

# Il gold digger del terzo millennio

Nel mare magnum delle informazioni personali che quotidianamente affidiamo al web, attraverso messaggistica, social e ricerche, si nasconde un autentico tesoro per chi, come il Data Scientist, sa raccogliere, analizzare, interpretare questi dati, trasformandoli in un prodotto di conoscenza utile per le aziende

➡ Ginevra Bartoli

Se ne parla ormai da molti anni. Ai non addetti ai lavori sembrava un'aleatoria e dispersiva attività da geek generation. E, invece, oggi è di un'evidenza lampante che i nostri dati personali, i nostri commenti sui social, le nostre ricerche sui motori di ricerca si sono rivelati l'oro nero del terzo millennio. Sono i Big Data, l'enorme mole di informazioni a disposizione di chi – aziende pubbliche o private, mezzi di comunicazione, partiti politici e persino organizzazioni criminali – voglia servirsene per indirizzare messaggi pubblicitari o comprendere tendenze di

mercato e di costume. Mark Zuckerberg ha affermato qualche tempo fa: "dimenticatevi la privacy". E non gli si può dare torto, perché tutto quello che condividiamo nel web, sui social network o anche sui sistemi di messaggistica istantanea come Whatsapp lo affidiamo al mondo. Certo una quantità così vasta di informazioni e l'accumulo

costante di dati rende ardua la loro gestione e se non vengono trasformati in un prodotto di conoscenza utile, allora sì la loro esistenza è aleatoria e dispersiva. Ecco, dunque, che è nata, non da molto, una nuova figura professionale. È quella del Data Scientist, definita dal New York Times come la più sexy del terzo millennio. È l'analista di dati che cerca nel grande mare dei Big Data le informazioni che gli interessano, le raccoglie (ad esempio per conto di un'azienda che vuole approcciare un nuovo mercato) e le analizza secondo il suo scopo, anche attraverso software specifici.

## Il primo manuale del Data Scientist

Il primo dettagliato manuale del Data Scientist è stato recentemente pubblicato da Tecniche Nuove ed è opera di **Antonio Teti** (foto), esperto di cyber security e cyber intelligence, nonché docente





dell'Università Gabriele D'Annunzio di Chieti-Pescara. «Questo è un testo che vuole essere un manuale di riferimento per quella che viene indicata come la nuova professione del terzo millennio», ci ha spiegato Teti. «L'analisi dei dati è fondamentale, soprattutto in un contesto come quello attuale nel quale il cyberspazio è l'ecosistema del terzo millennio. E di questo si occupa Data Scientist, attraverso una serie di competenze a largo spettro, che vanno dalla statistica alla comunicazione, dalle conoscenze tecniche per reperire informazioni dal cyberspazio a quelle informatiche, fino alla capacità di trasformarle in informazioni utili. Si parla ormai di intelligence economica, sociale, politica». E di questa intelligenza nessuno può e deve più fare a meno. È proprio sui dati che si gioca il futuro delle aziende. «Che sia piccola, media o grande tutte le aziende hanno lo stesso problema: come possono essere trasformati i Big Data in

una conoscenza utile per i vertici dell'azienda al fine di prendere le decisioni migliori per ottimizzare i processi? Anche nel settore del turismo, naturalmente, grazie a questa analisi si possono realizzare attività molto interessanti: percepire quali possono essere gli interessi del viaggiatore significa poter creare una offerta verticalizzata e personalizzata; capire una particolare clientela, magari di una determinata provenienza, vuol dire poter creare percorsi turistici e pacchetti con grande valore aggiunto in termini di ospitalità o culturale».

### Un analista di dati deve capire cos'è un algoritmo, non lo deve sviluppare

Anche le piccole e medie imprese, dunque, sono in qualche modo ormai costrette ad adoperarsi per raccogliere questi dati. Ma oggi l'economicità di questa attività è un vantaggio. «Prima le informazio-





### La questione della privacy

Tutti noi contribuiamo alla creazione dei Big Data. Siamo produttori e fruitori di informazioni, molto spesso dati personali e riservati. Se decidiamo di immettere sui social informazioni personali, ad esempio attivando la geolocalizzazione, dobbiamo essere coscienti che le stiamo condividendo con il mondo. Più informazioni mettiamo nella rete più cediamo un pezzo della nostra integrità. Le prime aziende ad aver compreso l'importanza dell'acquisizione delle informazioni del cyberspazio sono riconducibili a importanti testate giornalistiche (per l'analisi di tendenze politiche o del sentiment e della fiducia nei confronti di determinate notizie). Ma, ahimè, anche la criminalità organizzata (per non parlare del terrorismo) sta sfruttando le opportunità dei Big Data. A questo si aggiunge ancora la violenza intrusiva dei messaggi spamming.

Ciò che conta comprendere è che il cyber spazio rappresenta uno strumento eccezionale, che può essere usato a fini diversi e per una molteplicità di scopi.

**Le prime aziende ad aver compreso l'importanza dell'acquisizione delle informazioni del cyberspazio sono riconducibili a importanti testate giornalistiche (per l'analisi di tendenze politiche o del sentiment e della fiducia nei confronti di determinate notizie)**

ni provenivano da fonti diverse, principalmente i cosiddetti *Old Media* e difficilmente producevano dati omogenei. Se un'azienda voleva cominciare a creare una sorta di contenitore di saggezza era costretta ad affidarsi a fonti diverse (società di marketing) con dispendio di energie, risorse umane e risorse economiche. Oggi attraverso il cyberspazio, basta avere un minimo di cultura in materia di ela-





**Una delle fonti di informazioni utili più importanti è Facebook: dai post è possibile raccogliere tutte le persone che inseriscono i like su un prodotto specifico, una destinazione o un'attività, e creare un data base da vendere ad una azienda per le sue azioni di marketing**

borazione delle informazioni per ottenere risultati. Basta una persona interna all'azienda (o meglio almeno un paio per una azienda di 50 dipendenti) o, in alternativa, rivolgersi a professionisti esterni. Le competenze necessarie al Data Scientist potrebbero sembrare troppo elevate – dimestichezza con il campo della statistica, data mining, data base management, competenze web e social network – ma bisogna ricordare che un analista di dati deve capire cos'è un algoritmo, non lo deve sviluppare. Deve avere una base di competenze tecniche e statistiche associata alla capacità di comunicazione, perché i prodotti di intelligence devono essere comunicati poi al management. Ecco perché la formazione del Data Scientist prevede diversi percorsi di laurea tra quelli maggiormente orientati: lauree in statistica, matematica, informatica, ingegneria informatica, ma anche scienze del-

la comunicazione arricchita con una parte scientifica e statistica».

### Non solo like

Il lavoro dei Data Scientist va a pescare informazioni di vario genere dal web, dai social network e dai programmi di instant messaging. Una delle fonti di informazioni utili più importanti sono i social, come Facebook: dai post e dalle pagine è possibile raccogliere tutte le persone che inseriscono i like su un prodotto specifico, una destinazione o un'attività, per creare un data base che possa essere venduto ad una azienda interessata, quale target specifico a cui rivolgere le azioni di marketing. Ma ci sono anche altre attività aggiuntive svolte all'interno del cyberspazio che possono trasformarsi in preziose informazioni per l'analista di dati: le richieste di contatto, i commenti, le immagini postate e la duplice comunicazione testuale e visual... Il Data Scientist raccoglie ed elabora questa mole di dati, utilizzando appositi software anche gratuiti e opensource, per poi comunicare i risultati dell'analisi – data base profilati e analisi predittive di tendenze – al management che su questa base potrà definire strategie commerciali e comunicative.