

Maurizio Massaro e Marika Gon

4. Metodo e analisi qualitativa dei risultati del Forum Turismo in Friuli Venezia Giulia

Riassunto: Elemento centrale di ogni ricerca scientifica è la definizione della metodologia e del metodo di riferimento. Nel caso del Forum TurismoFVG, la ricerca ha voluto analizzare la prospettiva di molteplici operatori e stakeholder del settore, organizzati in tavoli di discussione. Partendo dagli obiettivi specifici del lavoro e dalle fonti informative disponibili, si è ritenuto che uno studio qualitativo basato sulla metodologia della content analysis fosse particolarmente adatto. L'utilizzo di approcci sia manuali che computer based ha consentito di ampliare la base di elaborazione a disposizione, arricchendo i processi interpretativi dei risultati ottenuti. Nel capitolo si presentano gli approcci di ricerca utilizzati per processare ed analizzare le informazioni prodotte dai partecipanti al Forum FVG.

Parole chiave: Analisi qualitativa, Analisi testuale, Bag of words, Leximancer

Abstract: The central element of every scientific research is the definition of the methodology. In the case of the Tourism Forum of Friuli Venezia Giulia region, the research goals wanted to analyze the perspective of multiple operators and stakeholders in the tourism sector, organized into discussion tables. In agreement with these goals, it was deemed that a qualitative methodology, based on content analysis, was particularly suitable. The use of both manual and computer based approaches made it possible to expand the available processing base, enriching the interpretative processes of the results obtained. The chapter presents the research approaches used to process and analyse the information produced by the participants in the FVG Forum.

Keywords: Quality analysis, Text analysis, Bag of words, Leximancer

Contenuto in: Processi partecipativi nella progettazione turistica

Curatori: Francesco Marangon, Marika Gon, Maurizio Massaro e Andrea Moretti

Editore: Forum

Luogo di pubblicazione: Udine

Anno di pubblicazione: 2018

Collana: Ambiente e territorio

ISBN: 978-88-3283-078-1

ISBN: 978-88-3283-079-8 (versione digitale/pdf)

Pagine: 43-50

Per citare: Maurizio Massaro e Marika Gon, «4. Metodo e analisi qualitativa dei risultati del Forum Turismo in Friuli Venezia Giulia», in Francesco Marangon, Marika Gon, Maurizio Massaro e Andrea Moretti (a cura di), *Processi partecipativi nella progettazione turistica*, Udine, Forum, 2018, pp. 43-50

Uri: <http://forumeditrice.it/percorsi/scienza-e-tecnica/ambiente-territorio/processi-partecipativi-nella-progettazione-turistica/4-metodo-e-analisi-qualitativa-dei-risultati-del>

4. METODO E ANALISI QUALITATIVA DEI RISULTATI DEL FORUM TURISMO IN FRIULI VENEZIA GIULIA

Maurizio Massaro e Marika Gon

4.1. Introduzione

Elemento centrale di ogni ricerca scientifica è la definizione della metodologia e del metodo di riferimento (Yin, 2017). Nel caso specifico del Forum Turismo FVG, la ricerca vuole analizzare la prospettiva di molteplici operatori e stakeholder del settore, organizzati in tavoli di discussione, i cui pensieri sono organizzati in forma testuale. Partendo dagli obiettivi specifici del lavoro e dalle fonti informative disponibili, si ritiene che uno studio qualitativo basato sulla metodologia della content analysis (Stemler, 2001) sia particolarmente adatto. Secondo Krippendorff, la content analysis è una metodologia che cerca di rendere replicabili e affidabili i risultati di analisi di stringhe testuali (Krippendorff, 2013). L'utilizzo di questo approccio ha visto una progressiva diffusione in moltissimi ambiti di ricerca nelle scienze sociali (Dumay e Cai, 2015; Stepchenkova *et al.*, 2009), con il conseguente sviluppo e diffusione di metodi differenti di applicazione. Secondo Salvatore *et al.* (2012), una delle principali distinzioni all'interno della content analysis riguarda lo sviluppo di metodi computer-based e di metodi manuali. Mentre nei metodi manuali l'interpretazione dei contenuti avviene, per l'appunto, in forma non automatica con gli operatori che sintetizzano le informazioni contenute nei testi in modo manuale; negli approcci *computer-based*, l'operazione di estrazione di significato dal testo avviene con l'assistenza più o meno intensa di software specifici. Naturalmente, è possibile immaginare anche approcci misti con l'utilizzo sia di codifiche manuali che computer-based. Nel caso in esame, sono state utilizzate entrambi i metodi, con sintesi delle informazioni effettuate da operatori manualmente e mediante l'uso di supporti informatici. L'utilizzo di entrambi gli approcci ha consentito di ampliare la base di elaborazione a nostra disposizione, arricchendo i processi interpretativi dei risultati ottenuti. Di seguito si presentano pertanto entrambi gli approcci di ricerca utilizzati per processare ed analizzare le informazioni prodotte dai partecipanti al Forum FVG.

4.2. L'analisi *computer-based* delle informazioni

L'approccio metodologico computer-based è stato realizzato nello specifico dal team di ricerca DIES dell'Università degli studi di Udine. L'analisi è stata

Tab. 4.1. Fasi del metodo di analisi *computer-based*.

Fase	Obiettivo	Supporto informatico
1. Raccolta e pulizia errori	Analisi dei testi e pulizia di errori di digitazione	-
2. Analisi testuale con approcci <i>bag-of-words</i>	Estrarre i concetti più utilizzati	R statistics
3. Analisi testuale delle cooccorrenze	Aggregare i concetti in temi analizzare le relazioni e tra concetti e temi	Leximancer

Fonte: elaborazione propria.

realizzata seguendo un processo articolato in tre fasi tra loro interdipendenti e sintetizzati nella tabella 4.1 sotto riportata.

La prima fase del processo di analisi ha visto la raccolta delle discussioni avvenute nei tavoli di lavoro e sintetizzati in appunti presi da specifici operatori denominati ‘scribi’. In totale sono stati così raccolti i dati di 19 tavoli organizzati nelle quattro domande fondamentali (perché, come, cosa e chi) cui gli operatori erano chiamati a rispondere per ciascuno dei cinque temi di analisi (significato del turismo, sostenibilità, nuove imprenditorialità, nuove tecnologie e formazione life long learning). Il testo raccolto è stato sottoposto ad un processo di controllo al fine di correggere eventuali errori di digitazione. Nello specifico sono stati utilizzati i correttori di testo normalmente utilizzabili nei software di videoscrittura.

La seconda fase del processo ha visto lo sviluppo di approfondimenti supportati da software con approcci *bag-of-words*.

Questa tipologia di analisi ha l’obiettivo di estrarre i concetti utilizzati più di frequente nelle discussioni ricorrendo a specifici dizionari. Questo approccio non approfondisce le relazioni tra i concetti, né enfatizza gli aspetti grammaticali che normalmente caratterizzano il linguaggio umano.

La rappresentazione delle informazioni avviene attraverso forme grafiche quali le *wordclouds*.

La terza fase del processo di analisi ha visto invece lo sviluppo di approfondimenti supportati da software orientati all’estrazione di concetti, alla loro aggregazione in temi e al successivo approfondimento delle relazioni tra concetti e temi. Questo approccio di ricerca approfondisce la dinamica dei concetti analizzandone la vicinanza sintattica.

I due sotto-paragrafi seguenti offrono approfondimenti su ciascuno specifico metodo di ricerca.

4.2.1. L’analisi con approcci *bag-of-words*

La prima fase della ricerca prevede l’importazione dei commenti ottenuti nei singoli tavoli di discussione all’interno della piattaforma R (R Core Team,

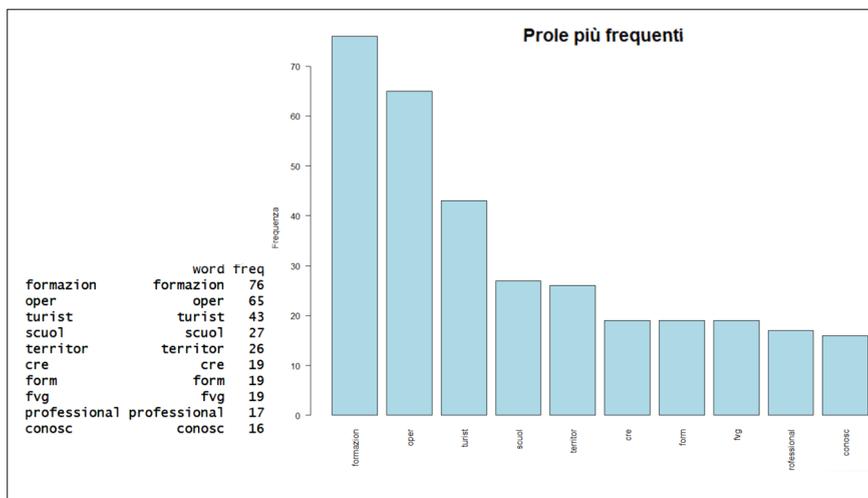


Fig. 4.1. Tabella per matrice di frequenza. Fonte: elaborazione propria.

2014). I dati, una volta importanti, hanno subito un processo di analisi caratterizzato dalle seguenti fasi:

- *Tokenisation*. Si tratta di una fase volta alla identificazione delle singole parole. Il testo viene suddiviso riconoscendo le singole parole che lo compongono definite, per l'appunto, *token*.
- *Lower case translation and stop word removal*. Si tratta di una fase volta alla identificazione e rimozione di pronomi (es: io, tu, egli, etc.), articoli (es. il, lo, un, uno, ecc.), preposizioni (es. per, con ecc.). Tutte le parole sono poi convertite in minuscolo, eliminando le maiuscole.
- *Stemming and lemmatisation*. Si tratta di una fase volta alla identificazione di parole con significati semantici simili (es. azienda, aziende, ecc.).
- Analisi dei termini ottenuti e loro aggregazione per significato simile. Si tratta di una fase manuale volta ad accorpare e/o eliminare parole dal significato simile.

Questa fase viene realizzata in modo semi assistito. In particolare, il software accorpa i sinonimi utilizzando specifici dizionari. L'operatore decide poi di effettuare ulteriori aggregazioni e/o eliminazioni non riconosciute dal software manualmente.

Una volta ottenuta la lista di parole, vengono costruite specifiche matrici di analisi (testo documento e/o documento testo), misurando la frequenza di ripetizione di ciascuna parola. Con gli approcci *bag-of-words*, infatti, ogni documento è rappresentato da un vettore di parole, in cui ognuna è associata in termini di presenza assenza nel caso binomiale o in termini di frequenza nel caso multinomiale. La figura 4.1 sotto riportata evidenzia un esempio di matrice di frequenza e un grafico rappresentante le parole a maggiore frequenza. I dati così ottenuti sono poi normalmente rappresentati attraverso una word-cloud. In questa tipologia di rappresentazione le parole a maggior frequenza sono rappresentate con caratteri più grandi, dando quindi una rappresentazio-

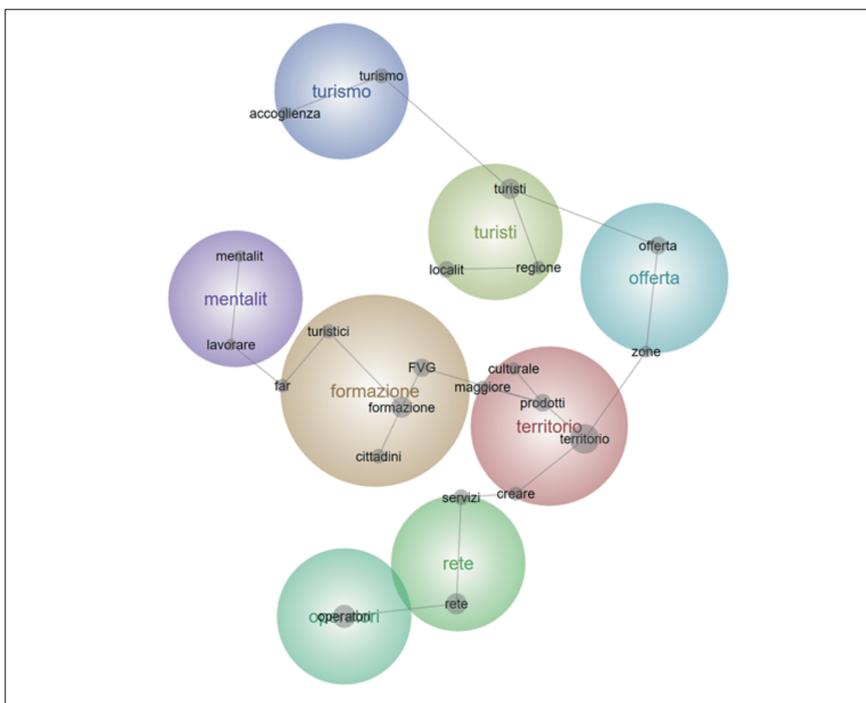


Fig. 4.3. Esempio di rappresentazione delle co-occorrenze nel testo. Fonte: elaborazione propria.

so lavoro manuale ad alto grado di soggettività». In questo modo viene quindi maggiormente garantita la replicabilità dei risultati.

Le fasi del processo di analisi prevedono:

- *concept seeds setting*. Si tratta della fase iniziale in cui il ricercatore setta i parametri della ricerca. Nello specifico vengono predefiniti il numero di frasi da utilizzare per individuare eventuali co-occorrenze nei concetti (normalmente due frasi) e i vocabolari di riferimento (italiano, inglese, ecc.);
- *compound settings*. Si tratta della fase più delicata in cui il ricercatore aggrega, elimina o ricodifica i concetti estratti dal software. In questo caso il ricercatore analizza i concetti proposti in automatico ed aggrega concetti che il software non riesce a riconoscere (es. Unione Europea e UE), elimina eventuali *stopwords* non identificate dal software (es. perché, potrebbe non essere riconosciuto per la presenza di diversi accenti sulla lettera ‘e’ finale).

Leximancer supporta sia l’analisi tematica che semantica (Tseng *et al.*, 2015). I risultati di questa fase di analisi sono molteplici. In primo luogo, viene rappresentato un diagramma di co-occorrenze in cui i vari concetti sono raggruppati in temi e connessi tra loro in una risultante mappa concettuale. La figura 4.3, sotto riportata, offre un esempio di questa rappresentazione. Le bolle rap-

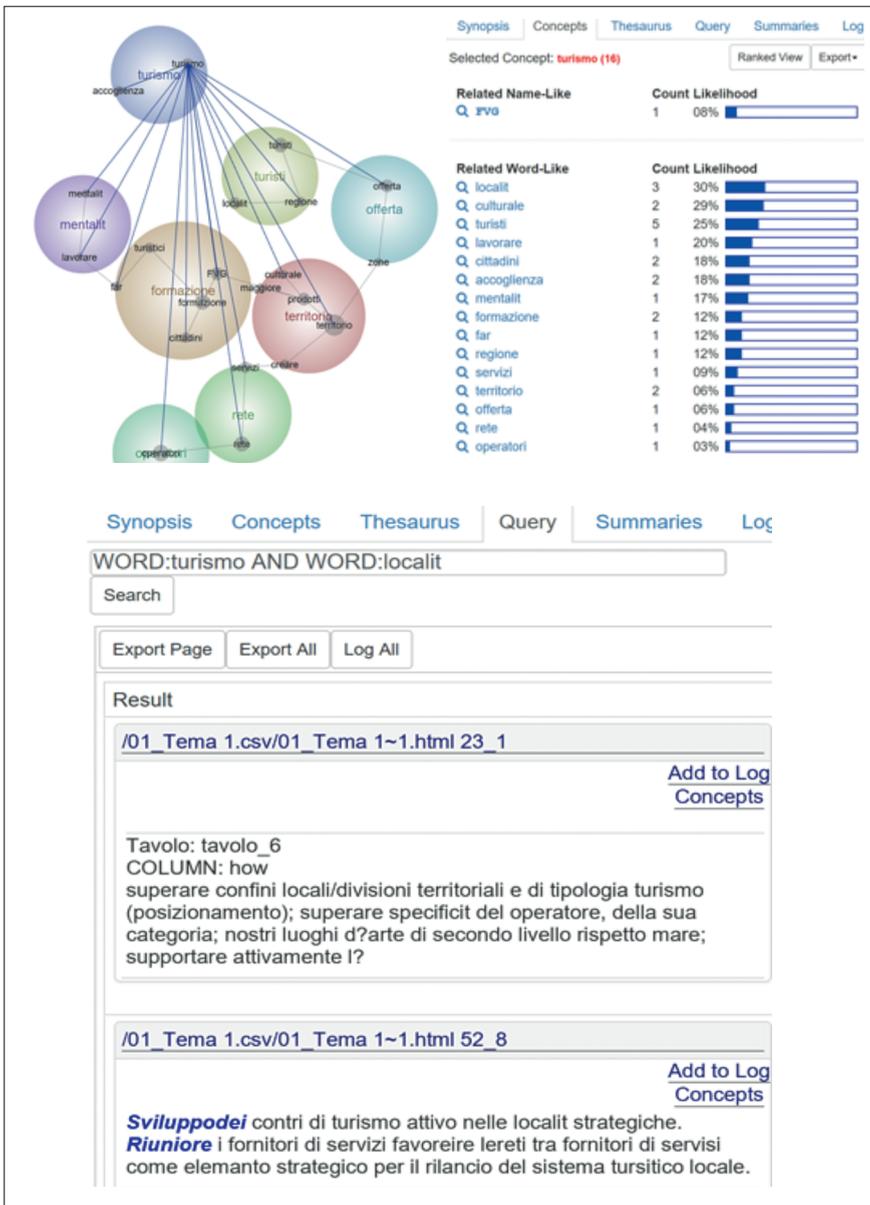


Fig. 4.4. Esempio di approfondimento delle co-occorrenze nel testo. Fonte: elaborazione propria.

presentano concetti, aggregati per co-occorrenza dei termini nelle singole unità di analisi (frasi), le linee di connessione rappresentano le relazioni che emergono tra le diverse tematiche (Sotiriadou *et al.*, 2014). Le bolle più grandi rappresentano tematiche chiave, mentre i punti raffigurano concetti puntuali.

La disposizione dei cerchi e dei punti corrisponde alla relativa vicinanza o distanza semantica tra i concetti: quando due bolle sono vicine, siamo in presenza di forti legami semantici (He e Chen, 2014; Tseng *et al.*, 2015).

Tale metodologia di analisi è stata di recente utilizzata in alcuni lavori scientifici (Sotiriadou *et al.*, 2014; Tseng *et al.*, 2015) che ne avallano l'alto livello di affidabilità e replicabilità.

L'elaborazione viene sviluppata attraverso un approfondimento dei singoli elementi di connessione. In questo modo il ricercatore può effettuare degli approfondimenti sui singoli testi per meglio comprendere i concetti espressi nel testo analizzato. La figura 4.4 offre un esempio di questa analisi.

4.3. Conclusioni

Il presente capitolo intende offrire una sintesi della metodologia di analisi utilizzata nel progetto di ricerca. Nello specifico è stata inquadrata la metodologia di riferimento, riconducendo l'analisi all'interno della *content analysis methodology*. Questa particolare tipologia di analisi si pone l'obiettivo di interpretare stringhe di testo rendendo replicabili e affidabili i risultati delle elaborazioni. Nello specifico all'interno della metodologia di analisi sono presenti molteplici metodi di applicazione.

Tra i molteplici approcci sono stati rappresentati i metodi noti come *bag-of-words* e analisi delle co-occorrenze utilizzati. L'utilizzo dei software R statistics e Leximancer hanno permesso una maggiore obiettività e sostenuto il processo di analisi della ricerca. I risultati emersi dall'analisi delle diverse prospettive dei partecipanti al Forum Turismo in Friuli Venezia Giulia sono descritti e presentati nel capitolo che segue.

Bibliografia

- Dumay J. (2014), *Reflections on interdisciplinary accounting research: The state of the art of intellectual capital*, in «Accounting, Auditing & Accountability Journal», 27 (8), pp. 1257-1264.
- Dumay J., Cai L. (2015), *Using content analysis as a research methodology for investigating intellectual capital disclosure: A critique*, in «Journal of Intellectual Capital», 16 (1), pp. 121-155.
- Krippendorff K. (2013), *Content Analysis. An Introduction to Its Methodology*, Sage Publications, Thousand Oaks, CA.
- He W., Chen Y. (2014), *Using blog mining as an analytical method to study the use of social media by small businesses*, in «Journal of Information Technology Case and Application Research», 16 (2), pp. 91-104.
- R Core Team (2014), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna.
- Salvatore S., Gennaro A., Auletta A.F., Tonti M., Nitti M. (2012), *Automated method of content analysis: A device for psychotherapy process research*, in «Psychotherapy research journal of the Society for Psychotherapy Research», 22 (3), pp. 37-41.

- Sotiriadou P., Brouwers J., Le T.A. (2014), *Choosing a qualitative data analysis tool: A comparison of NVivo and Leximancer*, in «Annals of Leisure Research», 17 (2), pp. 218-234.
- Stemler S. (2001), *An overview of content analysis*, in «Practical Assessment, Reserach & Evaluation», 7 (17), 137-146.
- Stepchenkova S., Kirilenko A.P., Morrison A.M. (2009), *Facilitating content analysis in tourism research*, in «Journal of Travel Research», 47 (4), pp. 454-469.
- Tseng C., Wu B., Morrison A.M., Zhang J., Chen Y.C. (2015), *Travel blogs on China as a destination image formation agent: A qualitative analysis using Leximancer*, in «Tourism Management», 46, pp. 347-358.
- Yin R.K. (2017), *Case study research and applications: Design and Methods*, Sage Publications, Thousand Oaks, CA.