

AIS 2017.

Big Data, Big Challenges. Convegno di metà mandato della Sezione di Metodologia.

Titolo: Twitter, Twitter delle mie brame, come sono le università italiane?

Esplorare la percezione sociale delle università italiane su Twitter tramite i Big Data

Area tematica: 6. L'analisi dei Big Data

Autori

Biella Marco, m.biella6@campus.unimib.it, Università degli Studi di Milano – Bicocca,

Zogmaister Cristina, cristina.zogmaister@unimib.it, Università degli Studi di Milano – Bicocca,

Ceolato Stefania, s.ceolato@campus.unimib.it, Università degli Studi di Milano – Bicocca,

Parozzi Elena Alessandra, e.parozzi@campus.unimib.it, Università degli Studi di Milano – Bicocca

Long Abstract (13656 caratteri)

Ogni Università italiana sta diventando, tra le altre cose, un brand che compete nel mercato con altri brand simili, per attrarre i migliori studenti e per ottenere finanziamenti pubblici e privati. Per questa ragione è importante capire in quale modo essa venga percepita e quale immagine dia ai suoi interlocutori, in altre parole quale sia la sua brand image. Tipicamente esistono diversi modi per indagare la brand image, i più tradizionali sono quelli basati su intervista, questionario e focus group, ma in epoca recente e con la diffusione di internet e dei social network, anche questi strumenti possono diventare importanti fonti d'informazione per capire come un'organizzazione viene percepita all'esterno.

Nel presente lavoro proponiamo un metodo per indagare l'immagine delle università italiane, o in altre parole per studiare alcuni aspetti importanti della percezione che le persone hanno di tali università, a partire dalle tracce digitali che queste lasciano quando utilizzano i social network. Il nostro metodo si propone di fornire le proprietà quantitative necessarie per un'indagine rigorosa rimanendo in contatto con il contenuto della brand image delle università e la rappresentazione qualitativa della realtà in cui gli

elementi dell'indagine si muovono. Un elemento fondamentale di qualsiasi lavoro di ricerca è l'osservazione di un fenomeno e il suo collocamento nel contesto nel quale si manifesta. Per poterlo inquadrare correttamente è necessario svolgere un'operazione di misurazione. Tali necessità ci hanno portati a focalizzarci sullo studio della brand image delle università italiane e ad elaborare una tecnica per poterle confrontare le une con le altre.

All'interno del contesto dei Social Network, nello specifico ci siamo focalizzati su Twitter e pertanto il nostro studio riguarda in particolare il modo in cui gli utenti di Twitter percepiscono le università italiane. La scelta di utilizzare questo Social Network in particolare è imputabile a due ragioni: la prima è che Twitter è un Social Network molto diffuso tra il pubblico generale, la seconda è che la maggior parte degli account Twitter sono pubblici, o in altre parole accessibili in forma libera, a differenza di altri social network come Facebook che invece sono caratterizzati dalla presenza di molti account privati, ossia accessibili soltanto dopo che l'utente ha stabilito uno specifico contatto (quella che viene chiamata "amicizia").

L'obiettivo del nostro lavoro è quello di misurare alcune dimensioni, definite a priori sulla base degli scopi della nostra ricerca, attraverso le quali mappare il collocamento delle diverse università italiane. A questo scopo ci siamo basati su un precedente lavoro di Culotta & Cutler (2016; Permalink:

<http://dx.doi.org/10.1287/mksc.2015.0968>) in cui veniva studiata la brand image di organizzazioni commerciali. Ci siamo serviti di una versione modificata della loro tecnica applicandola alla brand image delle università considerate. Un'altra differenza con il lavoro da cui siamo partiti sono l'impiego di tecniche di analisi della brand image nel caso di organizzazioni che non sono produttori di prodotti o servizi ma produttori di conoscenza e che, in quanto tali, verranno percepiti inevitabilmente in maniera diversa dai primi. Il principale contrasto nella percezione tra organizzazioni profit e le università risiede nella peculiarità delle dimensioni principali di valutazione (es. ricerca, occupabilità lavorativa degli iscritti, internazionalizzazione, ecc...). Il principale punto di contatto tra i due lavori invece risiede nell'analizzare la brand image e la percezione sociale indagando alcuni aspetti delle 'tracce digitali' che rimangono al seguito dei online nei Social Network.

In entrambi i casi sono state infatti analizzate alcune tracce del comportamento online e come parametro di confronto abbiamo utilizzato alcuni account prototipici per ognuna delle diverse dimensioni considerate.

Il comportamento oggetto della presente analisi è infatti costituito dalla scelta di seguire o meno un determinato account. Gli account rilevanti sono sia quelli considerati prototipici per le specifiche dimensioni (come vedremo meglio in seguito) sia gli account target della nostra indagine, ossia gli account universitari. Tale approccio ci ha permesso di sfruttare l'enorme mole di dati già presente in rete, i Big Data, differenziando questo metodo dalle principali metodologie di analisi della brand image (es. interviste, focus group, ecc...), che invece richiedono la raccolta di nuovi dati e il recupero attivo dell'informazione da parte dell'intervistato.

Uno dei principali vantaggi dell'uso di questo tipo di dati è la loro elevata ecologicità. Essi possono essere infatti considerati come il risultato di comportamenti prodotti spontaneamente, in aperta contrapposizione con i dati "causati" dagli sperimentatori nei tipici studi sulla percezione sociale. Tale prospettiva consegue dalla libera scelta di seguire un account fatta dagli utenti ai quali non viene posto nessun vincolo. Questo aspetto di spontaneità che caratterizza i comportamenti nei social network – e perciò le tracce di questi comportamenti - li qualifica come una risorsa particolarmente interessante per comprendere il modo in cui le persone si rappresentano un concetto d'interesse, poiché il comportamento avviene naturalmente, senza interferenze esterne dovute alla richiesta di esprimere un'opinione, una valutazione o un atteggiamento come invece accade con le tradizionali tecniche della domanda.

Il metodo proposto si basa su tre fasi. Nella prima fase vengono scelte delle dimensioni rilevanti per la brand image. Nella seconda fase vengono individuati degli account d'interesse - in questo caso gli account delle diverse università di cui si vuole analizzare la brand image - e un gruppo di account prototipici per ciascuna delle dimensioni. Infine, viene fatta un'analisi del grado di similarità tra gli account target (ossia gli account universitari) e gli account prototipici di ogni singola dimensione. In seguito a questa terza fase, ad ogni account target viene assegnato un punteggio per ciascuna delle dimensioni d'interesse, in base alla similarità con una serie di account prototipici. Più nel dettaglio, la similarità è operazionalizzata in termini di sovrapposizione di follower tra l'account d'interesse e gli account prototipici della dimensione. Ciò avviene tramite il calcolo di un Jaccard Similarity Coefficient. Tale indice riassume il rapporto tra l'intersezione dei follower dell'account prototipico e del target e la loro unione. Per ogni account target è calcolato un indice

di similarità con ogni account prototipico e l'aggregazione di tali indici fornisce un punteggio dell'account target lungo la dimensione.

Dal punto di vista teorico il nostro metodo si basa due presupposti. Il primo è che la decisione di seguire account prototipici di una certa dimensione sia un indicatore di interesse per tale dimensione (se una persona segue per esempio molti account che hanno a che fare con il tema dell'internazionalizzazione, è plausibile che questa persona abbia un elevato interesse per l'internazionalizzazione). Il secondo è che se molte persone interessate a una certa dimensione segue l'account target, ciò sia un indicatore che anche l'account target sia percepito come rilevante per tale dimensione. In altre parole, se per esempio dall'analisi emergesse che molti follower dell'Università di Milano - Bicocca sono follower di account relativi all'internazionalizzazione, è plausibile che anche l'Università di Milano Bicocca sia percepita come rilevante per l'internazionalizzazione. L'intersezione tra follower dell'account target e follower degli account prototipici di ogni dimensione si propone di quantificare questo aspetto.

Diretta conseguenza di questo legame è la possibilità di quantificare la percezione delle università lungo le dimensioni basandoci sulla scelta degli account prototipici. Tale conseguenza può garantire la possibilità di indagare dimensioni formate ad-hoc.

Nella fattispecie della nostra ricerca, le principali dimensioni rilevate riguardano aspetti peculiari delle università di interesse per gli studenti, come ad esempio la didattica, l'internazionalizzazione o l'occupabilità degli iscritti, o per possibili enti finanziatori, come ad esempio la ricerca, l'innovazione e la tecnologia, l'economia. Una volta chiarite le dimensioni lungo le quali profilare gli account target sono stati individuati account prototipici per ogni dimensione. Per quanto riguarda la dimensione "ricerca", per esempio, sono stati individuati account come quello dell'European Research Council (@ERC_Research), quello dell'ISTAT (@istat_it) o quello dell'ufficio stampa del CNR (@StampaCnr). La scelta degli account prototipici è stata svolta tramite la negoziazione tra gli autori e la discussione della rappresentatività del singolo account. Una volta definite dimensioni e account si è passati alla fase di calcolo del punteggio.

Andando in maggiore dettaglio circa il punteggio, dal punto di vista analitico esso è un'aggregazione degli indici di similarità tra l'account target e tutti gli account prototipici di una particolare dimensione. Tale aggregazione pondera il peso di ogni account prototipico per l'inversa della sua popolarità,

operazionalizzata come numero di follower. Diretta conseguenza di questo metodo è il diverso peso in termini di punteggio finale attribuito ad account specifici e generalisti. Un account prototipico particolarmente generalista, quindi con molti follower, avrà un impatto minore sul punteggio finale rispetto ad un account considerato specifico per la dimensione considerata. Così facendo viene limitata l'influenza di account estremamente popolari che potrebbero presentare similarità con gli account target imputabili esclusivamente alla loro estrema diffusione e non a una similarità di contenuti o di percezione, ovvero dovuta al fatto che chi segue l'account prototipico lo faccia per ragioni non legate alla specificità dell'account stesso.

Per valutare la bontà delle nostre dimensioni abbiamo proceduto tramite un confronto tra i punteggi ottenuti e criteri esterni. Ad esempio, tra le dimensioni abbiamo scelto "economia" non solo per la rilevanza generale, ma anche perché una delle università target (Università Bocconi) è fortemente caratterizzata su questa dimensione. Un'altra dimensione, "religione cattolica", sebbene in generale non sia rilevante nell'immagine di un'organizzazione universitaria, è stata inclusa nell'analisi perché un'altra università target (Università Cattolica del Sacro Cuore) ha un forte legame con questa dimensione. Perciò il primo passo nel processo di validazione delle dimensioni è consistito nel verificare se i punteggi sulle dimensioni rispecchiassero queste aspettative. Coerentemente con le nostre ipotesi, l'università con background cattolico ha ottenuto un punteggio nettamente superiore a tutte le altre per quanto concerne la dimensione "religione cattolica" e l'università a chiara vocazione economica ha ottenuto il massimo punteggio per la dimensione "economia".

Una volta profilati gli account target il passo successivo riguarda la mappatura del panorama percettivo entro il quale si collocano. Tale obiettivo è stato raggiunto tramite una cluster analysis che, sfruttando la natura quantitativa dei punteggi lungo le dimensioni appena ottenute, permette di apprezzare le posizioni relative dei singoli account gli uni rispetto agli altri. Questa tecnica consente di aggregare i diversi account in gruppi basati sulla loro similarità secondo il principio per il quale account simili sono uniti tra di loro in un gruppo i cui elementi condividono una percezione comune agli occhi degli utenti.

A seguito della mappatura della percezione degli account target si ottiene un quadro d'insieme dell'ecosistema consentendo di apprezzare le collocazioni relative dei diversi elementi considerati. Questa

mappa permette di raggruppare gli account d'interesse in cluster basati sulla similarità ed apprezzare come questa popolazione si stratifichi in termini di percezione sociale.

A fianco della mappatura dello spazio sociale nel quale si muovono i nostri target il metodo proposto consente dei confronti diretti. Infatti una volta "misurata" la brand image di un target la si può confrontare direttamente con quella dei suoi competitor, siano essi organizzazioni profit o non-profit. Quest'analisi fornisce indici utili sia per capire il panorama sociale dove gli elementi d'indagine agiscono sia per implementare una strategia di comunicazione mirata ad avvantaggiarsi rispetto a quelli che sono competitor a tutti gli effetti. Così facendo un ente di ricerca, un'organizzazione commerciale o un'università possono ottenere informazioni utili per riposizionarsi rispetto ad altri agenti con i quali competono per attrarre finanziamenti, studenti o quote di mercato.

Infine un ulteriore vantaggio del nostro metodo è la possibilità di assegnare agli utenti un'importanza relativa per ogni dimensione osservata. Infatti un follower condiviso tra più account prototipici, una volta acquisito da un account target, avrà un impatto maggiore sul punteggio lungo la specifica dimensione. Tale informazione è estremamente utile sia in un'ottica di intervento che in un'ottica d'indagine. Se infatti un'università interessata a riposizionarsi in termini di percezione sociale può individuare una popolazione specifica di follower da acquisire, i ricercatori possono sfruttare queste informazioni per aggregare le specifiche popolazioni peculiari di ogni dimensione della percezione sociale.