

Introduzione

Togliamoci il pensiero: hai fatto bene a sfogliare questo libro e farai benissimo ad acquistarlo. E se lo hai già comprato, be', vorrei esserti vicino per stringerti la mano. Nelle prossime pagine ti spiegherò il perché di questo entusiasmo, ma intanto sappi che, al termine della lettura di questo libro, avrai acquisito un sacco di tecniche e strumenti per scovare informazioni, spesso nascoste, dai social media e dal Web. Immagina il potenziale nel tuo lavoro e nelle piccole e giocose investigazioni che tutti, prima o poi, facciamo su qualcuno con cui abbiamo a che fare. Ma prima, qualche spiegazione.

Introduzione all'Introduzione

Che tu ci creda o no, l'Introduzione è una delle parti più difficili da realizzare in un libro e di solito si scrive alla fine. È un mondo che va alla rovescia, ma questo lo sapevamo già. Un'Introduzione, idealmente, ti deve dare un assaggio di quel che andrai a leggere. Per convincerti ad acquistare il libro, nel caso tu lo stia sfogliando in libreria, o per confermarti che hai fatto bene ad acquistarlo, nel caso tu sia già passato alla cassa, fisica o virtuale che sia. Sono un tipo pragmatico e te lo ripeto: compralo, che fai bene, e se lo hai già acquistato hai fatto benissimo. Stabilito questo, potresti già passare al Capitolo 1, ma siccome oltre a essere pragmatico sono anche volitivo e ligio al dovere, voglio spiegarti per filo e per segno perché hai bisogno di questo libro, perché può cambiare la tua vita (conteniamoci e diciamo il tuo lavoro, suvvia), che cosa ci trovi dentro e perché mi sono azzardato a rivolgermi a te col "tu". Insomma, per la prima volta in quindici anni e dopo quasi quaranta libri scritti, ecco un'introduzione degna di questo nome.

Social media... che?

Questo è un libro che parla del *social media mining* e il suo contenuto sta tutto nel titolo. *Social media* è facile da comprendere, ma nel Capitolo 1 capirai che è una definizione ballerina e che nasconde dettagli tecnici che non avresti mai considerato. Comunque, per intenderci, mi riferisco a quei servizi web che allietano le nostre giornate e fanno dannare capi e clienti di fronte a qualche calo di produttività. E quindi ecco i vari Facebook, Instagram, Twitter e via dicendo. Nel novero rientra anche LinkedIn, che è un

social media legato all'ambito lavorativo ma che si sta progressivamente avvicinando, per concetti e interfaccia, agli altri esimi colleghi. E poi c'è quella parolina, *mining*, spesso fonte di fraintendimenti, che in italiano sta per "estrazione". Per esempio, se ne sente parlare spesso con i bitcoin e le criptovalute, ma in quel caso è sbagliato, perché la moneta virtuale non si estrae ma si genera, differenza sottile ma importante. Nel nostro caso, invece, la scelta del vocabolo è perfetta, perché *social media mining* significa, in soldoni, estrarre dati dai social media. Ti è facile intuire quanti dati siano stipati in un social media. Facebook, per esempio, qualche anno fa dichiarava di gestire, ogni giorno, 4 petabyte di dati, vale a dire 4 milioni di gigabyte o 4000 terabyte, ma al giorno d'oggi si stima che questa cifra sia circa dieci volte più grande. Non è difficile da credere se fai un piccolo esperimento, cioè scaricare i dati che Facebook ha di te, da quando hai creato il tuo account, e osservarne le dimensioni (il file coi miei dati, per esempio, pesa 1,3 gigabyte, e non sono nemmeno un utente molto attivo). Non lo sai fare? In questo libro ti spiego tutto per filo e per segno, ma sarà solo un assaggio di quel che imparerai leggendomi.

Informazioni alla portata di tutti

L'obiettivo di *Social Media Mining*, infatti, è quello di spiegarti come funzionano il Web e i social media da "dietro le quinte", e poi come si creano e sfruttano strumenti e tecniche per raccogliere buona parte dei dati che li popolano. Non solo i tuoi, sia chiaro, ma soprattutto quelli degli altri. È legale? Troverai risposta anche a questa domanda a breve, ma in linea di massima dipende dai dati e da come sono stipati. Questo libro non nasce per trasformarti in un criminale informatico ma ti vuole solo far capire come chiunque, oggi, sia in grado di estrarre informazioni, pubbliche e non, di un singolo individuo o di una grande massa di utenti. E perché mai qualcuno dovrebbe fare una cosa simile? Be', considera che viviamo nell'epoca del mercato dei dati, che sono un po' l'oro del presente e del futuro. E anche con tutte le precauzioni in fatto di privacy di cui si sono dotati proprio Facebook, Twitter, Instagram e colleghi vari e assortiti, resta il fatto che buona parte delle informazioni viene condivisa dagli utenti in modo consapevole o accettando condizioni che le espongono al mondo intero. I problemi, per chi ne vuole attingere, a questo punto sono tre.

- Il primo è che i dati, in genere, sono custoditi all'interno del rispettivo social media, quindi sono difficili da raggiungere da chi ne è all'esterno o da chi non fa parte di determinate "bolle" di utenti.
- Il secondo è che un conto è estrarre informazioni di un individuo, un altro farlo in modo massiccio, ponendosi come obiettivo quelle di un'ampia cerchia.
- E poi c'è la sempiterna questione dell'abbondanza. È pur vero che i latini dicevano *melius est abundare quam deficere*, ma di fronte a una mole esorbitante di dati si rischia di perdere di vista quelli davvero importanti o che servono ai tuoi scopi.

La soluzione a tutti i problemi

Leggendo questo libro troverai soluzione a tutti e tre i problemi (e qualcuno in più, come scoprirai), ma senza pozioni magiche o riti propiziatori come scendere dal letto al

mattino appena svegli e saltellare 99 volte su una gamba sola. La magia, in questo caso, sta tutta nella tecnologia: scrivere alcuni programmi e, se vuoi integrarli o non hai tempo, scoprire dei servizi già pronti all'uso per estrarre informazioni dai social media. Non sai programmare? Sorpresa: il libro che tieni in mano risolve anche questo problema, fornendoti il minimo indispensabile per scrivere codice Python e arrivare all'obiettivo. A questo punto potresti scambiare davvero *Social Media Mining* per un cilindro magico, ma siccome parliamo di informatica e non di divinazione runica, è bene specificare che questo è un libro dedicato a chi non sa nulla di questi argomenti e non ha mai sfiorato un linguaggio di programmazione. Ci tengo a sottolinearlo, perché è un testo che fa delle semplificazioni uno dei principali punti di forza, ma qualche esperto potrebbe storcere il naso. Ma anche se sei tra questi, sono piuttosto convinto che troverai materiale interessante nei capitoli più avanzati.

A chi si rivolge?

Social Media Mining si rivolge a chi ha bisogno di estrarre informazioni dai social media, per lavoro o diletto, che si tratti di concentrarsi su un account o su una pagina specifica, o di considerare un pubblico esteso. Acquisite le nozioni necessarie, potrai impiegarle nei più disparati contesti. Marketing, pubbliche relazioni, *advertising*, ricerche di mercato, ma anche investigazioni digitali, le fasi iniziali di un *penetration test*, giornalismo investigativo e tanto altro ancora. Senza contare che, quanto imparerai in queste pagine, ti tornerà utile anche se vuoi capire come fanno i social media ad acquisire così tanti dati di noi. Se c'è una cosa che non manca a *Social Media Mining* è la versatilità.

Ma è legale?

Rispondo a questa domanda nel Capitolo 9, ma se questo dubbio ti assilla fin da ora ti darò un assaggio della risposta: "Dipende...". *Social Media Mining* ti mette in mano degli strumenti; come li utilizzerai dipende solo da te e dai tuoi obiettivi. Di base, queste pratiche non hanno alcun tipo di vocazione, né benevola né malevola. Estrarre informazioni, come ho appena detto, si presta a una grande varietà di scopi e questi a loro volta possono avere finalità etiche oppure no. Un coltello da sushi impugnato da uno chef diventa uno strumento capace di dare gioia e godimento ai clienti; mettilo nelle mani di Jack lo Squartatore e... non credo che chi gli sarà di fronte potrà esserne felice. In linea di massima, puoi tarare uno strumento di *social media mining*, come quelli che imparerai a creare e utilizzare in questo libro, per fare in modo che rispetti tutte le normative sulla privacy, ma forzando certi parametri correrai sul limite della legalità. *Social Media Mining* è dunque un libro maledetto? Assolutamente no. Nozioni ed esempi che trovi in queste pagine sono presentati a scopo educativo e senza alcuna finalità illegale, quindi imparare e sperimentare nei limiti di legge non farà di te un cattivone del Web. Insomma, leggere questo libro non ha effetti collaterali, se non quello di imparare qualcosa di nuovo.

Quali sono i prerequisiti?

Per leggere questo libro non hai bisogno di sapere nulla, anche se una minima conoscenza di cosa sia un social media è quanto mai gradita. Hai Facebook o Twitter o Instagram o LinkedIn o addirittura tutti e altri ancora? Ottimo! È quel che ti serve per leggermi. Poi ti serve un computer, e te ne parlerò meglio nei prossimi capitoli. Fare *social media mining* con smartphone o tablet è possibile, ma difficile. Meglio quindi puntare su un computer Windows o Mac. Non ci sono richieste hardware o software particolari, anche se è meglio poter contare su una buona connessione Internet.

Perché questo stile?

In questo libro utilizzo il “tu” e un tono (molto) colloquiale, con qualche pizzico di ironia, perché... è il mio stile, fin dal primo libro che scrissi, molto tempo fa. C'è un buon motivo per cui lo utilizzo: creare un rapporto “faccia a faccia” col lettore, cioè con te. Mi piace l'idea di portarti dentro queste pagine e fartele vivere, più che leggere. E poi mi piace pensare che leggerai questo libro anche nel tempo libero. Dato che so bene quanto sia prezioso il tempo libero, voglio che leggermi sia un'esperienza rilassante.

Struttura del libro

Questo libro è suddiviso in dieci capitoli, che spiegano l'argomento in modo graduale, passo dopo passo. Questo approccio ha il vantaggio di presentare un percorso accessibile a chi non sa nulla, anche se qualcuno, più esperto, potrebbe trovare ridondanti alcune nozioni. Il mio consiglio per questi ultimi è di leggere comunque tutto, sia mai che abbia infilato qualcosina di nuovo. Se poi vuoi un'anticipazione di quel che troverai, eccoti servito.

- *Capitolo 1, Com'è fatto il Web?* – Se credi di sapere tutto sul Web, riducendolo a un posto virtuale dove si naviga tra siti, leggendo queste pagine potresti scoprire diversi dettagli tecnici che forse non conosci. Anche la semplice navigazione e l'utilizzo del browser, in realtà, non sono così semplici come sembra.
- *Capitolo 2, Come sono fatti i social media?* – Facebook, Twitter, Instagram, LinkedIn e tanti altri: i social media sono sulla bocca di tutti, ma in pochi sanno davvero come funzionano dal di dentro. In questo capitolo, tante curiosità e, soprattutto, l'architettura interna e il funzionamento delle API dei social media.
- *Capitolo 3, Tutto quello che serve* – Un “maxi-capitolo” che ti insegna, passo dopo passo, tutto ciò che devi conoscere per gettarti nella mischia del *social media mining* vero e proprio. E quindi HTML, CSS, JavaScript e poi via, verso i rudimenti della programmazione Python. Ci sarà poi tempo, nei capitoli successivi, per approfondirla.
- *Capitolo 4, Mining di Facebook: i primi passi* – Facebook è un social media così vasto e complicato da richiedere ben due capitoli per essere trattato a dovere. Nel primo, una rassegna di tutti i dati che Facebook è in grado di assorbire, una dissertazione su come si utilizza la sua API e poi via, con i primi esperimenti con Python.

- *Capitolo 5, Mining di Facebook: diamoci dentro* – Il secondo capitolo dedicato al social media di Mark Zuckerberg entra nel vivo del discorso, parlando dell’analisi di una pagina Facebook e passando poi allo *scraping*, una tecnica di *mining* particolare e che si rivelerà un vero *passe-partout* per la raccolta di dati anche da normali siti web.
- *Capitolo 6, Twitter e i suoi segreti* – Leggendo questo capitolo non solo scoprirai tutta la tecnologia che muove questo social media e che ti è utile per effettuare un buon *mining*, ma ti appassionerai (spero) a realizzare alcuni software in Python per sfruttarne la API. In più, capirai perché Twitter è il social media più amato da chi è a caccia di informazioni.
- *Capitolo 7, Capire (e carpire) Instagram* – Per certi versi Instagram è un’“anomalia” tra i social media e questo capitolo te ne spiega le ragioni, sia dal punto di vista tecnico, sia per tipologia delle informazioni e tecniche di mining. Bonus: un’introduzione ai *tool* professionali, le ragioni e le tecniche per usarli.
- *Capitolo 8, Nei meandri di LinkedIn* – Mai sottovalutare l’unico vero social media dedicato al business! Questo capitolo ti spiega perché dovrebbe sempre rientrare in una buona strategia di *mining* e come raggranellare dati da uno dei servizi web meglio protetti dagli sguardi indiscreti.
- *Capitolo 9, Caccia grossa* – Qui si tirano le somme: tutte le nozioni e tecniche imparate finora sono messe in un bel calderone, insieme a qualche altro trucchetto, per confezionare dei software avanzati per la raccolta di informazioni e contenuti. Per trasformare quanto appreso in una professione o un potente strumento per il lavoro che già si fa.
- *Capitolo 10, Analisi dei dati* – Che te ne fai di informazioni e dati se poi non li sai analizzare? Sebbene la *data analysis* non sia il fulcro del libro, qui ti introduco ai suoi concetti di base con un approccio molto pratico, grazie al quale svilupperai potenti software per calcolare il *sentiment* nei social media. Insomma, se si parla di un’azienda, un *brand* o una persona nei social media, riuscirai a capire se lo si fa con tono positivo o negativo.

Ringraziamenti

Un libro è un percorso. Un percorso che va dalla sua prima alla sua ultima pagina, innanzitutto. Un percorso che va dal primo giorno in cui inizi a scriverlo a quando consegni l’ultimo paragrafo (che è proprio quello che stai leggendo ora). Soprattutto, un libro è un percorso che abbraccia tutta la tua vita: esperienze, persone, nozioni imparate, momenti esaltanti e momenti di sconforto, libri letti e altri libri scritti. Più libri scrivo, e più questo percorso si fa lungo e ricco di stimoli. Figurarsi le persone incrociate, o che mi stanno vicino, e che magari hanno dato spunto a una parola, un titolo, una pagina, un paragrafo o un intero capitolo. Siete tantissime, e vi ringrazio tutte, ma quelle importanti sanno perfettamente chi sono, perché vivono in me e io in loro.

Un grazie va alla pazientissima Giulia, che ha stabilito il record di puntualità di una mia consegna (e prometto di aiutarti a migliorarlo), e a Fabio, che ridendo e scherzando... son quindici anni passati a lavorare insieme.

Grazie a chi crede in me e non ha mai smesso di crederci, grazie a chi ci crede meno e mi dà la carica per darci ancora più dentro. Grazie a chi c'è sempre e grazie a chi non c'è più ma, chissà da dove, sorride fiero e con gli occhi lucidi e commossi. Sono certo che la pioggia che segna la finestra del mio studio, ora, sono le vostre lacrime. Qualcuna, in effetti, riga anche il mio viso mentre scrivo. Vi renderò orgogliosi di me.