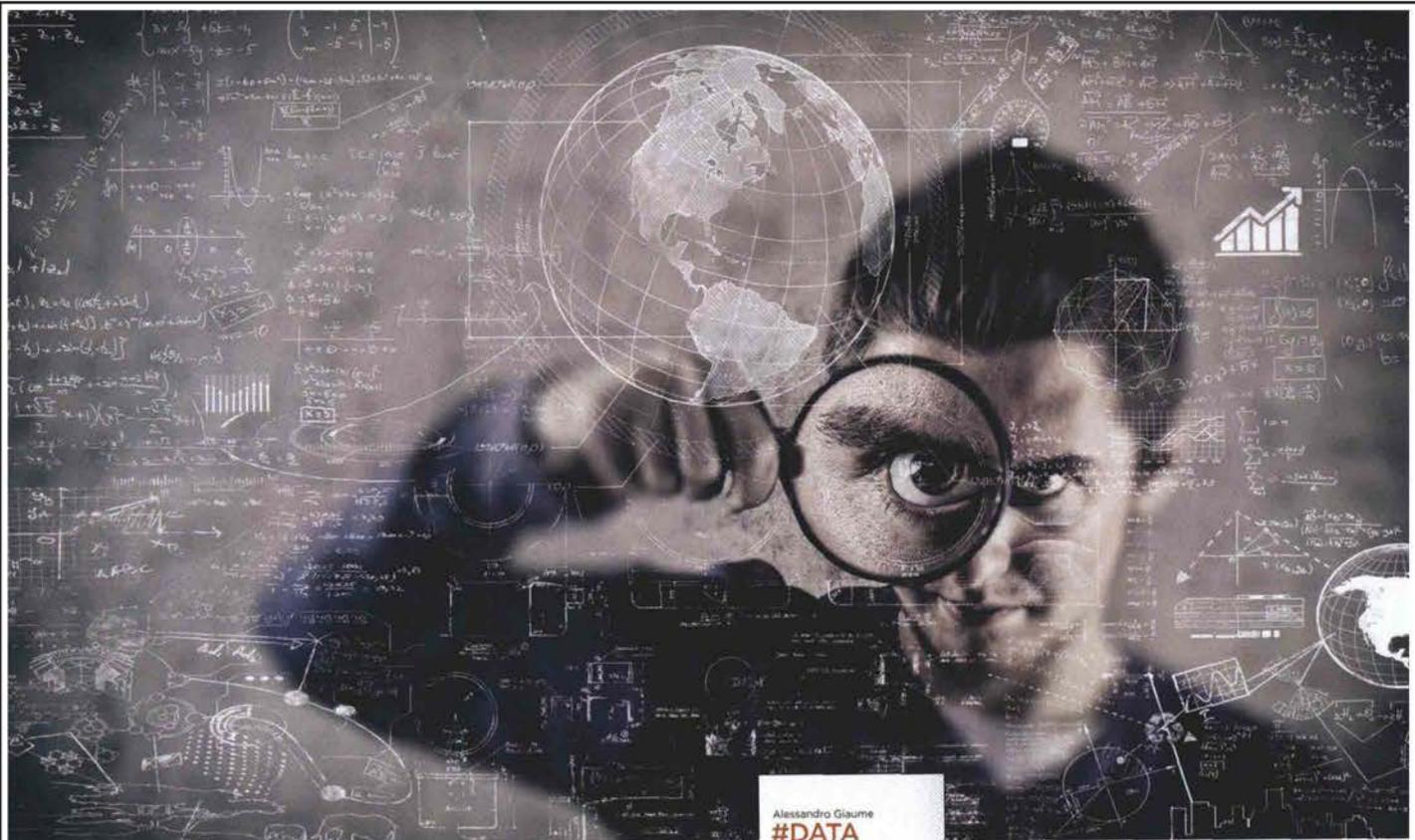


m WEB
INNOVATORI
di Tiziana Tripepi, t.tripepi@millionaire.it

I NOSTRI COMPOR-
TAMENTI GENERANO
UN'ENORME QUAN-
TITÀ DI DATI. LA VERA
OPPORTUNITÀ? CONO-
SCERLI, ANALIZZARLI
ED ELABORARLI. PER
FARNE UN LAVORO O
UN NUOVO BUSINESS.
TUTTO SUI BIG DATA

Big Data sono una parte importante della nostra vita e lo saranno sempre di più nei prossimi anni. Siamo noi stessi a generarli: quando corriamo con il nostro orologio con-
tapassi, quando andiamo in giro con il nostro smartphone e siamo geolocalizzati, quando ci colleghiamo a un sito per vedere cosa propone, concludiamo un acquisto su Amazon o parliamo sui social dei nostri gusti e le nostre preferenze. Sono caratterizzati da 5 V: Volume, Velocità, Varietà, Veracità, Valore. In un minuto vengono scritti più di tre milioni di post su Facebook, condivise 540mila foto su Snapchat, caricati 120 video. Un'enorme massa di dati generata dalle nostre abitudini che è di grande interesse per le aziende per migliorare le loro performance o prendere decisioni strategiche.

BIG DATA BIG BUSINESS



DATI STRUTTURATI+ DATI NON STRUTTURATI= BIG DATA



Alessandro Giaume

«Le aziende hanno da sempre utilizzato i dati per prendere decisioni. Hanno sistemi gestionali per registrare le fatture, anagrafiche per clienti e fornitori, report di vendita. Sono i cosiddetti dati strutturati, quelli cioè generati all'interno dell'azienda, che costituiscono la *business intelligence* di un'impresa. Ma negli ultimi anni una nuova categoria di dati si è fatta strada e sta crescendo esponenzialmente: sono quelli prodotti da entità esterne all'azienda, che arrivano dai social network, dai sensori, le immagini da satellite, i segnali Gps» spiega Alessandro Giaume, Innovation Director di Ars et Inventio, del gruppo Bip (www.arsetinventio.com) e autore del libro *#Data Scientist* (FrancoAngeli, 20 euro. 5 copie in omaggio per

Alessandro Giaume
#DATA SCIENTIST
Tra competitività e innovazione



i lettori. Scrivete motivazione e indirizzo postale a: esperto@millionaire.it entro il 31 maggio).

Un'enorme massa di dati non strutturati, cioè non organizzati in maniera predefinita, è disponibile. E la cosa incredibile è che costituiscono l'80% dei dati che interessano un'azienda.

Anche se le imprese non li sfruttano ancora come dovrebbero». Il totale di tutti questi dati costituisce i Big Data.



3,1 MILIONI
di like su Facebook



3,3 MILIONI
di post su Facebook



42MILA
upload di Instagram



204 MILIONI
di email spedite



120
video caricati

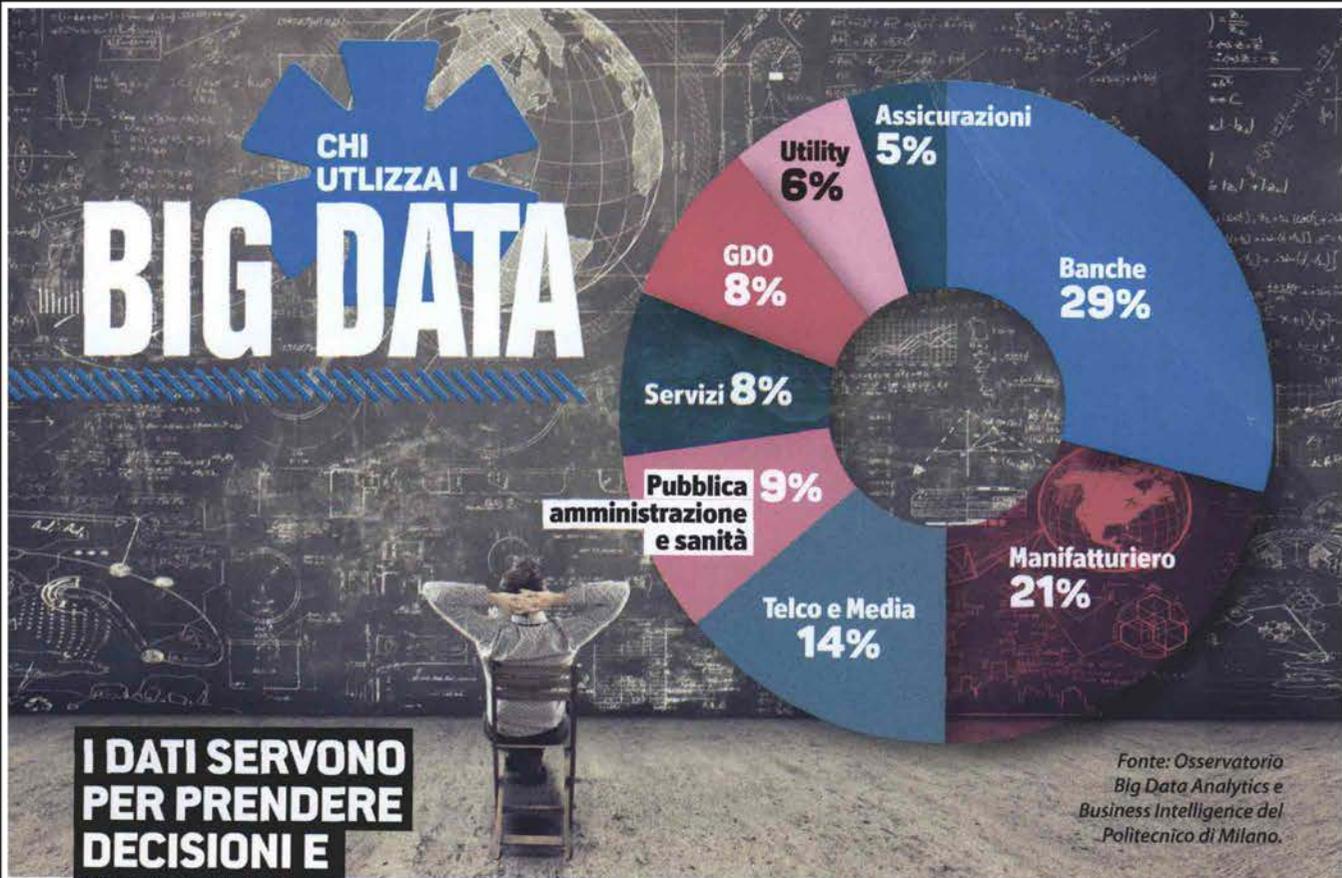


3MILA
articoli acquistati su Amazon

3 MILIONI
di ricerche su Google



Fonte: Osservatorio Big Data Analytics e Business Intelligence del Politecnico di Milano.



«Grazie ai Big Data si possono analizzare i comportamenti dei potenziali clienti, prevederli attraverso modelli previsionali, fino a influenzarli» continua Giaume. «Pensiamo a come si comporta Amazon: esamina il comportamento che teniamo quando ci colleghiamo al sito, costruisce dei profili, ci propone degli acquisti. Chi compra la racchetta da tennis comprerà anche le scarpe o un libro su Roger Federer». Per rimanere a casa nostra, i fondatori di LOVTheSIGN, e-commerce di oggetti di design, hanno deciso un paio di anni fa di lanciare anche prodotti propri sul mercato, perché analizzando i dati degli utenti si sono resi conto che c'erano buchi di mercato: «La clientela non cercava un divano dalla forma classica, per intenderci quello con i braccioli, ma un prodotto flessibile, modulare, e con la possibilità di essere configurato a proprio piacimento di volta in volta, per poter avere continuamente un divano "nuovo". È così che è abbiamo creato Rodolfo. Stesso lavoro per la sedia modulare Sofia, lanciata poche settimane fa».

DEVONO ESSERE SELEZIONATI, ELABORATI, CLASSIFICATI

I Big Data si caratterizzano per la grande varietà e per la velocità con cui si presentano e cambiano. Come fare a organizzarli per poterli elaborare? «Il primo passo è eliminare quelli inutili» continua Giaume. «Basti pensare che di questo 80% di dati non strutturati, solo il 5% ci serve davvero. Per fortuna la tecnologia ha compiuto passi da gigante: abbiamo a disposizione strumenti che permettono di elaborare grandissime quantità di dati e identificare quelli giusti. Per quanto riguarda questa fase, la parola magica è Hadoop, un software open source per l'archiviazione e l'analisi di quantità elevatissime di dati». Il secondo passo è la pulizia (in gergo *munging*) e l'organizzazione. I dati vengono "versati" in quello che in gergo si chiama "lago" (*data lake*), per poi essere selezionati ed elaborati. Per farlo, oltre ai software occorrono le competenze e la capacità di costruire algoritmi complessi. Una nuova professione è nata negli ultimi anni.

Grazie all'attività di Big Data, il sito LOVTheSIGN ha creato Rodolfo, un divano "nuovo".



DATA SCIENTIST: UNA PROFESSIONE "SEXY"

In America lo chiamano "il lavoro più sexy del XXI secolo". Quella del *data scientist* è una figura molto richiesta. Lo stipendio medio negli Usa è di 100mila dollari l'anno, ma può arrivare a 200mila-250mila. «In Italia non è ancora così diffusa, ma anche qui è gettonatissima» ci spiega Alessandro Piva, responsabile della ricerca dell'Osservatorio Big Data Analytics e Business Intelligence del Politecnico di Milano (www.osservatori.net). «La richiesta di queste figure proviene in primo luogo da grandi aziende, soprattutto nel settore bancario, assicurativo e utility. In secondo luogo, dalle tante aziende di consulenza, che offrono servizi di elaborazione dati a grandi e piccole imprese, che magari ancora non hanno integrato al loro interno queste professionalità. Il profilo? «Trent'anni, laurea in Ingegneria informatica, Matematica o Statistica. Competenze multidisciplinari: occorre conoscere i linguaggi di programmazione, saper utilizzare le tecnologie per estrapolare e aggregare i dati, saper sviluppare modelli e algoritmi. Ma anche conoscere il business dell'impresa per la quale si lavora (vedi box). Negli ultimi anni sono nati nelle principali università italiane molti corsi di laurea e master post laurea che preparano a questo lavoro. Dal Politecnico di Milano (www.mip.polimi.it) all'Università di Bologna (www.bbs.unibo.it). Per chi vuole iniziare c'è anche un percorso formativo online di Digital Learning (www.analyticsdigitallearning.com).



1. Matematico-Statistico

Le competenze devono essere di matematica e statistica. Ha una laurea in: Ingegneria, Matematica, Statistica, Fisica. Obiettivo: applicare il metodo scientifico e induttivo.

2. Programmatore

Deve conoscere linguaggi di programmazione: Hadoop, Python, R sono i più utilizzati.

3. Hacker

Non basta conoscere gli strumenti di programmazione, occorre anche essere un programmatore molto spinto, avere una capacità di elaborazione che travalica quella del "semplice" programmatore. Pensare oltre, avere un'attitudine "disruptive". Questa caratteristica è fondamentale per chi lavora per aziende come Airbnb e Uber.

4. Business person

Avere una visione del business e spirito imprenditoriale serve ad applicare il giusto modello matematico al modello di business dell'organizzazione.



I DATI SERVONO A TUTTE LE AZIENDE

«A seconda del business e delle competenze che hanno al loro interno, le aziende sfruttano i dati in gradi diversi» interviene Giaume.

1) Piccole e medie imprese. Aziende cosiddette di primo livello, in genere sotto i 200 milioni di fatturato, che hanno necessità di utilizzo basilari, e hanno bisogno di fare data analysis: estrazione dati, capacità di visualizzarli con i grafici.

2) Aziende di secondo livello, in genere grandi aziende, informatizzate negli anni 60, che hanno una grande mole di dati da analizzare e già utilizzano i Big Data. In Italia per esempio: Eni ed Edison. Il data scientist di queste aziende deve saper sviluppare software molto precisi ed essere autonomo.

3) Aziende data focused, come Google e Facebook, che sono cioè focalizzate sul business del dato in quanto tale. Qui lavorano data scientist "cresciuti", che si avvicinano agli hacker creativi, che devono saper costruire algoritmi futuribili.

4) Aziende data driven, come Amazon, Uber, Airbnb. Qui la conoscenza è ancora più spinta, i data scientist sono ai massimi livelli». In Italia, le nostre aziende si collocano a cavallo tra il secondo e il terzo gruppo.

BIG DATA

LE

5V

Volume

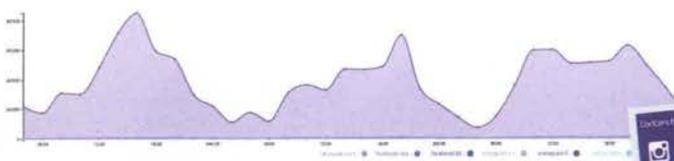
Velocità

Varietà

Veracità

Valore

BIG DATA



LA STORIA

«Abbiamo costruito il “termometro” dei social»



Si muovono con agilità tra i dati. Sono 9 ingegneri e 2 professori del Politecnico di Milano, che nel 2014 hanno costituito Fluxedo, una startup innovativa incubata nel Parco Scientifico e Tecnologico ComoNEXT, a Lomazzo (Co). Insieme a loro, Massimiliano Bancora, un business angel con conoscenze del mercato digitale, che ha deciso di investire tempo e competenze in questa impresa.

Che cosa fate? «Abbiamo costruito un prodotto, che si chiama “SocialOmeter”, in grado di aggregare in tempo reale i dati provenienti dai social network più popolari e rielaborarli sotto forma di grafici. Dato un determinato soggetto, il nome dell’azienda o del brand, analizziamo come viene discusso, come le persone interagiscono, come viene interpretato. Vi si può accedere dal sito, pagando una fee mensile. Oltre a questo, svolgiamo un’attività di *data science*, fornendo soluzioni su misura ad aziende o pubbliche amministrazioni. Per esempio un Comune ci può chiedere l’analisi dei flussi delle persone in una piazza della città perché vuole organizzare un evento, o un’azienda della grande distribuzione dall’analisi dei dati vuole capire quando i prodotti devono essere venduti o quando devono essere richiesti dai fornitori. Il costo per attività customizzate è da 3mila euro in su».

Che strumenti usate per elaborare i dati? «Ci sono software di pubblico dominio (open source), che possono essere utilizzati per elaborare i dati. Per il motore di ricerca sui social, entra in gioco l’analisi semantica, necessaria per trasformare delle conversazioni in dati: quando una persona “parla” sui social, devo tradurre le parole in un dato, e per farlo devo associarle ad altre parole

o argomenti, devo classificarle, associarle ad accezioni negative o positive. È la cosiddetta *sentiment analysis*. Infine, ci vogliono le competenze, le risorse umane sono la nostra più grande ricchezza».

Quali sono i vostri risultati economici? «Ci stiamo autofinanziando. Dopo aver costituito la Srl (con 10mila euro), abbiamo vinto un bando di 50mila euro all’interno di ComoNEXT, che ci ha dato la possibilità di affittare gli spazi e ricevere attività di *mentoring*. Recentemente abbiamo vinto un altro bando della Regione Lombardia che ci permetterà di intraprendere un progetto di analisi con i sensori. Abbiamo iniziato da subito a vendere i nostri servizi e stiamo crescendo. Il 2016 è stato chiuso con 100mila euro di fatturato, cifra già superata nei primi tre mesi di quest’anno».

Chi compra i vostri servizi? «Per il 50% sono istituzioni, che hanno bisogno di analizzare dati con la sensoristica, oppure vogliono sapere cosa succede sui social. Per il restante 50% si tratta di società di consulenza che ci “subappaltano” servizi richiesti da aziende e di cui loro non hanno le competenze. L’esigenza di elaborare questo genere di dati è un’esigenza sempre più grande da parte di ogni tipo di impresa».

INFO: www.socialometers.com

LE STARTUP DEI DATI

C’è chi con i dati decide di fare un business: questo è un campo di azione in cui a poter operare sono per loro natura le startup. «È uno degli ambiti in cui si registra più crescita, anche a livello internazionale» afferma Piva. «Sono 229 le startup a livello globale nel mercato dei Big Data e Business Intelligence, che hanno raccolto nel mondo 3,18 miliardi di dollari. In Italia se ne contano una trentina. Operano in tre campi principali. Il primo è quello delle applicazioni: inventano cioè strumenti che aggiungono funzionalità ad alcune funzioni che già esistono in azienda, per esempio riescono a tracciare meglio la *customer journey*, cioè cosa fa un cliente quando naviga sui social. La seconda è quella dei sistemi di Analytics, cioè la costruzione di piattaforme “orizzontali”, che riescono a fare analisi su più ambiti aziendali. La terza sono le cosiddette “tecnologie di abilitazione,” nuovi modi di analizzare i dati che vengono dai social o dall’Internet of Things o dai social. Il mondo sta andando in questa direzione: la capacità di accedere ai dati e di interpretarli è fondamentale in tutti i settori e chi non saprà adattarsi scomparirà».

