfondire queste tematiche dal punto di vista giuridico e provvedere a stabilire delle norme e dei criteri a tutela dei singoli, in modo che i cittadini siano

consapevoli dei possibili utilizzi, primari e secondari, dei dati che li riguardano, proteggendoli e tutelandoli da usi impropri e inadeguati.

Un altro rischio che si corre nell'ambito dei Big Data è quello di affidarsi ciecamente ai dati nel prendere le decisioni, di sottoporsi inconsciamente a una sorta di "dittatura dei dati" (Cukier, Mayer-Schonberger, 2013). La tendenza odierna è quella di raccogliere sempre più dati, ma è necessario considerare che questa non è sempre la soluzione ottimale, in quanto i dati raccolti possono essere sbagliati, fuorvianti, di cattiva qualità. Inoltre, possono essere misurati in maniera impropria: per questo motivo occorre prestare particolare attenzione, anche considerando le potenziali conseguenze di un utilizzo improprio dei dati.

Capitolo 2: Big Data e Management

2.1. BIG DATA: NUOVA RISORSA PER LE AZIENDE

Il management è uno dei campi che sono maggiormente influenzati dalla rivoluzione conseguente all'impiego dei Big Data. Infatti, i Big Data permettono di cambiare l'approccio nella gestione aziendale: si passa dall'*intuitus* del top management ad una cultura del *decision-making* supportata dai dati, cioè da una visione in cui un ruolo importante è rivestito dall'esperienza e dalla capacità di fare previsioni corrette sugli sviluppi del settore da parte dei vertici aziendali, ad una concezione secondo la quale i principali driver nel prendere le decisioni saranno i risultati delle analisi dei dati effettuate dagli analisti ed esperti del settore (per questo motivo si prevede un forte aumento della domanda di statistici e *data analyst* da inserire nelle aziende).

In ambito aziendale, i Big Data sono utilizzati principalmente per due ragioni. In primo luogo, a scopo analitico, in quanto forniscono alle imprese utili intuizioni e analisi sulle loro performance e dei supporti quantitativi per il *decision-making*.

In secondo luogo, i Big Data permettono alle aziende di sviluppare applicazioni e servizi in tempo reale che si avvalgono delle imponenti moli di dati digitali per creare valore per il cliente finale, valore che sarebbe impossibile ottenere altrimenti. (Pembe Muhtaroglu, Demir, Obali, Girgin, 2013). Alcuni esempi pratici sono i servizi post vendita di manutenzione e riparazione dei prodotti, tali per cui, grazie a dei sensori presenti all'interno del prodotto, si riesce a stabilire qual è il momento più adeguato per intervenire ed evitare rotture.

Tuttavia, nonostante i risultati positivi che si riscontrano dell'utilizzo dei Big Data, la maggior parte delle aziende non ha ancora implementato in modo sistematico delle tecniche di analisi dei dati per la gestione aziendale, sia per la mancanza di un adeguato *know-how* nella gestione e utilizzo effettivo di grandi moli di dati, sia per difficoltà di tipo tecnico. Ad ogni modo, date le tendenze attuali, si prevede che l'utilizzo dei big data sarà sempre più pervasivo in tutti i livelli delle organizzazioni.

2.2. COSA CAMBIA CON I BIG DATA?

Per quanto riguarda il *decision-making*, quella a cui stiamo assistendo è una trasformazione che coinvolge tutte le sfere del business. Se i dati sono scarsi o difficili da ottenere e analizzare, ci si affida principalmente alla capacità del top management di fare previsioni

sulla base dell'esperienza e del loro intuito. Tuttavia, ancora oggi, secondo McAfee e Brynjolfsson (2012), "le persone si affidano troppo all'esperienza e all'intuizione e non abbastanza ai dati".

Secondo gli autori, gli organi esecutivi dovrebbero innanzitutto chiedersi cosa dicono i dati e interrogarsi sulla loro provenienza, affidabilità e sul tipo di analisi che è stata condotta; in secondo luogo dovrebbero lasciarsi guidare nelle decisioni dalle evidenze portate dai dati stessi. Il ruolo del leader cambierà, e richiederà la capacità di porre la domanda giusta, anziché quella di dare una risposta.

Il processo decisionale viene dunque cambiato: si introduce la scienza nel management. (Manyka et al., 2011). I manager hanno ora a disposizione una sorta di processo scientifico sul quale basarsi per prendere le decisioni: si parte dalla formulazione di ipotesi, l'effettuazione di esperimenti per testarle, l'analisi rigorosa dei risultati quantitativi prima di prendere una decisione. Ci si basa dunque su risultati empirici (Manyka et al., 2011).

Implementare in azienda delle tecniche di analisi dei dati che fanno riferimento ai big data impone la valutazione di alcune conseguenze fondamentali, esplicate in seguito:

Nuove sfide per il management

Il management è posto di fronte a vari cambiamenti, in diversi aspetti. I leader dai vari team dovranno essere in grado di porsi le domande giuste, di definire obiettivi chiari. Anche utilizzando i Big Data, ci sarà bisogno di leader che sappiano cogliere le nuove opportunità e gli sviluppi del mercato, elaborare progetti di ampio respiro e coinvolgere in essi le persone e gli *stakeholders* in maniera costruttiva per realizzarli. Per raggiungere tali obiettivi, i leader hanno bisogno di valorizzare e al tempo stesso comprendere il valore dei dati che possiedono o ai quali hanno accesso. È fondamentale catalogare i dati, inclusi quelli ai quali si ha potenzialmente accesso (dati pubblici o acquistabili), per rendersi conto del loro valore effettivo, e successivamente integrare dati di proprietà con quelli provenienti dall'esterno (Manyka et al., 2011).

Cambia il ruolo del *decision-maker*, che richiede la capacità del leader di creare un'organizzazione flessibile e in grado di massimizzare la collaborazione tra le varie funzioni aziendali (McAfee, Brynjolfsson, 2013.).

Creazione di valore, opportunità e rischi

Un secondo aspetto fondamentale che i manager dovranno esaminare è quello dell'identificazione delle potenzialità di creazione di valore e dei rischi da tenere in considerazione, controllando periodicamente l'ambiente circostante e conducendo esperimenti utilizzando i big data, in modo da implementarli direttamente nei processi aziendali, anziché “calare dall'alto” modelli già predefiniti.

Per creare valore, non è necessario utilizzare fin dall'inizio complessi sistemi di analisi dei dati; al contrario è bene procedere per gradi al fine di costruire nel tempo una solida base.

Il primo passaggio consiste nel digitalizzare i dati, strutturarli e organizzarli in modo che siano utilizzabili per le analisi.

Successivamente, ci si occupa di rendere i dati disponibili, integrando vari *dataset* per avere informazioni più complete; il terzo livello di sofisticazione prevede di applicare tecnologie base di analisi dei dati, con tecniche abbastanza standard, che non richiedono competenze specialistiche. Infine, il livello più avanzato permette di applicare sofisticate analisi dei dati, come algoritmi automatici e analisi dei dati in tempo reale, che permettono di ottenere nuove prospettive e modelli (Manyka et al., 2011).

Il valore che le aziende possono ottenere deriva dunque non soltanto dall'utilizzo di dati propri, ma anche dall'aggregazione dei *dataset* a disposizione con altri dati, pubblici o ottenuti da terze parti a pagamento; inoltre, con un sistema efficiente di analisi dei dati sviluppato internamente, le imprese possono ambire a costruire un modello di business accessorio che affianchi il *core business*.

Costruire le competenze

È fondamentale per le imprese reclutare team di *data scientist*, *computer scientist* e statistici in grado di lavorare con grandi quantità di dati, al fine di avere analisi corrette e chiare a disposizione del proprio business, figure professionali che, secondo le previsioni, saranno sempre più richieste. Questo non è tuttavia sufficiente: è necessario infatti formare i manager e i dipendenti al nuovo approccio e alle nuove metodologie decisionali portate dai dati, facendo sì che vi sia collaborazione e una comunicazione efficace tra il team tecnico e i *decision-makers*.

Implementare una nuova tecnologia

Un aspetto importante è la tecnologia a disposizione dell'azienda: i costi per accedervi si sono molto abbassati negli ultimi anni, e non sono particolarmente costose: infatti molti software per analizzare i dati sono *open source* (come ad esempio *Hadoop*, il rivale open source di

MapReduce). Molte aziende avranno bisogno di investire nuovamente in software, hardware e servizi in grado di analizzare grandi set di dati.

La tecnologia è una condizione necessaria ma non sufficiente per implementare una strategia basata sui Big Data, in quanto non deve mancare una visione d'insieme della strategia aziendale, che deve essere focalizzata sul cliente. È opportuno però iniziare a implementare le nuove strategie con progetti precisi e di ampiezza limitata al fine di costruire delle competenze solide basandosi su procedure che funzionano su piccola scala per poi espanderle in progetti di dimensioni più ampie.

Problemi connessi all'utilizzo dei dati

Le implicazioni connesse alla privacy e alla sicurezza sono di primaria importanza per vari scopi. In particolare, l'attenzione alla tematica della privacy è fondamentale non solo per il rispetto delle leggi e dei regolamenti, ma anche per la costruzione di una relazione di fiducia con i clienti, partner, dipendenti e altri *stakeholder*. A questo scopo, le imprese devono rendere nota la loro politica di gestione dei dati, soprattutto nei confronti dei clienti, sempre più attenti a tali tematiche.

Le organizzazioni devono inoltre implementare strategie che proteggano dai rischi che incontrano le funzioni IT, dalla possibilità di rotture fisiche al pericolo degli hacker. Si hanno tuttavia delle soluzioni di tipo informatico che permettono di proteggersi da tali rischi connessi alla privacy e alla sicurezza.

È necessario prestare particolare attenzione a queste tematiche, che spesso richiedono consulenze legali, ma anche di considerare la strategia aziendale, le relazioni con i clienti, i dipendenti e gli altri partner.

Dalle sfide che vanno affrontate emerge la necessità di un cambio di approccio e di mentalità, che coinvolge tutte le funzioni aziendali. In particolare, con l'implementazione delle tecniche di analisi basate sui big data, ci si chiede innanzitutto cosa dicono i dati, quali sono le variabili e i fattori correlati tra loro e come questi cambiamenti congiunti condizionino l'operato dell'impresa.

2.3. LA RIVOLUZIONE DEI BIG DATA: VARIE REAZIONI

Non tutte le aziende hanno però lo stesso modo di affrontare la questione dei Big Data e dell'analisi dei dati: si notano infatti, secondo uno studio (Marshall, Mueck, Shockley, 2015),

si individuano tre modi di utilizzare i Big Data da parte delle aziende per innovare, ottenere vantaggio competitivo e creare valore.

Si identificano pertanto un primo gruppo di Leader, innovatori, che usano i Big Data e l'analisi dei dati per innovare attraverso un approccio strutturato; un secondo gruppo di imprese che, benché investa in strumenti a supporto dell'innovazione, è meno certo su quali siano le attività più importanti a tale scopo; infine, un terzo gruppo, il più debole da questo punto di vista, più avverso al rischio e meno portato a innovare.

L'analisi condotta rivela che le aziende che nel processo di innovazione usano i big data e l'analisi dei dati hanno una probabilità più alta del 36% di battere i competitor in termini di crescita del guadagno e dell'efficienza interna. La capacità delle imprese che realizzano delle performance migliori sta soprattutto nell'estrapolare dati da diverse fonti e nel tradurli in risultati concreti grazie ad approfondite analisi.

Le aziende leader nell'innovazione condividono alcune strategie su dati, *skills* e cultura aziendale nell'integrare innovazione e analisi dei dati. In primo luogo, si promuovono la qualità e l'accessibilità dei dati in tutta l'azienda: queste imprese riconoscono il valore dell'analisi dei dati per avere una migliore conoscenza del cliente e sono migliori nel dedurla. Inoltre, utilizzano questo tipo di analisi per tutto il processo di innovazione e lo applicano in più aree, sottolineando la connessione tra *analytics* e innovazione, rendendo i dati accessibili a tutta l'organizzazione e curando la conoscenza del cliente e l'innovazione.

In secondo luogo, nelle imprese leader si insiste sull'importanza di formare il personale all'utilizzo dei big data, per carpirne il valore e sfruttare il vantaggio generato dalla collaborazione tra le varie funzioni aziendali.

In terzo luogo, caratteristica delle imprese leader in questo campo, è quella di costruire una cultura dell'innovazione in tutta l'impresa e di incoraggiare comportamenti che rinforzino la collaborazione, la creatività e immaginazione. Un altro aspetto importante è la capacità di valutare e misurare il profitto generato dall'innovazione: l'investimento di per sé non è sufficiente, ma occorre avere un feedback sulla sua efficacia.

2.4. COME SFRUTTARE I BIG DATA IN MODO PROFITTEVOLE?

L'utilizzo dei Big Data pervade tutta l'organizzazione: per sfruttare al meglio i dati e le loro potenzialità essi dovrebbero essere accessibili ai vari reparti dell'impresa e non appannaggio

di alcune funzioni aziendali. Infatti, i big data si possono utilizzare lungo tutta la catena di creazione del valore: dall'ottimizzazione dei rapporti di fornitura e della produzione, al marketing e alla gestione dei servizi post-vendita quali manutenzione e riparazione. Per esempio, dal punto di vista del marketing i big data permettono di migliorare la *marketing intelligence*, mentre la logistica li utilizza per gestire al meglio i veicoli e programmare in maniera efficiente i percorsi; le operations se ne servono allo scopo di "ottimizzare le richieste della produzione, dallo stock in magazzino alla programmazione della produzione" (Sanders, 2016).

Un approccio costruttivo per sfruttare al meglio i dati che sono forniti dai clienti stessi è quello di orientarsi al cliente, al fine di creare valore per il consumatore finale. Infatti, "un orientamento che non sia centrato sul cliente limita fortemente l'abilità dell'organizzazione di sfruttare le potenti leve dei big data per creare nuovo valore" (Manyka et al., 2011). Per questo motivo, i dati dovrebbero essere disponibili alle diverse funzioni in maniera trasversale e completa, con una struttura omogenea e adeguata a tutte le differenti esigenze delle aree aziendali per essere utilizzati in modo efficace da tutti.

Nel management, l'impatto dei big data è foriero di enormi possibilità per il miglioramento della gestione aziendale, sebbene siano un numero ristretto le aziende che finora ne hanno implementato un utilizzo sistematico. Le aree e funzioni coinvolte in questo cambiamento sono molteplici; in questa sede ci si limita a esaminare più in dettaglio due campi che sono particolarmente interessanti: il supply chain management, che attraversa l'intero processo produttivo in modo trasversale, e il marketing, area che sta vivendo un'evoluzione rapidissima grazie alle nuove tecnologie.

2.5. BIG DATA E SUPPLY CHAIN MANAGEMENT

Le potenzialità fornite dai Big Data sono utilizzabili in varie aree aziendali, in modo particolare nella gestione della supply chain.

Si definisce *supply chain management* "la gestione delle interconnessioni tra organizzazioni che si relazionano a monte o a valle per la creazione di valore per il cliente finale", in un "approccio olistico" di management oltre i confini delle singole organizzazioni (Slack, Brandon-Jones, Johnston, 2013). La gestione della supply chain ha lo scopo di rendere i processi efficienti lungo la catena del valore di un'azienda, e coinvolge tutte le funzioni *core* (ricerca e sviluppo, operations e produzione, marketing e vendite) ma anche quelle di

supporto, dalla gestione delle merci in entrata, all'efficienza della produzione, a una distribuzione efficace e a una buona gestione dei servizi post-vendita. Tutti i processi contribuiscono alla creazione di valore per il cliente finale: a questo scopo una visione trasversale del flusso di valore che si crea attraverso i vari passaggi è fondamentale, per far sì che l'efficienza e l'efficacia nei processi siano rivolte alla soddisfazione del cliente. La supply chain costituisce infatti un sistema, formato da diverse funzioni che devono agire per un unico scopo: per questo motivo non possono essere colte singolarmente, ma devono essere interpretate in un'ottica integrata.

Utilizzare i Big Data può essere uno strumento utile per favorire questo tipo di approccio: per avere dei dati utili e completi su cui basarsi occorre l'apporto di molte variabili differenti, disponibili alle diverse funzioni aziendali; una volta elaborate, queste informazioni devono essere rese accessibili a tutte le aree dell'azienda: a tale scopo è pertanto necessaria una comunicazione costruttiva ed efficace tra i vari segmenti della supply chain e la concezione della stessa come un continuum nella creazione del valore per il cliente finale.

Infatti, lungo tutta la supply chain si “genera un immenso ammontare di dati, che possono dar luogo a indizi importanti attraverso le tecniche di analisi dei dati” (Sanders, 2016). Le informazioni che si ricavano possono contribuire al miglioramento dei processi, alla riduzione di sprechi e alla creazione di ulteriore valore, come dimostrano gli esempi di aziende come Amazon e Walmart, che hanno implementato con successo tali tecniche. Tuttavia, la maggior parte delle imprese incontra delle difficoltà nel “far leva sull'analisi dei big data per trasformare la propria supply chain” (Sanders, 2016): infatti, se per alcune imprese la tecnologia e i software necessari per questo genere di analisi è un investimento costoso e dal rendimento incerto, altre manifestano difficoltà nell'impiegare i dati per gestire la supply chain nonostante la grande quantità di dati disponibile, altre ancora cercano sì di implementarne l'utilizzo, ma ottenendo risultati frammentati e non facenti parte di una strategia più completa e a lungo termine.

Lungo tutta la supply chain vi sono molteplici possibilità di implementare strategie di analisi dei dati, nelle principali fasi di reperimento, produzione, trasferimento e vendita del prodotto. Le applicazioni dei big data nella fase del reperimento di risorse sono quelle che si stanno sviluppando più velocemente: per molte imprese manifatturiere il reperimento di materie prime costituisce la spesa più consistente per l'azienda, variando dal 50% al 90% dei ricavi: l'analisi dei dati può costituire un ingente risparmio, in quanto permette di ottimizzare i canali di fornitura e integrarsi al meglio con i propri fornitori. L'analisi dei dati è utilizzata

principalmente per determinare i mix e le quantità di rifornimento e scorte necessari per far fronte alle esigenze della produzione, ma anche a supporto delle contrattazioni con i fornitori, “fornendo una completa analisi delle preferenze del cliente e dei comportamenti d’acquisto” (Sanders, 2016).

Nella seconda fase della supply chain, ossia quella delle operations e della produzione, le applicazioni dei big data permettono di migliorare e rendere più efficiente la gestione del magazzino e delle scorte, di ottimizzare la manutenzione e la localizzazione dei servizi. È possibile monitorare la produttività della forza lavoro, migliorandola e programmando la produzione in modo più adeguato. Inoltre, si può ottenere un approfondimento ancora maggiore e disponibile quasi in tempo reale dei dati, potendo così intervenire tempestivamente.

La fase della logistica e della distribuzione è stata una delle prime ad essere influenzata dalle applicazioni offerte dai big data, soprattutto nell’ottimizzare il livello dello stock a magazzino, nell’individuare i luoghi ottimali per i centri distributivi e i percorsi per recapitare i prodotti in modo tale da minimizzare i costi di trasporto. Grazie ai big data, si può aumentare l’efficienza nei consumi di carburante, con la manutenzione preventiva dei mezzi di trasporto (alcuni sensori segnalano quando è il momento di effettuare la manutenzione, prevenendo rotture), ottimizzando i comportamenti dell’autista e dei percorsi seguiti in tempo reale.

Le applicazioni nel campo del marketing e delle vendite sono orientate al cliente, per catturarne la domanda attraverso tecniche quali la micro segmentazione e la previsione dei comportamenti del consumatore. Le aziende possono ottenere molte più informazioni sui clienti e tener traccia dei loro comportamenti anche in tempo reale, in modo tale da potersi adattare alle mutevoli esigenze del consumatore. Un’altra importante applicazione del marketing è l’ottimizzazione del prezzo in base ai diversi clienti, in modo da procurare l’offerta migliore per ciascuno di essi.

È fondamentale impiegare i Big Data integrando le diverse funzioni aziendali in maniera armonica. Oggi si hanno nuove opportunità per le indagini, che permettono di essere svolte a una profondità molto maggiore e più specializzata: infatti, le imprese leader nell’innovazione e nel cambiamento adottano un approccio che integra tra loro le funzioni aziendali.

Implementare questi cambiamenti è spesso difficile per le imprese, che incontrano di frequente varie difficoltà. In particolare, emerge in primo luogo la tendenza a non focalizzarsi

sui problemi, ma a cercare relazioni causali tra le variabili senza avere la piena consapevolezza della direzione da prendere e senza ottenere risultati utili. In secondo luogo, vi è il rischio di sviluppare sì ottime applicazioni, ma che esse restino isolate e non collegate con le altre, in quanto non si agisce integrando le varie funzioni, ma mantenendole separate e isolate. In terzo luogo, si rischia di cadere nella trappola di voler misurare troppi aspetti e di non riuscire a mantenerli sotto controllo tutti, rimanendo in un certo senso travolti dalla sovrabbondanza di dati e non essendo in grado di gestirli (Sanders, 2016).

Le aziende leader nell'innovazione in questo campo sono riuscite a far sì che le nuove applicazioni fossero di supporto alla loro *value proposition*, coordinando gli sforzi a tutti i livelli della supply chain per migliorare l'intera strategia aziendale e monitorando i risultati in un'ottica di progresso continuo.

Si possono dunque individuare, grazie agli esempi positivi forniti da alcune imprese, degli step comuni seguiti per implementare strategie positive.

Inizialmente, si creano dei segmenti all'interno della supply chain con degli attributi specifici e si stabilisce come essere competitivi in ciascuno di essi. Ognuno di essi si focalizza su un obiettivo specifico, ottimizzando i bisogni del cliente e le necessità della supply chain.

Il successivo step consiste nell'adeguare le funzioni organizzative, integrando i processi tra i vari livelli della supply chain e focalizzandosi sulla competitività nel segmento scelto.

Successivamente, si misurano le performances ottenute, il grado di integrazione tra i vari processi, di cooperazione tra le funzioni aziendali e l'effettivo appeal della value proposition dell'impresa. In un'ottica *kaizen*, cioè di miglioramento continuo, il feedback che si ottiene permette di ricalibrare gli sforzi nel caso in cui i risultati non fossero soddisfacenti o di avere conferma della giusta direzione intrapresa. (Sanders, 2016.)

2.5.1. ESEMPIO - DELL

Vi sono vari esempi di aziende che hanno implementato con successo una strategia di *big data analytics* nel gestire la propria supply chain. Una di queste, il cui esempio è riportato nello studio di Sanders (2016), è Dell, azienda tra le principali produttrici di PC al mondo. Il modello di business di Dell è detto *configure-to-order*: si basa infatti sulla vendita di prodotti *customizzati*, che il cliente configura in base alle proprie esigenze, potendo variare software, memoria, schermo, design e ogni altro aspetto personalizzabile, senza ricorrere a intermediari: gli ordini infatti si effettuano direttamente dal sito dell'azienda.

Una strategia di questo genere, che permette al cliente di personalizzare il prodotto a seconda delle proprie esigenze e di ridurre i tempi di consegna grazie all'eliminazione degli intermediari, è apprezzata dal consumatore finale. Tuttavia, si riscontravano notevoli difficoltà nella gestione del processo di approvvigionamento e di assemblaggio per creare i prodotti ad hoc: infatti le combinazioni possibili sono dell'ordine di 10^{24} , quantità impossibile da gestire.

Dal momento che il modello di business basato sulla personalizzazione del prodotto è ciò che permette a Dell di differenziarsi rispetto ai suoi competitors, soddisfacendo i bisogni di un'ampia varietà di clienti e al tempo stesso minimizzando i costi, l'azienda ha deciso di mantenerlo invariato e di implementare un sistema di gestione basato sull'analisi dei dati per semplificare la gestione della supply chain.

Avendo a disposizione i dati storici degli ordini già effettuati, Dell è riuscita a scoprire quali combinazioni hanno un maggior numero di componenti in comune. Di conseguenza, è stato possibile ridurre le configurazioni possibili a qualche milione.

Il passo successivo consiste nella segmentazione delle opzioni più scelte dai clienti: una volta identificate, è possibile coordinare tutta la supply chain in modo da ridurre i tempi di assemblaggio e poter stoccare i prodotti in magazzino. Tutto ciò si traduce sia in una riduzione dei costi per Dell, che riesce a velocizzare i suoi processi riducendo i tempi di permanenza delle scorte a magazzino e preparando alcune tra le combinazioni più popolari già pronte in magazzino, in modo da essere più veloci nel recapitare il prodotto al cliente, arrivando a ridurre i tempi di consegna a un solo giorno di attesa. Tali riduzioni di costo costituiscono dunque allo stesso tempo un incremento del valore aggiunto per il cliente.

L'innovazione implementata ha permesso a Dell di incrementare i ricavi di 40 milioni di dollari annui.

2.5.2. ESEMPIO – WALMART

Un esempio, riportato nello studio di Sanders (2016), di un'azienda leader nella gestione della supply chain attraverso i big data è fornito da Walmart. Si tratta del più grande *retailer* al mondo, con “oltre due milioni di dipendenti e un fatturato annuo intorno ai 450 miliardi di dollari” (Cukier, Mayer-Schönberger, 2013).

Walmart è leader anche nell'applicazione di tecnologie analitiche, che le permettono di collegare al meglio le funzioni tra loro e coordinare l'intera *supply network*, sia di fornitori

che di distributori, estesa in tutto il mondo, date le dimensioni del colosso fondato da Sam Walton.

Walmart raccoglie più di 1 milione di transazioni all'ora, che vanno ad alimentare database della dimensione di circa 2,5 petabytes (10^{15} byte) (The Economist, 2010): l'azienda possiede una quantità di dati enorme, che riesce a sfruttare lungo tutta la *supply chain*, attraverso sofisticate tecniche analitiche, fornendo dati utili ai manager dei diversi livelli.

In particolare, questo metodo ha permesso a Walmart di apprendere molto di più riguardo alle preferenze dei consumatori e ai loro comportamenti d'acquisto: per esempio, negli Stati Uniti, dall'analisi dei consumi dei clienti poco prima che si scateni un uragano, emerge un aumento delle vendite di merendine alla fragola; conoscendo questa tendenza, i commessi di Walmart possono provvedere a disporle negli scaffali all'ingresso, "assieme ad altre attrezzature anti-uragano, per facilitare il compito ai clienti che entrano e escono di corsa - facendo lievitare il fatturato" (Cukier, Mayer-Schonberger, 2013).

Walmart si serve della gran quantità di dati a sua disposizione per collaborare in maniera proficua con i suoi fornitori: ha messo a punto un sistema, chiamato "Retail Link", che consente di tenere sotto controllo i movimenti dei prodotti. Walmart condivide con i suoi fornitori (più di 17400 in 80 Paesi) le informazioni riguardanti i prodotti, i loro spostamenti e i livelli delle scorte: in questo modo i fornitori non devono attendere un ordine da parte di Walmart ma possono prevedere il momento in cui è necessario rifornire il *retailer* e pianificare la produzione di conseguenza.

Inoltre, il sistema permette ai fornitori di avere informazioni sull'assortimento e sulle quantità richieste per ogni punto vendita, sulla base dei consumi e delle previsioni: i fornitori possono dunque regolare il mix di prodotti richiesti da Walmart in base alla domanda del consumatore finale.

Emerge da questo esempio come Walmart consideri la *supply chain* un *continuum* nella creazione del valore, e sfrutti i big data e l'analisi dei dati al fine di collegarne al meglio le varie fasi, dalla produzione alla vendita. Infatti, le informazioni viaggiano lungo la *supply chain* per informare tutte le altre funzioni: si raccolgono dati di vendita dai POS e informazioni sullo stock da sensori RFID (*Radio-Frequency Identification*), che vengono poi collegati e resi disponibili ai diversi livelli della *supply chain* dalla piattaforma Retail Link. In questo modo, si possono intraprendere azioni coordinate e si ottimizzano le decisioni, risparmiando sui costi e migliorando il servizio per il cliente.

2.6. BIG DATA E MARKETING

Un ambito che sta subendo veloci trasformazioni grazie all'utilizzo dei big data è quello del marketing. Attraverso internet, infatti, è aumentata a dismisura la quantità di dati riguardanti i clienti che sono disponibili per le aziende: non solo i dati anonimi ricavabili dalle ricerche effettuate su internet, ma anche i dati forniti alle aziende dai clienti stessi, per esempio attraverso l'iscrizione alle newsletter, il numero di carta di credito ad essi collegato per fare acquisti online o le informazioni riguardanti dati personali e comportamenti d'acquisto fornite nella creazione di un account nei siti di *e-commerce*.

I *retailer* sono sempre più esperti nell'estrapolare indicazioni utili dai dati che raccolgono dai canali di vendita, cataloghi, negozi e interazioni online. L'utilizzo di dati riguardanti i singoli clienti permette ai *retailer* e ai venditori di essere più efficaci nel marketing e nelle vendite: i big data permettono di ridurre i costi e creare vantaggi competitivi per accrescere il margine (Manyka et al., 2011).

Al fine di ricavare dai dati informazioni utili e dunque valore, essi devono essere organizzati e rielaborati: I dati "grezzi" risultano spesso confusi e difficilmente interpretabili. Nella fase iniziale di adozione dell'analisi dei dati risulta spesso difficile avere le competenze per analizzare i dati in modo sufficientemente tempestivo ed efficace, perciò si preferisce esternalizzare la fase detta di "*data science*", cioè la parte analitica vera e propria, affidandola spesso a consulenti o ad aziende che possiedono l'adeguato *know-how*. Tuttavia, è importante mantenere la proprietà dei dati, al fine di renderli sempre disponibili per tutte le aree aziendali e poterli gestire meglio (Artun, Levin, 2015).

Nuove tecniche per il marketing

Per il marketing, in particolare, emergono dallo studio del *Mc Kinsey Global Institute* varie opportunità per incrementare i ricavi, utilizzando delle applicazioni pratiche che derivano dai big data. Tra i principali metodi per aumentare le vendite si ricordano il *cross-selling*, il *location-based marketing*, la *micro-segmentazione*, la *sentiment analysis* e la *multi-channel experience*.

Il *cross-selling*, in primo luogo, utilizza tutti i dati riguardo al cliente (demografici, su acquisti passati, sulla posizione attuale) per aumentare le dimensioni dell'acquisto.

Il *location-based marketing*, invece, si basa sulle informazioni fornite dagli smartphones riguardo alla posizione dei clienti: individuando il luogo in cui essi si trovano, è possibile rivolgersi direttamente a coloro che sono nelle vicinanze o all'interno dei negozi con offerte speciali, creando un legame tra shopping fisico e esperienza online. Inoltre, l'analisi del comportamento all'interno del punto vendita costituisce un'importante fonte di informazioni per decidere come meglio strutturare lo *store*, dal punto di vista del layout, del product mix e del posizionamento fisico dei prodotti. Tali analisi si possono effettuare tracciando i percorsi dei consumatori nel punto vendita con immagini che derivano dalle telecamere di sorveglianza o con le informazioni fornite dagli smartphones.

I big data hanno permesso anche un'evoluzione notevole della *micro-segmentazione* dei clienti: l'enorme aumento di dati disponibili e le nuove possibilità offerte dall'evoluzione tecnologica permettono ai *retailer* di concentrarsi sul marketing *one-to-one* e sulla personalizzazione, anche in tempo reale, aiutati altresì dalle informazioni sugli acquisti passati del singolo cliente e sul comportamento online.

Anche la *sentiment analysis* è uno strumento utile per il marketing: basandosi sull'enorme quantità di dati provenienti dai clienti dai vari social media permette di prendere delle decisioni congruenti con ciò che esprimono i clienti. Per esempio, è possibile conoscere le opinioni di una campagna pubblicitaria in tempo reale ed eventualmente correggerla se non suscita le reazioni desiderate. La *sentiment analysis* sta accrescendo la sua importanza in quanto sempre più frequentemente i consumatori si affidano a raccomandazioni di conoscenti o di *influencer* indipendenti.

I marketing manager possono sfruttare anche la leva dell'intensificazione della *multi-channel experience* per i clienti, che permette di aumentarne la soddisfazione, la *loyalty* e di far lievitare le vendite. Infatti, i *retailer* possono integrare le promozioni, per i vari canali di vendita a disposizione (online, *in-store*, nei cataloghi online e cartacei), adattando le offerte ai singoli clienti in modo più efficace.

Anche i *retailer* fisici possono trarre vantaggio dall'utilizzo dei big data, per esempio grazie all'ottimizzazione dell'assortimento in base a dati demografici e percezioni dei consumatori si possono aumentare sensibilmente i ricavi. Un altro strumento a disposizione dei *retailer* è l'ottimizzazione dei prezzi, che può essere ottenuta grazie alla crescente disponibilità di dati, ai nuovi modelli che si possono elaborare e ai trend che si possono estrarre dai comportamenti

dei consumatori: in tal modo è possibile applicare le politiche di promozione in maniera più individualizzata, per esempio con offerte speciali per i prodotti con maggiore elasticità della domanda.

Google Analytics: utile strumento di marketing e analisi dei dati

Uno dei canali di comunicazione più rilevanti ed efficaci nel campo del marketing è rappresentato dal sito Internet aziendale, che permette all'impresa di presentare un'immagine positiva di sé stessa a clienti, potenziali clienti e a diversi tipi di *stakeholders* aziendali. Per questo motivo, la costruzione del sito aziendale deve essere curata in modo particolare, facilitando il reperimento delle informazioni, aumentando la fedeltà dei clienti e trasmettendo un'immagine dell'azienda che sia coerente con i propri valori e la *value proposition*.

Nel management del sito aziendale è di particolare importanza condurre delle analisi per avere dei feedback sulle performance, per sapere se le azioni e le campagne intraprese sono orientate nella giusta direzione, o se sia necessario correggerle. Un utile strumento per tali analisi è rappresentato da Google Analytics, un pacchetto di servizi gratuiti elaborato da Google: si tratta di uno strumento semplice da utilizzare ma solido nei risultati e adattabile alle più diverse esigenze, permette di tracciare l'utilizzo sia da pc che da dispositivi mobili; è sufficiente creare un account e collegarlo all'indirizzo web di un sito per iniziare a monitorarne alcuni aspetti elementari.

È possibile analizzare le visite e il comportamento dei consumatori nel sito aziendale a vari livelli di profondità e prendendo in considerazione diversi aspetti. Inizialmente, vengono monitorati il numero delle visite, il *bounce rate* (il “tasso di rimbalzo”, cioè la percentuale di visitatori che ha lasciato immediatamente il sito e che non ha visitato più di una pagina), il numero di pagine visitate ad ogni collegamento, il tempo medio di permanenza sul sito per ogni visita, il numero medio di pagine visualizzate per ogni visita (totale pagine visitate/visite totali), la percentuale di nuovi visitatori.

Oltre a queste informazioni generiche è poi possibile ricavare altre informazioni che permettano analisi più complesse: è possibile creare alcuni “Advanced segments” per scoprire eventuali legami tra social media e sito aziendale, in particolare per indagare su quali social media rimandano al sito aziendale e con quale frequenza. È inoltre possibile indagare sulla tipologia di dispositivi, fissi o mobili, dai quali avviene l'accesso al sito.

Un'altra importante funzionalità è quella che consente di ricavare in che percentuale i visitatori del sito completino dei “*funnel path*”, ossia un percorso prestabilito ottimale per il raggiungimento di determinati obiettivi. Di conseguenza, l'azienda potrà migliorare la funzionalità dei processi al fine di incentivare i visitatori a completarli.

È possibile inoltre, ricavare degli spunti su come migliorare il proprio sito in base alle ricerche che vi vengono effettuate (si monitora la percentuale di visitatori che effettua ricerche interne al sito, quali parole sono usate, se si effettuano ricerche avanzate in altre pagine o si abbandonano il sito subito dopo), estraendo indicazioni su come i visitatori usino il sito e cosa si aspettino da esso, in quali punti si nota la necessità di essere più chiari e in quali di arricchire i contenuti.

In sostanza, Google Analytics rappresenta un'importante risorsa per controllare la performance di un aspetto rilevante per le organizzazioni qual è il sito aziendale, permettendo loro di ottimizzarne l'efficacia: si tratta di un esempio pratico di come l'analisi dei dati sia utile per le aziende, poiché fa sì che si direzionino le energie efficacemente, arricchendo il valore aggiunto per i clienti e consentendo vari risparmi di costo alle imprese.

Predictive Marketing: sviluppo di una nuova frontiera del marketing

Un aspetto importante del marketing che con i big data sta conoscendo uno sviluppo rapidissimo è quello del *predictive marketing*, il quale ha lo scopo di offrire al cliente un'esperienza migliore e più costruttiva in tutti i punti di contatto tra cliente e azienda al fine di aumentare la lealtà del cliente e, naturalmente, i ricavi. Esso sostanzialmente si serve dei risultati delle tecniche di analisi dei dati al fine di prevedere il comportamento del cliente, di anticiparne le mosse e i bisogni, proponendogli il prodotto desiderato al momento giusto e creando una relazione più profittevole. A questo proposito, in “*Predictive Marketing. Easy ways every marketer can use customer analytics and big data*”, Ömer Artun e Dominique Levin identificano un importante cambiamento nel marketing nel cambio di prospettiva: non ci si focalizza più sul prodotto o sui canali distributivi, ma sul cliente.

L'utilizzo del *predictive marketing* è dovuto principalmente a tre fattori: innanzitutto dalle richieste dei clienti stessi di un'esperienza di acquisto più coinvolgente, di una maggiore personalizzazione e servizi ad hoc; in secondo luogo per il successo che si riscontra tra i primi adottanti di queste tecniche; in terzo luogo per la disponibilità, ad un costo non più proibitivo, di una tecnologia che consente di ricavare più dati dai comportamenti dei clienti e di ricavare dei trend.

Il *predictive marketing* fa leva sull'analisi dei dati per acquisire nuovi clienti, accrescerne il potenziale valore e trattenerli più a lungo, creando una relazione personale con il cliente: grazie alle informazioni raccolte sul cliente, sia online che offline, infatti, è possibile personalizzare le offerte e le proposte di acquisto per ognuno. Gli strumenti a disposizione del marketing risultano pertanto più efficaci e performanti grazie alle nuove tecnologie di *machine learning* e *big data analytics*.

Gli usi del marketing predittivo permettono di migliorare la precisione del *targeting* e gli sforzi per acquisire nuovi clienti, per massimizzarne il valore potenziale nel rapporto con l'azienda e dunque massimizzare il valore totale dell'intero portafoglio dei clienti. Una relazione profittevole con i propri clienti costituisce infatti un aspetto di notevole importanza per le imprese, nell'ottica di *do ut des* (o di *give-to-get*) che caratterizza la creazione di valore nel marketing. All'inizio di una relazione con il cliente, è prassi in numerose imprese esplicitare al cliente la propria *value proposition* e permettere di farne esperienza attraverso offerte speciali o campioni gratuiti. La spesa iniziale che si sostiene non costituisce un *sunk cost* (costo affondato), ma un investimento che permetterà di costruire un rapporto duraturo e fruttuoso con i clienti, se queste iniziative sono rivolte a individui attraenti per l'impresa, identificati in base ai dati raccolti con le varie tecniche di *analytics*.

Le *predictive analysis* applicano algoritmi e modelli matematici complessi a enormi moli di dati per raggruppare i clienti in modo adeguato e prevederne i comportamenti, in modo da offrire loro i prodotti che cercano, consapevolmente o meno. A questo scopo, è importante raccogliere e aggiornare i dati in maniera continuativa, analizzare le informazioni a livelli micro e macro, utilizzare i dati che si hanno a disposizione per personalizzare le esperienze dei singoli clienti.

Alcune tecniche di predictive marketing

Si utilizzano diverse tipologie e tecniche per effettuare questo tipo di analisi per la previsione dei comportamenti, tra le quali si ricordano *clustering* o *unsupervised learning*, *propensity model* o *supervised learning*, *reinforcement learning* o *collaborative filtering*.

Il *clustering*, o *unsupervised learning* (apprendimento non supervisionato), consiste in una tecnica per raggruppare individui simili tra loro permettendo di scoprire automaticamente, utilizzando appositi algoritmi, quali siano i segmenti in cui si suddivide la *customer base* dell'impresa. Si possono prendere in considerazione centinaia di aspetti, finché l'algoritmo

non individua quelli discriminanti per la segmentazione: l'idea di fondo è che vi sono diversi gruppi statisticamente significativi che si comportano allo stesso modo.

I *cluster* che si ottengono con questo processo sono simili ai gruppi che si determinano attraverso la segmentazione; la differenza fondamentale tra le due metodologie è che, mentre la segmentazione prevede l'attribuzione dei clienti a gruppi già individuati preventivamente, dunque a priori, il *clustering* permette di identificare la disposizione "spontanea" degli individui nei diversi raggruppamenti: si lasciano parlare i dati, le caratteristiche di similarità non sono note in precedenza.

I *cluster* possono essere di diversi tipi, a seconda dei criteri con i quali i clienti sono raggruppati: si evidenziano i *cluster product-based*, in cui si riuniscono clienti che tendono a comprare prodotti o combinazioni di più prodotti della stessa categoria; i *cluster brand-based*, che accolgono clienti che preferiscono determinati *brand* al posto di altri (questo tipo di preferenza tende ad essere più forte rispetto a quella basata sui prodotti); i *cluster behavior-based*, che raggruppano consumatori accomunati da simili comportamenti d'acquisto, aiutando il marketing manager a individuare il modo più adatto per rivolgersi a ciascuno di loro.

Una seconda tipologia di *predictive analytics* è detta *propensity model* (o *supervised learning*): si basa su modelli probabilistici servendosi di tecniche avanzate di *machine learning* quali *neural networks*, *logistic regression*, *random forest*, and *regression trees*, che costituiscono delle metodologie che utilizzano algoritmi sofisticati, di cui non si tratta in questa sede. Il fine principale di queste tecniche consiste nella previsione di comportamenti futuri dei clienti sulla base di esempi passati. Con il tempo, questi algoritmi diventano più efficaci, potendo affinare le previsioni grazie a un maggior numero di dati.

In terzo luogo, il *reinforcement learning and collaborative filtering* costituisce una tecnica molto utilizzata nell'effettuare previsioni; un'applicazione importante è quella delle *recommendations*, ossia dei consigli su quali prodotti acquistare, che compaiono solitamente nella navigazione in Internet. Essi non devono essere troppo invadenti né inopportuni, altrimenti si ottiene l'effetto opposto a quello desiderato. Le *recommendations* sono "targettizzate", cioè sono "su misura" per il cliente, ottenute considerando l'intera relazione tra cliente e brand. Solitamente, i consigli sono del tipo "*upsell*", quando consigliano prodotti di un livello superiore, di tipo "*cross-sell*", quando i prodotti consigliati appartengono alla stessa categoria, di tipo "*next-sell*", se si consiglia il prodotto che in genere i consumatori

comprano successivamente, o di “*products-generally bought together*”, cioè i prodotti che di solito sono comprati assieme.

È importante, nell'intraprendere azioni di marketing, cercare di conoscere il più possibile i propri clienti; il *profiling* che si può ottenere con i big data assolve a questa funzione in modo ottimale. Prima di ottenere i risultati, però, è importante raccogliere i dati e “ripulirli” per prepararli all'analisi, cioè eliminare gli eventuali errori sistematici: a questo scopo è fondamentale la collaborazione con il reparto IT e con eventuali *data scientists*. Inoltre, dal momento che spesso i big data contengono miriadi di informazioni utili ma confuse, è necessario rendere questi database utilizzabili, cioè abbastanza maneggevoli da cogliere le informazioni importanti in essi contenute.

Le tipologie di informazioni che si hanno a disposizione sono diverse: demografiche, comportamentali, riguardanti gli acquisti effettuati, per esempio.

Customer lifetime value

In sostanza, lo scopo del marketing consiste nel “creare relazioni di valore con il cliente” (Kotler, Armstrong, Ancarani, Costabile, 2015): per ottenere il massimo dalle relazioni con i clienti è necessario massimizzare il *customer lifetime value*, che consiste nell'ammontare totale di ricavi ottenibili dalla relazione con un cliente nel corso del suo rapporto con l'azienda. A tale scopo, è necessario valutare il valore degli acquisti futuri del cliente con tecniche previsionali, che si basano sugli acquisti passati, sugli acquisti presenti e futuri di clienti simili e sul tasso di ritenzione, cioè la percentuale di clienti che rimane fedele all'azienda nel tempo.

Lo scopo del marketing manager non è quello di massimizzare il tasso di ritenzione, bensì di massimizzare il valore totale del *lifetime value* dell'intero portafoglio clienti: è meglio perdere dei clienti non redditizi e concentrare la propria attenzione su un numero minore di persone che permettono all'azienda di massimizzare il suo profitto.

Tuttavia, è importante far sì che il tasso di ritenzione sia sufficientemente elevato, in quanto una solida e fedele *customer base* assicura profitti stabili nel tempo. A tal fine, nei confronti del singolo cliente è importante fare attenzione alle fasi di: 1) *acquisizione del cliente*, che può richiedere svariati contatti e necessita la focalizzazione sulla tipologia più adatta di clienti, con tecniche di *targeting* adeguate;

2) *crescita*: si desidera fare in modo che il cliente, dopo il primo acquisto, diventi un cliente abituale;

3) *ritenzione*: lo scopo in questa fase è riconoscere e apprezzare i clienti leali, trattenendo i clienti ad alto valore.

D'altra parte, se si considera l'intero insieme di clienti, è importante innanzitutto cercare di fidelizzare clienti profittevoli, in numero maggiore o uguale a quello di coloro che lasciano l'azienda; in secondo luogo è necessario prevenire l'abbandono da parte dei clienti: recuperare i clienti persi è più difficile che mantenerli nella *customer base*; in terzo luogo è importante catturare nuovamente la fedeltà dei clienti inattivi: è più economico rispetto all'acquisizione di nuovi clienti, non noti e sprovvisti di *brand awareness*.

Il marketing, dunque, risulta essere particolarmente influenzato dall'utilizzo dei big data, che permettono una rapida evoluzione, dirigendosi sempre più verso un'esperienza di acquisto diversificata per ognuno, efficacemente rivolta a soddisfare bisogni e desideri talvolta non ancora espressi né noti al cliente stesso.

Di seguito è presentato l'esempio dell'azienda che ha fatto delle raccomandazioni personalizzate il suo punto di forza, pioniera del marketing *one-to-one* e dell'utilizzo dei big data: Amazon.

2.6.1. ESEMPIO – AMAZON

Un'azienda che utilizza tecniche di *predictive marketing*, al fine di prevedere i comportamenti del cliente ed offrirgli il prodotto adatto a lui è sicuramente Amazon. Infatti, il leader nel settore *e-commerce* ha basato parte del suo successo sulle *recommendations* che appaiono quando si acquista o si guardano i dettagli di un prodotto: “ti potrebbe interessare anche...”, “chi ha acquistato questo articolo ha acquistato anche...”.

Agli albori dell'e-commerce, nei primi anni Duemila, Amazon, che all'epoca vendeva libri online, si basava sulle recensioni scritte da un suo gruppo di critici letterari e redattori per proporre nuovi titoli. Questo gruppo era considerato la fonte del vantaggio competitivo dell'azienda. Tuttavia, il CEO di Amazon Jeff Bezos, propose di elaborare raccomandazioni a ogni cliente in base alle proprie preferenze: Amazon aveva a disposizione un'ingente quantità di dati, riguardante gli acquisti e i comportamenti di ciascun utente che si interfacciava con il sito.

Inizialmente, il metodo per suggerire si basava sul convenzionale campionamento per mettere in evidenza le affinità tra i clienti: esso, però, risultava poco efficace, in quanto proponeva

prodotti solo leggermente diversi dal primo acquisto effettuato. Di conseguenza, Amazon procedette ad un cambio di prospettiva: non si ricercavano più affinità tra i clienti, bensì tra i prodotti, preferendo delle raccomandazioni del tipo “*item-to-item*”, in cui si consigliavano alcuni prodotti sulla base di quelli visualizzati o acquistati in precedenza, prevedendo così le preferenze del consumatore. La nuova metodologia era molto più efficace e veloce, in quanto si potevano svolgere i calcoli in anticipo e proporre i suggerimenti in modo subitaneo. Inoltre, le vendite generate applicando questo algoritmo erano ben superiori al metodo delle recensioni scritte dai critici. Emergevano dunque delle importanti correlazioni tra i prodotti, ma non si era in grado di spiegarne la causalità (Cukier, Mayer-Schönberger, 2013).

Queste tecniche risultano fruttuose: attualmente il 35% delle vendite sono dovute alle raccomandazioni, e si aggiravano attorno a 26 miliardi di dollari nel 2013 (Artun, Levin, 2015). Le raccomandazioni si basano su tecniche avanzate di *machine learning*, algoritmi che sono in grado di diventare più efficienti con il tempo, poiché possiedono più dati disponibili e possono dare un peso maggiore a quelli più recenti. Per essere adeguati, infatti, i suggerimenti dovrebbero essere aggiornati giornalmente, in modo da tener conto dei comportamenti più recenti dei clienti e interpretare in anticipo i nuovi trend, monitorando i cambiamenti di acquisto dei clienti più simili. Gli algoritmi, inoltre, dovrebbero evitare di suggerire prodotti obsoleti o in sconto e privilegiare quelli ad alto margine.

Si evidenzia dunque l'importanza dei suggerimenti d'acquisto: essi sono in grado di far lievitare i ricavi consigliando al cliente i prodotti giusti per lui e allo stesso tempo riescono ad aumentare *l'engagement* e il coinvolgimento nei confronti del brand, suggerendo contenuti che potrebbero essere interessanti. È altresì importante che le raccomandazioni non risultino troppo intrusive, insistenti o inadeguate, e che compaiano al momento giusto, per esempio quando si sta effettuando l'acquisto (Artun, Levin, 2015).

Le metodologie implementate inizialmente da Amazon si sono poi diffuse e sono oggi adottate da quasi tutti i siti di *e-commerce*, rivoluzionando il modo di fare previsioni.

Capitolo 3: I Big Data e le conseguenze per le statistiche ufficiali

3.1. BIG DATA E STATISTICHE UFFICIALI, NOVITÀ E TRADIZIONE: SCONTRO O FUSIONE?

I big data, come è già stato illustrato in precedenza, sono forieri di cambiamenti che coinvolgono gli aspetti più disparati. La statistica ufficiale è particolarmente coinvolta in tali mutamenti: infatti i big data mettono a disposizione dell'analista o del manager enormi quantità di dati che possono essere utilizzati per le statistiche ufficiali.

I big data si contraddistinguono per l'alto volume, l'alta velocità e l'elevata varietà. L'enorme volume dei dati garantisce una maggiore precisione e livello di dettaglio; la grande velocità nella loro generazione permette di ottenere statistiche più tempestive e più velocemente; la maggior varietà permette di ottenere stime per ambiti nuovi e diversi. (Braaksma, Zeeleberg, 2015).

D'altro canto, le statistiche ufficiali si basano su fonti più tradizionali, quali indagini campionarie, censimenti, dati provenienti da apparati amministrativi.

Le indagini campionarie hanno una solida base metodologica: si basano infatti sul pieno controllo nella fase di raccolta dei dati, che permette di ottenere database gestibili in maniera relativamente semplice e di buona qualità. I punti deboli sono collegati alla resistenza nei confronti della responsabilità delle risposte e gli elevati costi dovuti al crescere delle dimensioni.

I censimenti utilizzano dati costantemente aggiornati in possesso degli apparati amministrativi, permettono di ottenere buoni risultati anche per piccole aree e ridotti sottogruppi della popolazione, ma comportano alti costi per la realizzazione del censimento stesso e richiedono la collaborazione tra i detentori dei dati e centri di statistica per la condivisione dei dati: alcuni di essi sono posseduti da apparati amministrativi, ma in quantità crescente sono generati da enti privati quali aziende (per esempio i dati sulle vendite ottenuti dalla scansione dei codici a barre nei supermercati): a tal proposito è necessario accordarsi in modo proficuo per l'accesso ai dati da parte degli istituti ufficiali (Hackl, 2016).

Le statistiche ufficiali, in sostanza, si basano su un approccio di tipo tradizionale, sviluppatosi in situazioni di scarsità dei dati: ottenerli risulta costoso e difficile, ma ne garantisce il pieno controllo. Tali modalità di reperimento di dati fanno della buona qualità dei risultati il loro punto di forza, ma è necessario considerare la rapidità con cui la realtà circostante cambia:

raccogliere dati esatti è certamente un buon punto di partenza per svolgere analisi corrette, ma questo spesso comporta un insostenibile allungarsi dei tempi necessari per l'elaborazione.

Utilizzare dati ottenuti da altre fonti, come i big data, permette di avere accesso a grandi quantità di informazioni molto velocemente, ma riduce il controllo sui dati, quindi anche sulla loro qualità. Spesso, infatti, si tratta di dati raccolti da enti che non hanno come principale scopo l'elaborazione di statistiche.

Le fonti principali da cui provengono le informazioni che costituiscono i big data e che sono utilizzati nelle statistiche ufficiali sono soprattutto gli “*scanner data*”, cioè i dati generati dalla scansione del codice a barre nei Point-Of-Sale (POS) dei negozi, i dati provenienti dagli smartphones, dati sul traffico e altre fonti di dati, dati provenienti da internet e dai social media.

Per ciò che concerne dati provenienti dai codici a barre, essi permettono di ottenere statistiche sui prezzi medi e sugli acquisti effettuati, ma richiedono una buona collaborazione con le imprese che generano i dati stessi; inoltre, tale genere di dati non è sempre disponibile per molte categorie di prodotto.

I dati provenienti da cellulari e smartphones, come ad esempio telefonate, sms, posizione, sono processati in primo luogo dalle compagnie telefoniche (con le quali è importante collaborare) e permettono di elaborare statistiche sui comportamenti degli individui, sui flussi e gli spostamenti della popolazione, sul turismo, sulle migrazioni.

I dati riguardanti il traffico stradale, invece, provengono principalmente da sensori e webcam posizionati sulle strade e sui veicoli, e permettono di ottenere informazioni interessanti sui flussi della popolazione, sulle quantità di merci importate ed esportate e dunque sui flussi economici tra zone diverse.

I dati generati dai social media sono difficili da raccogliere e organizzare (quasi il 50% dei contenuti dei messaggi scambiati online costituiscono un “chiacchiericcio privo di significato”, secondo uno studio effettuato sull'utilizzo di Twitter in Olanda, da Braaksma e Zeeleberg, 2015), ma costituiscono delle buone basi per individuare trend su opinioni e umore della popolazione. La velocità con cui questi dati si generano permette inoltre di conoscere la situazione in modo tempestivo e rapido.

Altre fonti di dati riguardano immagini provenienti da satelliti o *remote sensing data*, cioè dati ricavati da sensori e che si ottengono a distanza mediante radiazioni elettromagnetiche.

3.2. BIG DATA NELLE STATISTICHE UFFICIALI: NUOVE SFIDE E OPPORTUNITÀ

I big data sono fonte di enormi potenzialità per le statistiche ufficiali: è possibile infatti per gli istituti ufficiali utilizzare queste enormi quantità di dati, ma anche ricombinarli con i dati raccolti in modo tradizionale.

I principali vantaggi risiedono nella tempestività delle informazioni, nel maggior approfondimento e precisione che la mole di dati consente di ottenere, nella riduzione dei costi per raccogliere i dati stessi, nella possibilità di effettuare analisi statistiche anche in nuovi ambiti e acquisendo nuovi e diversi indicatori.

Tuttavia, si pongono anche delle sfide nel governo delle nuove modalità di analisi. I big data non sono infatti semplici da gestire: non sono ben strutturati, mancano di un campione di riferimento e spesso la qualità non è garantita, ma si tratta di “un indicatore tempestivo, anche se imperfetto dello sviluppo della popolazione” (Braaksma, Zeelemborg, 2015), che non si può non sfruttare.

Gli istituti ufficiali di statistica devono affrontare varie sfide nell’impiegare anche i big data per le loro analisi: innanzitutto è fondamentale che il personale sia formato e preparato nella nuova modalità di effettuare le statistiche e che la gestione dei dati avvenga in maniera corretta e nel rispetto delle norme, ma ci si pone di fronte anche a sfide diverse.

In particolare, si devono affrontare eventuali limitazioni legislative per l’accesso ai dati: in alcuni casi non è possibile avere il pieno accesso a tutti i dati ed emerge la necessità di integrare le informazioni mancanti per poter effettuare analisi complete.

Altra tematica importante è quella della privacy e della gestione dei dati personali: spesso gli utilizzatori di dispositivi elettronici o di internet non hanno la piena consapevolezza di quanto comunicano e di come i loro dati possono essere utilizzati e ricombinati, specialmente per utilizzi secondari. È importante pertanto informare i cittadini dei possibili impieghi delle informazioni che li riguardano ed educare soprattutto i più giovani ad un uso consapevole delle tecnologie.

Un altro aspetto da considerare è quello finanziario: acquisire i dati può risultare oneroso, specialmente se a fornirli sono enti privati. È perciò importante confrontare qualità, tempestività, precisione, accessibilità e interpretabilità dei dati con i costi necessari per acquisirli e la riduzione della responsabilità delle risposte.

Si pongono anche questioni di tipo metodologico: tradizionalmente gli istituti che producono le statistiche ufficiali si sono sempre basati su una popolazione target per effettuare le proprie analisi, selezionando gli aspetti del campione da indagare, costruendo un modello e ottenendo dati in maniera strutturata. Ora accade il contrario: a partire dai dati, si deve costruire un algoritmo in grado di “districare la matassa”, per ottenere informazioni rilevanti e comprendere i trend; è necessario pertanto un cambio di approccio nel costruire le statistiche ufficiali. Occorrono dei metodi per svelare rapidamente le informazioni utili dall’ammasso informe di dati e per integrarle nel processo statistico.

Inoltre, emerge la necessità di sviluppare metodi che consentano di valutare la qualità dell’output che deriva da dati esterni e difficili da gestire e di superare alcune problematiche relative all’integrazione di dati provenienti da fonti esterne, dalle quali è più difficoltoso estrapolare dati ad alto valore.

Si pone infine una sfida anche da punto di vista tecnologico: per processare i dati più velocemente e per integrare i dati di tipo amministrativo con quelli provenienti da altre fonti (commerciali, da dispositivi di *tracking* o da sensori fisici, da social media e motori di ricerca), migliorando la tempestività dell’analisi, è importante utilizzare dei sistemi informatici innovativi. In effetti, una questione importante e che offre nuove possibilità di sviluppo è quella della raccolta e integrazione di diversi tipi di dati provenienti da varie fonti, anche combinando dati esterni con quelli rilevati tramite le metodologie tradizionali (United Nations Economic Commission For Europe - Conference Of European Statisticians, 2013).

Un altro problema notevole è quello della rappresentatività della popolazione fornita dai big data: si può notare se ci sono disparità tra la popolazione target e quella ripresa dal campione; operazione, questa, spesso difficile in quanto mancano le informazioni disponibili per effettuare il confronto. Tuttavia, basandosi su previsioni e algoritmi si possono ottenere dei termini di paragone efficaci.

Nell’utilizzare i big data per le statistiche ufficiali, si incontrano varie problematiche anche di tipo pratico. Innanzitutto è indispensabile una fase di indagine sui dati a disposizione, in quanto essi spesso provengono da fonti esterne agli istituti nazionali, frequentemente da enti privati. Questa fase consente di comprendere il contenuto e la struttura dei dati, per poterli utilizzare in modo efficace.

In secondo luogo, ci si interfaccia con il problema dei dati mancanti: i database forniti possono essere lacunosi e gli istituti devono sviluppare modelli che fronteggino l'incompletezza e metodi che permettano l'aggregazione di dati e l'estrazione di informazioni. Inoltre, i dati provenienti da sensori o da *sentiment analysis* risultano spesso caratterizzati da un'alta volatilità e da fluttuazioni che non danno informazioni importanti sul fenomeno in quanto si presentano ad una scomposizione troppo elevata: è pertanto fondamentale riuscire a fronteggiare la volatilità.

In sostanza, con lo sviluppo delle tecnologie legate ai big data il ruolo degli istituti di statistica ufficiale non passa in secondo piano, ma mantiene l'importanza che ha finora rivestito nell'elaborare statistiche e stime valide e attendibili: gli istituti nazionali di statistica, infatti, possono affidarsi a delle ottime strutture per garantire l'affidabilità, l'accuratezza e l'interpretabilità dei risultati ottenuti. Anzi, i big data permettono di elaborare stime e valutazioni su ambiti che prima non erano accessibili, spesso per mancanza di dati, e di aumentare la velocità e la precisione con cui le statistiche sono prodotte, rendendole in tal modo più utili. Inoltre, i big data riducono la responsabilità sull'esattezza dei dati e i costi della produzione delle statistiche.

È fondamentale pertanto integrare i dati ricavati con le tradizionali metodologie con quelli provenienti dai big data, per esigenze di completezza e per supplire alle attuali mancanze tecnologiche.

Non vi è un unico metodo per utilizzare i dati nella produzione di statistiche, ma ogni fonte di dati necessita di essere gestita ad hoc: per questo motivo è importante sviluppare delle soluzioni per i diversi tipi di problemi, aumentando anche le competenze delle risorse umane impiegate. Inoltre, è necessario definire degli standard per le qualità dei set di dati e dei prodotti statistici che ne derivano (Hackl, 2016).

Capitolo 4. Big data e Internet of Things (IOT)

4.1. INTERNET OF THINGS: L'ORIGINE DEI DATI

Un campo in espansione, correlato a quello dei Big Data, è costituito dall'Internet of Things: si tratta di un fenomeno in piena crescita e che presenta ancora molte questioni aperte dal punto di vista degli impieghi presenti e futuri e dei possibili sviluppi tecnologici.

L'IoT (Internet of Things) costituisce una sorta di evoluzione e di estensione dell'Internet tradizionale: se Internet si focalizza sul collegamento tra i computer, l'Internet of Things pone l'accento sulle cose, anziché sui computer, e mira a collegare tra loro oggetti di uso quotidiano connettendoli ad Internet al fine di permettere la comunicazione tra loro e con l'ambiente circostante, riuscendo a interpretare i segnali e i cambiamenti che dall'ambiente provengono (Qin, Sheng, Falkner, Dutstar, Wang, Vasilakos, 2016).

Il concetto base del “nuovo paradigma che sta guadagnando visibilità sullo scenario delle comunicazioni wireless”, è appunto la “presenza pervasiva attorno a noi di una varietà di cose o oggetti che, attraverso uno schema unitario, sono in grado di interagire tra loro e cooperare con gli oggetti vicini per un comune scopo” (Atzori, Iera, Morabito, 2010).

L'impatto dell'Internet of Things si riflette in vari ambiti, dalla salute alla domotica, dal business al management delle *smart cities*: attraverso la connessione ad Internet di oggetti di uso quotidiano si avrà accesso a molti vantaggi ottenuti grazie al flusso costante di dati e informazioni che provengono dai vari dispositivi.

Il cambio di prospettiva che richiede l'IoT riguarda principalmente l'andare oltre lo scenario tradizionale del *computing* connesso a portatili e smartphones, passando ad un mondo in cui si connettono oggetti di uso quotidiano ad Internet ottenendo molte più informazioni sull'ambiente (Gubbi, Buyya, Marusic, Palaniswami, 2013).

L'IoT, dunque, costituisce una sorta di superamento dell'attuale Internet, per giungere ad una rete di oggetti interconnessi, che non solo raccolgono informazioni e interagiscono con l'ambiente circostante, ma utilizzano gli standard di Internet per trasferire informazioni, analizzare i dati e comunicarli (Gubbi et al., 2013).

Il fenomeno IoT si basa su dispositivi connessi con tecnologia wireless, tramite Bluetooth, WiFi, RFID tags (etichette con Radio Frequency IDentification, cioè microchip che sono installati sulla superficie di vari oggetti e che generano soprattutto informazioni riguardanti la tracciabilità e la visibilità degli stessi), dati telefonici: queste tecnologie permettono di identificare la posizione e di ottenere altre informazioni relative all'oggetto e all'ambiente in

cui esso si trova. Infatti, i dispositivi riescono (e in futuro saranno sempre più precisi e accurati in queste funzioni) a raccogliere informazioni dall'ambiente circostante, a interagire con esso e a svolgere funzioni di trasferimento di informazioni, analisi, comunicazione (Gubbi et al., 2013).

4.2. APPLICAZIONI

L'IoT è utilizzabile in diversi campi: applicazioni in ambito personale e della salute; impieghi nel campo del business e dell'impresa; utilizzi nelle *smart cities* e in campo ambientale.

Il campo della salute e della cura della persona possono avere molti effetti positivi utilizzando i dati provenienti dai sensori e da *RFID tags* che si possono disporre su medicinali, su macchinari utilizzati per effettuare analisi e sugli strumenti medici. Sarà possibile ottenere esiti degli esami più precisi e affidabili, monitorare in tempo reale lo stato di salute delle persone, programmare controlli medici in modo più efficiente.

Dal punto di vista delle imprese, la tecnologia IoT permetterà di migliorare l'efficienza: si ridurranno i costi per il processo produttivo, per la logistica e per la gestione del magazzino, grazie al maggior numero di informazioni raccolte. Inoltre, i consumatori potranno avere più informazioni riguardo ai prodotti e alla loro provenienza, si potranno prevenire guasti monitorando lo "stato di salute" dei prodotti, la maggior tracciabilità garantirà più trasparenza e assistenza.

Un altro importante aspetto riguarda l'impatto dell'IoT nelle *smart cities* e nella domotica: rilevando informazioni nelle città, queste possono essere più efficienti, migliorando i trasporti e la sicurezza, riducendo i costi e permettendo di risparmiare tempo ed energia, dai consumi per riscaldare le abitazioni ai carburanti per gli spostamenti urbani.

Anche la protezione dell'ambiente può trarre vari vantaggi dalle tecnologie IoT: infatti è possibile, grazie alle maggiori informazioni disponibili, monitorare i livelli di inquinamento presenti nell'ambiente e intervenire tempestivamente ove necessario, risparmiando sui costi di manutenzione e regolamento dei sensori (Qin et al., 2016).

4.3. CONNESSIONE TRA BIG DATA E IOT

Big Data e IoT sono due fenomeni fortemente intrecciati: l'origine dei Big Data è dovuta allo sviluppo tecnologico che ha consentito la generazione di enormi quantità di dati e il loro

immagazzinamento, mentre l'IoT è proprio uno degli elementi trascinanti nello sviluppo e nell'analisi dei Big Data.

Nel 2012, “2,5 quintilioni ($2,5 \times 10^{18}$) di bytes di dati erano generati ogni giorno”, mentre con l'utilizzo sistematico dell'Internet of Things sarà possibile connettere molti più dispositivi, e si genereranno quantitativi di dati sempre più ingenti: si pongono pertanto problemi di gestione e immagazzinamento dei dati per poi effettuare le analisi che permetteranno di individuare trend e isolare informazioni che daranno risposte più precise e adatte ai singoli fenomeni.

Per velocizzare e migliorare la fase di *analytics*, che avviene nella maggior parte dei casi solo in seguito alla raccolta e al trasferimento dei dati, ma anche per gestire in modo più efficiente i dati, occorre sviluppare dispositivi tecnologici che riducano i problemi di incertezza, ambiguità e inconsistenza dei dati e che li trasmettano ripuliti da errori e compressi, in modo da ridurre lo spazio necessario (Qin et al., 2016).

In sostanza, si può affermare che i Big Data e l'*analytics* sono due aspetti dello stesso fenomeno: lo sviluppo tecnologico che ha permesso la nascita dell'IoT e che in futuro è destinato a crescere in modo sempre più rapido ha generato le enormi moli di dati che stanno rivoluzionando il modo di condurre le analisi e che sono alla base del fenomeno dei Big Data. Per questa ragione, tali aspetti vanno considerati congiuntamente; le questioni che essi pongono, come quelle relative alla privacy o all'immagazzinamento dei dati, vanno colte da entrambe le prospettive: da una parte, la generazione dei dati, dall'altra l'analisi degli stessi, che permette di ricavare, dall'ammasso confuso di informazioni, tendenze precise, un'interpretazione specifica della realtà che ci circonda.

Capitolo 5. Conclusioni

La trattazione fin qui condotta ha lo scopo di evidenziare il notevole impatto dei Big Data in vari campi, soffermandosi in maniera particolare sulle conseguenze di questa innovazione in campo aziendale.

Il fenomeno dei Big Data è in piena espansione, ma è difficile prevedere a che punto arriveranno gli sviluppi futuri della tecnologia. Si può però affermare con sicurezza che i miglioramenti tecnologici condurranno ad una maggior efficienza nel trattare i dati e ad una più efficace e precisa analisi degli stessi.

Il mutamento principale apportato dai Big Data, tuttavia, è già avvenuto, e risiede nel cambio di prospettiva evidenziatosi nella metodologia di analisi. Come si è sottolineato, infatti, le tecniche utilizzate nell'era degli “*small data*” sono difficilmente applicabili ai dati a nostra disposizione oggi, data la loro estensione molto maggiore: la regressione lineare non è più la tecnica migliore per decifrare le informazioni nascoste nei dati, in quanto si richiedono *dataset* relativamente ridotti, omogenei e organizzati.

Ora, nella maggior parte delle situazioni, una riorganizzazione dei dati antecedente all'analisi risulta eccessivamente dispendiosa in termini di tempo, se non addirittura impossibile, data la quantità di dati presente. Ciò su cui si pone l'accento nella fase dell'analisi dei dati è dunque la correlazione tra le variabili: si osserva come esse variano congiuntamente, senza necessariamente conoscerne la ragione. Risulta perciò essere più importante questo aspetto rispetto ad un'analisi più approfondita qual è la relazione di causalità.

Il capovolgimento di prospettiva richiede un adattamento da parte degli analisti e di chi utilizza i dati: essi non privilegeranno più l'assoluta precisione e l'esattezza, piuttosto riterranno più importante riuscire a decifrare tra ammassi di dati non ordinati delle tendenze, che spiegano in maniera sufficientemente accurata ma non priva di un'area di imprecisione, l'andamento e la possibile evoluzione dei fenomeni che si prendono in esame.

È pertanto il *modus operandi* ad essere capovolto: anziché applicare un modello prestabilito ai dati, è ora necessario che siano i dati stessi a indicare i trend, che possono emergere grazie ad appropriate tecniche statistico-matematiche, a strumenti informatici adeguati e, soprattutto, all'interpretazione dei dati ottenuti.¹

¹ 14567 parole

Bibliografia

- ARTUN, O., LEVIN, D., 2015. *Predictive Marketing. Easy ways every Marketer can use customer analytics and Big Data*. Wiley
- ATZORI, L., IERA, A., MORABITO, G., 2010. The Internet of Things: a survey, *Computer Networks*, 54, PP. 2787-2805
- BRAAKSMA, B., ZEELEMBERG, K., 2015. Re-make/Re-model: should big data change the modelling paradigm in official statistics?. *Statistical Journal of the IAOS*, pp. 193-202
- CHEN, Philip, ZHANG, Chun-Yang, 2014. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275 (2014), pp. 314-347
- CUKIER, K. 2010. Data, data everywhere. *The Economist*. Disponibile su <http://www.economist.com/node/15557443/print>, [data di accesso: 29/05/2016]
- GUBBI, J., BUYYA, S., MARUSIC, M., PALANISWAMI, M., 2013. Internet of Things (IoT): a vision, architectural elements, and future directions, *Future Generation Computer Systems*, 29, pp. 1645-1660
- HACKL, P. 2016. Big Data: what can official statistics expect?, *Statistical Journal of the IAOS*, 32, pp. 43-52
- < https://www.google.com/intl/it_it/analytics/ > [data di accesso: 21/09/2016]
- KOTLER, P., ARMSTRONG, G., ANCARANI, F., COSTABILE, M., 2015. *Principi di Marketing. Quindicesima edizione, Milano: Pearson*.
- MANYKA, J., CHUI, M., BROWN, B., BUGHIN, J., DOBBS, R., ROXBURGH, C., HUNG BYERS, A. 2011. Big Data: the next frontier for innovation, competition and productivity, Mc Kinsey Global Institute
- MARSHALL, Anthony, MUECK, Stefan, SHOCKLEY, Rebecca, 2015. How leading organizations use big data and analytics to innovate. *Strategy & Leadership*, 43 (5), pp. 32-39
- MAYER-SCHÖNBERGER, V., CUKIER, K., 2013. *Big Data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*. Milano: Garzanti.
- MC AFEE, Andrew, BRYNJOLFSSON, Erik, 2012. Big Data: the Management Revolution. *Harvard Business Review*,
- PEMBE MUHTAROGLU, C., DEMIR, S. , OBALI, M., GIRGIN, C., 2013. *Business model canvas perspective on big data applications*, 2013 IEEE International Conference on Big Data

- PLAZA, Beatriz, 2010. Google Analytics. Intelligence for Information professionals, www.onlinemag.net [data di accesso: 21/09/2016]
- QIN, Y., SHENG, Q. Z., FALKNER, N., DUTSTAR, S., WANG, H., VASILAKOS, A., 2016. When things matter: a survey on data-centric Internet of Things, *Journal of Network and Computer Applications*, 64, PP. 137-153.
- SANDERS, Nada., 2016. How to use big data to drive your supply chain. *California Management Review*, 58 (3), pp. 26-48
- SLACK, N., BRANDON-JONES, A., JOHNSTON, R., 2013. *Operations Management. Seventh edition, Pearson.*
- United Nations Economic Commission For Europe- Conference Of European Statisticians, High-Level Group for the Modernisation of Statistical Production and Services (HLG), 2013. *What Does Big Data Mean For Official Statistics?*
- WINSNIEWSKI, J., 2012. Advanced Google Analytics Tips and Tricks, www.onlinemag.net [data di accesso: 21/09/2016]
- WINSNIEWSKI, J., 2012. Getting started with Google Analytics, www.onlinemag.net [data di accesso: 21/09/2016]