

ALMA MATER STUDIORUM - UNIVERSITA' DI BOLOGNA

SCUOLA DI SCIENZE POLITICHE

Corso di laurea magistrale in

Comunicazione Pubblica e d'Impresa

**BIG DATA, COMUNICAZIONE E MARKETING NEL TURISMO 2.0:
UNO STUDIO DI CASO PER L'HOSPITALITY SECTOR**

Tesi di laurea in

Big Data for the Social Sciences

Relatore: Prof. Marco Albertini

Correlatore: Prof.ssa Pina Lalli

Presentata da: Lorenzo Marcucci

Appello
secondo

Anno accademico
2017-2018

Indice

Introduzione

- 1. I Big Data e la comunicazione d'impresa**
 - 1.1. Che cosa sono i Big Data?
 - 1.2. Cambiamento di paradigma
 - 1.3. Il prezzo dei grandi numeri: *messiness* dei Big Data
 - 1.4. L'uso dei big data nella comunicazione e gestione di impresa
 - 1.4.1. La profilazione dei clienti
 - 1.4.2. La rilevanza dell'accesso ai dati e delle competenze di analisi nel settore della comunicazione aziendale
 - 1.5. Big Data e customer satisfaction
 - 1.6. Prospettive future
- 2. L'hospitality sector in Italia**
 - 2.1. La situazione attuale
 - 2.1.1. Caratteristiche del settore hospitality in Italia
 - 2.1.2. La forza lavoro nell'hospitality sector
 - 2.2. Il progresso tecnologico e la sfida del turismo 2.0
 - 2.2.1. L'hospitality sector italiano e il turismo 2.0
 - 2.3. Big data analytics per il settore hospitality
- 3. Lo studio di caso: recensioni di hotel di fascia alta in Italia**
 - 3.1. Selezione e raccolta
 - 3.2. Pulizia e strutturazione dei dati
 - 3.3. Pre-elaborazione e analisi del testo
 - 3.4. Segmentazione per nazionalità
- 4. Risultati osservati e proposte di sviluppo**
 - 4.1. Analisi descrittiva
 - 4.1.1. Le differenze tra gruppi
- 5. Approfondimenti: la messiness nelle recensioni e il caso di Bologna**

5.1. Debolezza intrinseca dell'analisi automatizzata delle reviews

5.2. Recensioni degli hotel nei giorni di fiera

5.3. Il caso di FICO

Conclusione

Allegato 1. Comunicato stampa Federalberghi, 20 agosto 2018

Allegato 2. Workflow di Knime per l'elaborazione delle recensioni su Bologna

Bibliografia

Introduzione

In Italia il settore alberghiero sta vivendo in modo complicato il rapporto con il web 2.0 e l'evoluzione della domanda turistica. I nuovi intermediari online hanno del tutto – o quasi – rimpiazzato le agenzie di viaggio e la lettura delle recensioni online degli utenti sembra essere il criterio principale nel prendere decisioni d'acquisto sui prodotti turistici. Si è deciso di sviluppare questo elaborato con l'intento di presentare una proposta analitica a vantaggio del settore, basata sulle stesse tecnologie che sembrano causare tanti problemi agli addetti ai lavori. In seguito alla crescente popolarità e al relativo consenso acquisiti recentemente dalle tecniche di Big Data analysis, in questa sede se ne studieranno alcune delle potenzialità per l'hospitality sector. A partire dall'ambito del marketing ma non limitatamente ad esso. Infatti, sono numerose le opportunità da esplorare per le aziende in generale e per gli operatori del sistema ricettivo italiano in particolare.

L'obiettivo di questo elaborato è proporre un nuovo metodo di analisi di mercato, basato sui contenuti generati dagli utenti, auspicabilmente in grado di suscitare ulteriore interesse verso questo approccio, mettendo in evidenza le potenzialità di creazione di valore aggiunto per quelle imprese che volessero adottare un approccio simile. Nello specifico, si è proceduto indagando su quali fossero i fattori sottostanti alla soddisfazione o insoddisfazione dei clienti rispetto al loro soggiorno, cercando di individuare i criteri principali alla base dei loro giudizi. Per raggiungere questo obiettivo è opportuno osservare quali siano i termini più frequentemente utilizzati dagli utenti per esprimere le loro opinioni su Booking.com, uno dei canali di prenotazione più utilizzati per le strutture alberghiere. Lo stesso processo è stato eseguito sia considerando l'utenza nel suo insieme, sia andando a selezionare gruppi di persone in base all'area geografica di appartenenza, in modo tale da gettare le basi per una segmentazione del mercato di riferimento. Per fare ciò, è stato sviluppato un algoritmo di analisi che elabora testi non strutturati, in questo caso estratti dalle recensioni online relative ad alberghi italiani di fascia alta, liberamente reperibili su Booking. Nel dettaglio, la selezione è ricaduta sulle recensioni scritte in lingua inglese per tutti gli alberghi a quattro e cinque stelle presenti nelle città di Bologna, Firenze, Torino, Verona e Venezia. Dopo aver effettuato l'estrazione dei dati in questione tramite la creazione dell'apposito *web crawler*, si è proceduto all'elaborazione degli stessi mediante un software per il data mining. Entrambi gli strumenti scelti sono open-source - ad accesso

libero - e si basano sulla semplicità di utilizzo, tanto da non richiedere competenze tecniche di programmazione informatica.

La tesi è articolata in cinque capitoli. Nel primo, verranno introdotti i concetti alla base del lavoro, partendo dalla definizione stessa di Big Data e ragionando sulle potenzialità che possono offrire alla ricerca e alle imprese in generale. Nel secondo, invece, ci si concentrerà sull'hospitality sector italiano, prima presentandone un quadro generale aggiornato, poi andando ad osservare l'effetto che i Big Data e i relativi metodi e tecnologie stanno apportando al giorno d'oggi in questo ambito. Il terzo capitolo racchiude la parte tecnico-pratica dell'elaborato, presentando tutte le fasi di estrazione ed elaborazione dei dati in questione. I risultati di questo trattamento saranno esposti nel dettaglio ed analizzati nel quarto capitolo, dove si cercherà di ricostruire i profili di consumo turistico delle varie categorie di utenza. Nell'ultimo capitolo, infine, si proporranno degli approfondimenti relativi a casi specifici: la debolezza intrinseca delle recensioni di Booking.com; l'effetto delle fiere nella città di Bologna sulla soddisfazione degli utenti; l'indagine sull'apertura del nuovo parco agroalimentare FICO Eataly World per questa fascia di utenti. Nella conclusione finale dell'elaborato si riassumerà come, attraverso questo lavoro, sia stato possibile mettere alla prova l'approccio induttivo tipico delle tecniche di Big Data analysis e ottenere degli spunti interessanti per il mercato dell'hospitality sector.

Conclusione

Prendendo le mosse dal processo di *text mining* condotto sulle recensioni di Booking.com per gli hotel a quattro e cinque stelle di Bologna, Firenze, Torino, Venezia e Verona, è stato possibile ottenere diversi spunti per l'hospitality sector. La cosa particolarmente interessante è che la quantità di interrogativi a cui abbiamo cercato di dare una risposta, o per cui si ritiene valga la pena approfondire gli studi, è superiore al numero di domande di ricerca che hanno dato le mosse a questo lavoro. Ciò è possibile grazie all'adozione dell'approccio induttivo, reso possibile dai Big Data e dalle tecniche di analisi che vi si applicano. Partendo da una domanda di ricerca poco definita, si è lasciato che i dati raccolti ci guidassero nella ricerca di aspetti che valesse la pena approfondire, semplicemente tramite poche operazioni di strutturazione e analisi visiva.

I risultati sembrano promettenti: è stato possibile identificare tendenze generali, come la grande importanza attribuita alla colazione consumata in albergo, e anche a individuare differenze nelle abitudini di consumo degli utenti di diverse nazionalità. Successivamente, sulla falsa riga di quanto auspicato dagli studi teorici sui Big Data visti in apertura, si è proceduto ad approfondimenti specifici basati sull'applicazione di metodi di ricerca più tradizionali, fondati sull'approccio deduttivo, per rispondere ad interrogativi sorti solo dopo una prima fase di raccolta e pre-elaborazione dei dati. L'esempio più stringente è quello del lavoro sulla stima del tasso di confusione, o *messiness*, intrinseca alle reviews (par. 5.1): dopo aver formulato una domanda di ricerca, è stato eseguito un campionamento dei dati e testata l'ipotesi sui casi selezionati.

I risultati finali di questo progetto, dunque, possono essere ritenuti interessanti sia per il valore concreto ottenuto nel merito dello studio specifico sul settore alberghiero, sia per la dimostrazione di applicabilità della metodologia sottostante allo stesso. Infatti, a questo punto diviene interessante interrogarsi sulle potenzialità dell'approccio per altri ambiti più o meno strettamente legati al caso in esame, comprese le scienze sociali. Gli studi pregressi basati su approccio induttivo e tecniche di analisi automatizzata dei dati, di cui abbiamo presentato alcuni esempi nei primi due capitoli, sembrano offrire risultati promettenti e numerose potenzialità da approfondire. L'auspicio finale è che lavori di questo tipo possano contribuire al grado di fiducia riconosciuto alla metodologia in questione così che si possa averne sempre maggiore considerazione per applicazioni future.

Bibliografia

- Anderson, C. (2008, Giugno 23). *The End of Theory: The Data Deluge Makes The Scientific Method Obsolete*. Tratto da Wired: <https://www.wired.com/2008/06/pb-theory/>
- Associazione Italiana Confindustria Alberghi. (2018, Agosto 1). *Confindustria Alberghi: ad agosto cresce la voglia di #mareitalia ma non solo*. Tratto da Confindustria Alberghi: <http://www.alberghiconfindustria.it/tema/stampa-comunicazione/notizia/30163/confindustria-alberghi-ad-agosto-cresce-la-voglia/>
- Berezina, K., Bilgihan, A., Cihan, C., & Fevzi, O. (2016, Aprile 22). Understanding Satisfied and Dissatisfied Hotel Customers: Text Mining of Online Hotel Reviews. *Journal of Hospitality Marketing & Management*, p. 1-24.
- Buhalis, D., & Law, R. (2008, January 17). Progress in information technology and tourism management: 20 years on and 10 years after the internet-the state of eTourism research. *Tourism Management*, 29(4), p. 609-623.
- Cardon, D. (2016). *Che cosa sognano gli algoritmi - Le nostre vite al tempo dei big data*. Milano: Mondadori.
- Cicerchia, V. (2018, Agosto 7). *Starhotels gestirà l'albergo di Fico Eataly World Bologna*. Tratto da Il Sole 24 Ore: <http://www.ilsole24ore.com/art/impresa-e-territori/2018-08-07/starhotels-gestira-l-albergo-fico-eataly-world-bologna-113445.shtml?uuid=AEMuOvXF>
- Coleman, J. S. (1986, Maggio). Social Theory, Social Research, and a Theory of Action. *The American Journal of Sociology*, 91(6), p. 1309-1335.
- Confindustria Alberghi. (2018, Giugno 11). *Bene lavorare su classificazione e contro l'abusivismo, ma anche l'on-line deve fare la sua parte*. Tratto da Confindustria Alberghi: <http://www.alberghiconfindustria.it/tema/stampa-comunicazione/notizia/29825/bene-lavorare-su-classificazione-e-contro/>
- Court, D. (2015). Getting big impact from big data. *McKinsey Quarterly*.

- Cukier, K. (2010, Febbraio 25). Data, data everywhere. (T. Economist, Intervistatore)
- De Mauro, A., Greco, M., & Grimaldi, M. (2016). A formal definition of Big Data based on its essential features. *Library Review, Vol. 65, Issue 3*, p. 122-135.
- Del Chiappa, G. (2013). Internet versus travel agencies: The perception of different groups of Italian online buyers. *Journal of vacation marketing 19(1)*, p. 55-66.
- Dinis, G., Costa, C., & Pacheco, O. (2016). The Use of Google Trends Data as Proxy of Foreign Tourist Inflows to Portugal. *International Journal of Cultural and Digital Tourism*, 66-75.
- Federalberghi. (2018, Agosto 13). Abusivismo - una piaga da fermare. (ANSA, Intervistatore)
- Federalberghi. (2018, Agosto 20). Portali di prenotazione - Prosegue la pubblicità ingannevole. *Federalberghi chiede allo IAP di bloccare lo spot di Trivago*. Roma, Italia.
- Federalberghi, EBNT. (2018). *Datatur : Trend e Statistiche sull'Economia del Turismo*. Roma: Edizioni ISTA.
- Gandomi, A., & Heider, M. (2014, Dicembre 3). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, p. 137-144.
- Germann, F., Lilien, G. L., & Rangaswamy, A. (2012, Novembre 27). Performance implications of deploying marketing analytics. *Internal Journal of Research in Marketing*.
- Grandi, R. (2007). *La comunicazione pubblica, Teorie, casi profili normativi*. Roma: Carocci.
- Guiggiani, R. (2018, Giugno 8). *Fico Eataly World: bello senz'anima*. Tratto da Siena News: <http://www.sienanews.it/in-evidenza/fico-eataly-world-bello-senzanima/>

- Jingjing, L., Lizhi, X., Ling, T., Shouyang, W., & Li, L. (2018, Aprile 10). Big data in tourism research: A literature review. *Tourism Management*(68), p. 301-323. Tratto da <https://doi.org/10.1016/j.tourman.2018.03.009>
- Kitchin, R. (2014, Aprile 2014). Big Data, new epistemologies and. *Big Data & Society*, p. 1-12.
- Laney, D. (2001, Febbraio 6). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*.
- Marr, B. (2016). How Big Data and analytics are changing hotels and hospitality sector. *Forbes*.
- Mayer-Schonberger, V., & Cukier, C. (2014). *Big Data: A Revolution that will transform how we live, work and think*. Eamon Dolan/Mariner Books.
- McAbee, S. T., Landis, R. S., & Burke, M. I. (2016). Inductive reasoning: The promise of big data. *Human Resource Management Review*, p. 277-290.
- McKinsey Global Institute. (2011, Maggio). Big Data: The Next Frontier for Innovation, Competition, and Productivity. *McKinsey Global Institute Report*.
- Mingming, C., & Xin, J. (2018, Aprile). What do Airbnb users care about? An analysis of online review comments. *International Journal of Hospitality Management*(76), p. 58-70.
- Montagnoli, L. (2018, Aprile 18). *Enrico Bartolini lascia Fico. Chiude il ristorante Cinque all'interno di Eataly World*. Tratto da Gambero Rosso: <http://www.gamberorosso.it/it/news/1047259-enrico-bartolini-lascia-fico-chiude-il-ristorante-cinque-all-interno-di-eataly-world>
- Obal, M., & Kunz, W. (2013, Vol. 24 Issue: 1). Trust development in e-services: a cohort analysis of Millennials and Baby Boomers. *Journal of Service Management*, p. 45-63.

- Opresnik, D., & Taisch, M. (2015, January 5). The value of Big Data in servitization. *International Journal of Production Economics*.
- Oracle. (2018). *What is Big Data?* Tratto da Oracle.com: <https://www.oracle.com/big-data/guide/what-is-big-data.html>
- Ritzer, G. (2010, March 9). Production, Consumption, Prosumption. The nature of capitalism in the age of the digital 'prosumer'. *Journal of Consumer Culture*, p. 13-36.
- Seymour, S. (2017, Novembre 17). *Eataly World opens but leaves a bad taste in Bologna*. Tratto da The Guardian: <https://www.theguardian.com/travel/2017/nov/17/eataly-world-bologna-megamarket-betrayal-italian-gastronomy>
- Silveri, A. (2018, Aprile 26). *Voi avete capito se Fico Eataly World è un successo o un fiasco?* Tratto da Dissapore: <https://www.dissapore.com/notizie/fico-eataly-world-e-un-successo-o-un-fiasco/>
- Talón-Ballester, P., González-Serrano, L., Soguero-Ruiz, C., Muñoz-Romero, S., & Rojo-Álvarez, J. L. (2018). Using big data from Customer Relationship Management information. *Tourism Management*, 68, p. 187-197.
- Tonelli, M. (2017, Novembre 13). *Fico Eataly World a Bologna. Cos'è, com'è nato, come funziona, cosa ci piace e cosa no*. Tratto da Gambero Rosso: <http://www.gamberorosso.it/it/food/1046289-fico-eataly-world-a-bologna-cos-e-com-e-nato-come-funziona-cosa-ci-piace-e-cosa-no>
- Zheng, X., Qianzhou, D., Yufeng, M., & Weiguo, F. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*(58), p. 51-65.