

# Text mining e social media: Anatomia di una crisi di governo

Francesca Greco<sup>1</sup>, Andrea Fronzetti Colladon<sup>2</sup>, Alessandro Polli<sup>3</sup>

<sup>1</sup>Sapienza University of Rome, Rome, Italy – francesca.greco@uniroma1.it

<sup>2</sup>University of Perugia, Department of Engineering, Perugia, Italy – andrea.fronzeticolladon@unipg.it

<sup>3</sup>Sapienza University of Rome, Rome, Italy – alessandro.polli@uniroma1.it

## Abstract

In August 2019, when Matteo Salvini asked to return to the ballot box, the Italian government entered into crisis. This generated a large debate that also took place on social media – particularly on Twitter, which has now become increasingly important as a place for political debate. This study presents the analysis of the Twitter discourse related to government crisis, in the period following the resignation of the Prime Minister. Text mining and social network analysis techniques have been integrated to study the positioning of political leaders and the topics of major interest. In particular, we used the Emotional Text Mining (ETM) and Semantic Brand Score (SBS) techniques. The SBS served as an indicator of the most relevant topics in the discourse. It was used to determine the importance of the themes emerging from the EMT. This importance was measured along the three dimensions of prevalence, diversity, and connectivity – considering textual association patterns and co-occurrence networks. The integration of the two procedures allowed to describe the public perception of the crisis and to identify the symbolic-cultural matrix and the sentiment related to each discourse topic.

**Keywords:** social media, SBS, ETM, social profiling

## Abstract

Nel mese di agosto del 2019, il governo italiano è entrato in crisi a fronte della richiesta di tornare alle urne del leader della *Lega*, scelta questa che ha promosso un intenso dibattito sulle ragioni e sulle possibili soluzioni della crisi. Una gran parte di tale dibattito si è svolta sui social media, che attualmente hanno assunto una crescente rilevanza quale luogo del dibattito politico. Questo studio presenta l'analisi delle comunicazioni sui social media relative alla crisi di governo, nel periodo successivo alla presentazione delle dimissioni da parte del Presidente del Consiglio. A questo scopo, sono state integrate tecniche di text mining e social network analysis, per studiare i comportamenti di voto e il posizionamento dei leader politici. In particolare, si sono utilizzati l'Emotional Text Mining (ETM) e il Semantic Brand Score (SBS). Il SBS è servito quale indicatore di importanza dei temi emersi dall'EMT. Dall'analisi delle associazioni, ovvero delle co-occorrenze testuali, il SBS estrae informazioni utili a valutare le dimensioni di prevalenza, diversità e connettività di ciascun tema all'interno del discorso. L'integrazione delle due procedure ha consentito di esplorare la percezione pubblica della crisi, attraverso l'identificazione della matrice simbolico-culturale, delle rappresentazioni dei cittadini, del sentiment relativo ai leader politici e degli elementi di priorità all'interno del discorso.

**Parole chiave:** social media, SBS, ETM, social profiling

## 1. Introduzione

In questo studio presentiamo l'analisi della percezione pubblica della recente crisi di governo a partire dalle comunicazioni prodotte sui social media. Come sottolinea Garland (2001),

l'attenzione per una specifica tematica, che caratterizza a fasi alterne l'opinione pubblica, non è il prodotto diretto della comunicazione mediatica o della retorica politica, benché entrambi questi fattori abbiano un ruolo rilevante rispetto alla sua declinazione. Infatti, la sensibilità al tema è inevitabilmente connessa con l'esperienza collettiva di vicende concrete che accadono nella vita quotidiana e dagli adattamenti pratici che ne derivano. Tuttavia, la diffusione delle informazioni e il posizionamento assunto dalle istituzioni politiche influiscono sul modo attraverso il quale gli eventi sono collettivamente rappresentati.

In Italia ad agosto del 2019, dopo poco più di un anno dalle elezioni politiche del 2018, il governo è entrato in crisi, apparentemente a fronte del disaccordo inerente impegni presi durante la campagna elettorale tra i due partiti di governo. La crisi di governo ha sollecitato un forte dibattito sulle ragioni e sulle possibili soluzioni della stessa. Una gran parte di tale dibattito si è svolta sui social media, che attualmente assumono una crescente rilevanza quale luogo del dibattito politico. In particolare, i social sono diventati veicolo della libera espressione delle opinioni e dei sentimenti delle persone, compresi quelli *politically incorrect*.

Come evidenziato dalla letteratura, l'ampia diffusione di Internet ha aumentato la possibilità per milioni di persone di navigare sul web, di creare account e di cercare o condividere informazioni quotidianamente. Il costante aumento del numero di utenti nelle piattaforme dei social media rende disponibili una grande quantità di dati, che rappresentano una fonte importante per esplorare le rappresentazioni, le opinioni e i sentimenti delle persone. I social media ed i siti di social network, come Facebook e Twitter, hanno iniziato ad assumere un ruolo crescente nella politica del mondo reale (Castells, 2007; Cepernich et Novelli, 2018; Ceron et al., 2013); ciò ha portato allo sviluppo, negli ultimi anni, di un numero crescente di tecniche di analisi, che sono diventate strumenti utili, veloci ed economici, rispetto alle rilevazioni tradizionali per l'esplorazione dei fenomeni sociali. Come l'analisi delle manifestazioni e azioni di rivolta, della partecipazione e istituzione di movimenti sociali e partiti politici, di elezioni, dell'attività del mercato azionario, ecc. (e.g. Ceron et al., 2013; Greco et al., 2017; Hopkins & King, 2010 ; Gloor, 2017; Schoen et al., 2013). In tal senso, il presente studio esplora i discorsi pubblicati su Twitter attorno al tema della scelta del leader della *Lega* di uscire dal governo, per tornare al voto nelle fasi successive alla presentazione delle dimissioni del Presidente del Consiglio.

Nello specifico, per studiare la percezione della crisi di governo, sono state utilizzate due tecniche di text mining, l'Emotional Text Mining (ETM) (Greco, 2016; Greco & Polli, 2019a) e il Semantic Brand Score (SBS) (Fronzetti Colladon, 2018), che vengono usate nell'ambito delle scienze sociali per studiare e prevedere i comportamenti di voto e il posizionamento dei leader politici. Esse consentono di esplorare la percezione pubblica della crisi attraverso l'identificazione della matrice simbolico-culturale, delle rappresentazioni e del sentiment dei cittadini, individuando candidati e argomenti prioritari.

In particolare, sui risultati ottenuti con l'ETM in uno studio precedente (Greco et Polli, 2019c) è stato calcolato il SBS, allo scopo di individuare un indicatore di importanza dei temi emersi dalla prima analisi. In questo senso, la procedura ha integrato tecniche di text mining e di social network analysis, che si basano sulla trasformazione dei testi in reti di parole. Dall'analisi delle associazioni e delle co-occorrenze testuali, l'SBS estrae informazioni utili a valutare gli elementi di prevalenza, diversità e connettività di ciascun tema all'interno del discorso.

## 2. Anatomia di una crisi di governo: il caso italiano

### 2.1. *Analisi di contesto*

Nel linguaggio comune, il concetto di "istituzione" si riferisce a un apparato che svolge funzioni e compiti legati all'interesse pubblico, in senso lato, mentre nelle scienze sociali l'istituzione è un modello di comportamento che ha una forza vincolante in una data società.

Tuttavia, la forza vincolante non è l'unico requisito che identifica un'istituzione. Limitando il campo alle istituzioni politiche, che svolgono un ruolo centrale nella ricerca del benessere collettivo, è essenziale considerare, accanto alla funzione cogente, quella relazionale, indicata da Poulantzas come un "campo strategico" di confronto e strutturazione delle relazioni di potere, in cui si stabilisce una dialettica tra "società civile" e "società politica" (Poulantzas, 1973).

La ricerca del benessere collettivo da parte di un'istituzione sembra essere indissolubilmente legata alla sua capacità di soddisfare le esigenze dei cittadini, come la sicurezza, che è una delle principali determinanti del benessere collettivo. Più specificamente, il perseguimento dei due obiettivi rappresenta la principale fonte di legittimazione delle istituzioni politiche, che costruiscono il loro consenso sulla presunta capacità di garantire entrambi. Il legame tra istituzioni politiche, benessere e consenso è profondamente cambiato a seguito della crescente interdipendenza globale innescata dai processi di globalizzazione. Più specificamente, il trasferimento di potere dagli Stati ai mercati (2) ha portato all'eradicazione delle istituzioni politiche dalla loro base territoriale (Giordano, 2016). Pertanto, sia la rappresentazione del benessere collettivo sia le misure più idonee per raggiungerlo sono cambiate. Per quanto riguarda la sicurezza, nel recente passato essa includeva il concetto di sicurezza sociale garantito da un sistema di welfare, mentre oggi sembra essere sostanzialmente limitata agli aspetti geografico-strutturali (Coluccia et al., 2008). La conseguente crisi della rappresentanza ha causato l'ascesa di movimenti con forti connotazioni identitarie, che si concentrano sul superamento delle attuali forme di democrazia rappresentativa e su questioni di sicurezza, fecendone elementi centrali della loro agenda politica.

### 2.2. *Emotional Text Mining*

L'uso delle tecniche di text mining per misurare il sentimento di un testo è stato ampiamente discusso in letteratura (e.g., Liu, 2012; Ceron et al., 2013; Gloor, 2017; Greco e Polli, 2019a; Hopkins et King, 2010). Tra le varie procedure utilizzate per misurare il sentiment l'Emotional Text Mining (ETM) (Greco, 2016; Greco et Polli 2019a) è una metodologia di classificazione di grandi quantità di dati non strutturati, che utilizza un approccio semiotico e una procedura non supervisionata di tipo bottom-up; vale a dire che non necessita di dizionari di riferimento, o di un learning set, per la classificazione dei testi. Essa consente di misurare il sentiment a partire dall'identificazione delle rappresentazioni collettive e della relativa matrice simbolico-culturale che le sostanzia sulla base della co-occorrenza delle parole che caratterizzano la comunicazione. L'ETM è una procedura automatica, veloce e relativamente semplice, che può essere utilizzata per estrarre informazioni in merito a quelle che potremmo definire le opinioni emotivamente guidate.

L'EMT si colloca entro una prospettiva di matrice socio-costruttivista a orientamento psicodinamico, che interpreta il comportamento degli attori sociali come il prodotto dei

processi di categorizzazione conscia e di simbolizzazione inconscia della realtà, processi che sono socialmente costruiti e condivisi. Gli attori sociali interpretano e condividono la conoscenza del mondo che li circonda attraverso le rappresentazioni sociali, che generano e, allo stesso tempo, sono generate dalle interazioni. Esse non sono solo condivise ma ricostruite, ri-prodotte, ri-rappresentate interattivamente dalla collettività (Moscovici, 2005). Alla base della genesi delle rappresentazioni sociali vi sono dei processi inconsci di simbolizzazione emozionale che organizzano le relazioni, i comportamenti e le narrazioni degli attori sociali entro i contesti (Carli, 1990). Gli attori categorizzano consciamente la realtà e, contemporaneamente, la simbolizzano affettivamente a livello inconscio (Fornari, 1976). Questi due processi sono il prodotto di un funzionamento bi-logico della mente (Matte Blanco, 1981) che consente l'adattamento dell'uomo al suo contesto sociale. In questa prospettiva, i processi inconsci sono sociali, poiché gli attori generano interattivamente e condividono le stesse simbolizzazioni affettive del contesto ed esse organizzano il loro comportamento così come i loro processi di comunicazione (Salvatore et Freda, 2011).

Alla luce di tale prospettiva, è possibile utilizzare l'analisi del testo, guardando alla comunicazione come un fenomeno sociale che, alla stregua dei comportamenti, sono il prodotto della mente e, dunque, sono riconducibili sia al suo modo di essere cosciente, che al suo modo di essere inconscio. Da ciò consegue che la comunicazione può essere studiata sia a partire dai suoi elementi semantici e razionali, sia attraverso l'analisi della dimensione simbolica che genera e istituisce il valore *emozionale* della narrazione. Se il funzionamento conscio organizza il *cosa* viene narrato, vale a dire i contenuti manifesti della narrazione, il funzionamento inconscio può essere inferito attraverso il *come* viene narrato, vale a dire attraverso la scelta delle parole e la loro associazione nel testo. Le parole sono il prodotto della simbolizzazione emozionale che il narrante ha dell'oggetto narrato. In tal senso, attraverso una procedura di analisi statistica multivariata, è possibile considerare i nessi associativi fra le parole e formulare delle interpretazioni circa l'ipotetica matrice simbolica che determina la coesistenza di termini in relazione a un oggetto e in funzione del contesto. Pertanto, l'ETM consente di effettuare una profilazione degli attori sociali a partire dalle loro comunicazioni. Essa è stata utilizzata in diversi ambiti (e.g., Greco, 2016; Cordella et al., 2018; Greco et al., 2018; Greco, 2019) tra cui lo studio del dibattito politico in relazione alle scelte di voto (Greco et al., 2017; Greco et Polli, 2019a; 2019b; 2019c).

L'interpretazione dei risultati delle analisi dei cluster consente di identificare gli elementi che caratterizzano la rappresentazione emozionale, mentre i risultati dell'analisi delle corrispondenze consentono di individuare le dimensioni latenti costitutive della matrice simbolica che caratterizza la cultura. Se il processo statistico procede per successivi livelli di strutturazione e di sintesi, dalla selezione delle parole alla identificazione delle dimensioni latenti caratterizzanti lo spazio fattoriale, il processo di interpretazione procede in modo inverso, partendo dalla definizione dell'ipotetica matrice simbolica sottostante la comunicazione, per identificare le rappresentazioni. Il vantaggio di questa procedura risiede nella possibilità di misurare il sentiment in funzione del numero di testi classificati e caratterizzanti la rappresentazione sociale. A differenza delle tecniche supervisionate, si ha il vantaggio di esplorare il posizionamento (positivo, negativo o neutro) di chi comunica, comprendendone altresì le motivazioni e le rappresentazioni.

### **2.3. ETM di una crisi di governo**

In un precedente studio (Greco et Polli, 2019c) è stata analizzata la percezione pubblica relativa alla decisione del leader della *Lega* di uscire dal governo con la metodologia

dell'Emotional Text Mining (ETM) (Greco et Polli, 2019). A questo scopo, erano stati raccolti tutti i messaggi in italiano di Twitter che contenevano il nome del leader della *Lega* per sei giorni nel periodo critico della crisi, vale a dire quando il Presidente del Consiglio aveva rassegnato le proprie dimissioni e non era ancora stato definito un nuovo equilibrio politico. Da Twitter, erano stati raccolti tutti i messaggi in lingua italiana che contenevano la parola "Salvini" dal 20 al 25 agosto 2019. La raccolta dati era stata effettuata con il pacchetto *twitteR* del software R statistics (v.1.1.9, Gentry, 2016).

A causa dell'elevato numero di messaggi prodotti nell'arco di 6 giorni, e delle limitazioni imposte da uno dei software utilizzati per effettuare l'ETM (T-Lab Plus, Lancia, 2018), erano stati selezionati i primi 99.000 tweet (il 77,6% dei quali erano retweet), ottenendo un corpus di grandi dimensioni (token = 1.846.011). Allo scopo di studiare le caratteristiche lessicali del corpus in analisi e di valutare la possibilità di effettuare un'analisi statistica dei dati testuali, erano stati calcolati il rapporto del numero delle forme grafiche sul totale delle occorrenze (TTR) e la percentuale delle forme grafiche che occorrono solo una volta (H) (TTR = 0,024; H = 57,9%).

Dopo una fase di pre-processamento dei dati, in cui sono state rimosse le stop word ed è stata effettuata una lemmatizzazione, sono stati selezionati tutti i lemmi, escludendo dall'analisi le parole utilizzate per selezionare i testi e quelle del rango basso di frequenza (Greco et al., 2017; Greco and Polli, 2019a). Successivamente, sulla matrice lemmi selezionati per tweet è stata effettuata un'analisi dei cluster con un algoritmo bisecting *k*- means, limitato a 20 partizioni, con l'esclusione dei testi che non presentavano almeno una co-occorrenza, ed è stato utilizzato il coseno come misura di similarità (Savaresi et Boley, 2004). Al fine di individuare la partizione ottimale, sono stati utilizzati tre indici: il Calinski-Harabasz, il Davies-Bouldin e il coefficiente di correlazione intraclasse ( $\rho$ ). Infine, sulla matrice parole chiave-cluster è stata effettuata una analisi delle corrispondenze (Lebart et Salem, 1994) e si è misurato il sentiment in base al numero di messaggi classificati nei cluster e alla loro interpretazione.

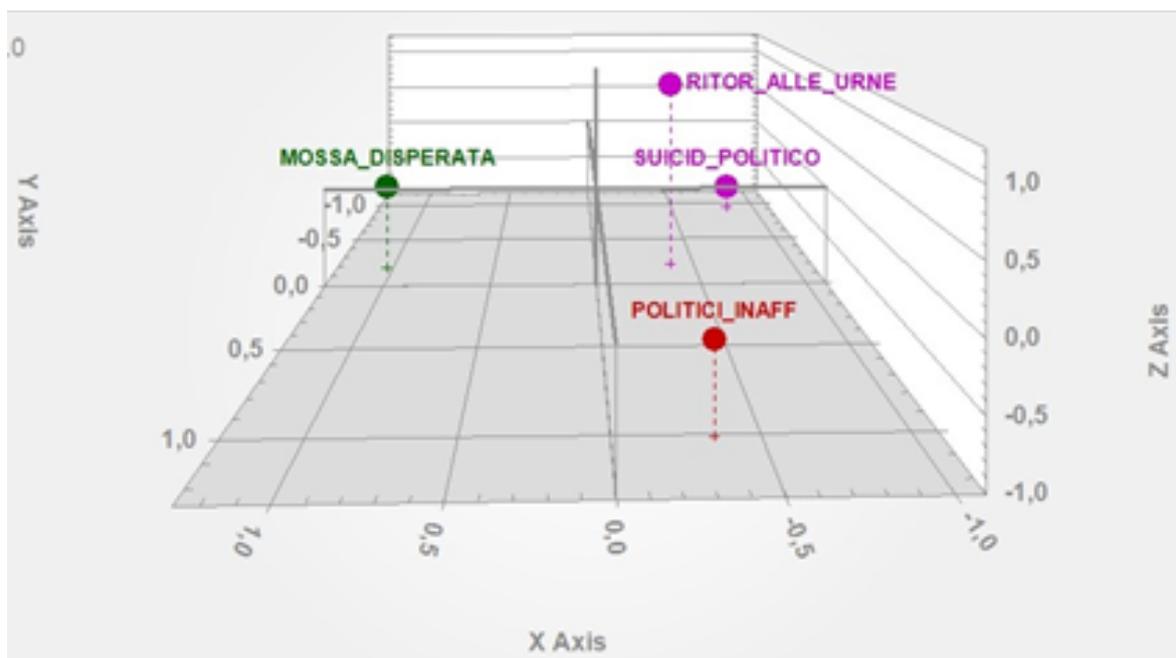
#### 2.4. La mappa emozionale della crisi

Dall'analisi dell'ETM sono emersi quattro rappresentazioni della scelta del leader politico di uscire dalla coalizione di governo (Greco et Polli, 2019c). L'utilizzo dello spazio fattoriale consente di rappresentare uno spazio simbolico entro il quale le quattro rappresentazioni assumono senso (Figura 1). Come evidenziato in Tabella 1, gli utenti di Twitter simboleggiano il leader della Lega per mezzo di tre dimensioni principali: il leader, il motivo della sua scelta e le conseguenze.

Tabella 1 Sintesi dell'interpretazione dell'ETM

<i>Cl. Etichetta</i>	<i>Tweet</i>	<i>Tweet %</i>	<i>Fattore 1</i> (39.6% inerzia) <i>Leader</i>	<i>Fattore 2</i> (33.2 % inerzia) <i>Motivazione</i>	<i>Fattore 3</i> (27.2% inerzia) <i>Conseguenze</i>
<b>1</b> Ritorno alle Urne	23.623	27,1	Consapevole		Votare
<b>2</b> Politici Inaffidabili	22.213	25,4	Consapevole	Incompetenza	Dibattere
<b>3</b> Mossa Disperata	25.071	28,8	Impulsivo		
<b>4</b> Suicidio Politico	16.264	18,7	Consapevole	Intenzionalità	Dibattere

Il primo fattore riguarda la categorizzazione del leader della Lega, simbolizzata come una persona riflessiva, su cui si può fare affidamento, o una persona impulsiva, potenzialmente inaffidabile. Il secondo fattore riflette il motivo della sua scelta, che viene simbolizzata come intenzionalmente dannosa, vale a dire dettata dall'intento di trarre vantaggio a scapito delle conseguenze, o inesperta, vale a dire come il prodotto di una valutazione inaccurata. Infine, il terzo fattore simboleggia la reazione dei cittadini: andare a votare, che rimanda ad una decisione individuale degli stessi, o discutere del problema e delle possibili soluzioni, che rimanda all'intervento esperto dei rappresentanti politici ed, eventualmente, ad una successiva



decisione collettiva.

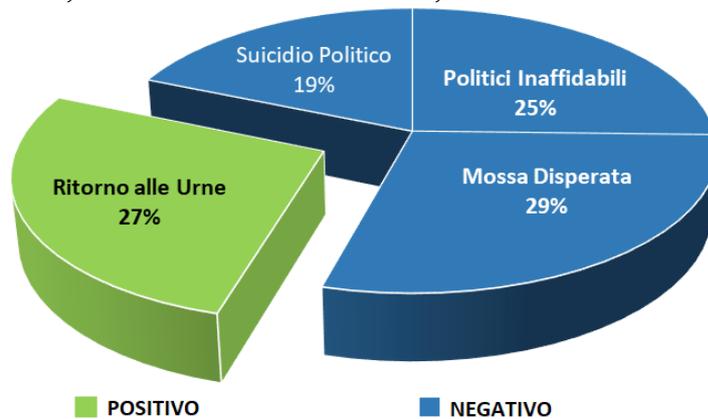
Figura 1: Lo spazio simbolico

I quattro cluster sono di dimensioni diverse (Tabella 1) e riflettono diverse rappresentazioni della decisione del leader della Lega. Nel primo cluster, gli utenti di Twitter sembrano concordare con la proposta del leader di ritornare alle urne, percependolo come consapevole e in grado di prendere decisioni politiche responsabili; il secondo cluster lo rappresenta come l'espressione dell'inaffidabilità della classe politica in generale, insensibile all'interesse generale; il terzo cluster percepisce la scelta come una mossa disperata, dettata da una decisione impulsiva; infine, il quarto cluster, che è il più piccolo (19%), giudica la decisione del leader come un suicidio politico, che necessita dell'intervento di esponenti politici capaci di individuare una soluzione per ridurre l'impatto della scelta del leader della *Lega*.

Per quanto concerne il sentiment (Figura 2), sono stati considerati negativi (73%) i temi dei *Politici Inaffidabili*, della *Mossa Disperata* e del *Suicidio Politico*, e positivo (27%) quello del *Ritorno alle Urne*. È interessante notare come due delle tre rappresentazioni negative si caratterizzano per una valutazione del leader della Lega (*Mossa Disperata* e *Suicidio Politico*), mentre l'altra (*Politici Inaffidabili*) generalizza il giudizio estendendolo all'intera classe politica.

### 3. Semantic Brand Score

I risultati ottenuti con la metodologia dell'ETM, precedentemente descritta, sono stati sottoposti ad analisi col Semantic Brand Score. La misura del SBS è stata applicata ai set di parole, ovvero ai cluster lessicali, identificati attraverso la procedura dell'Emotional Text



Mining (Greco et Polli, 2019). Di questi cluster, espressione di specifici temi e concetti, sono stati valutati l'importanza e il posizionamento, anche in termini di sentiment. In aggiunta, sono state indagate le similarità tra cluster, considerandone le associazioni testuali. Una parte importante di questo lavoro riguarda appunto lo studio della complementarità tra i due approcci, l'ETM e il SBS.

Figura 2 – Analisi del sentiment

Il Semantic Brand Score (SBS) è un nuovo e innovativo indicatore dell'importanza dei brand (A Fronzetti Colladon, 2018), progettato per adattarsi ai contesti più vari, sia culturali, sia linguistici. Si tratta di una metrica adatta per l'analisi di (grandi) dati testuali, che ha le sue radici nella Social Network Analysis (Wasserman et Faust, 1994). La concettualizzazione del SBS è stata in parte ispirata da noti modelli di brand equity e dai concetti di brand image e brand awareness (Keller, 1993).

Uno dei vantaggi, rispetto alle misure tradizionali, è che il calcolo del SBS non richiede questionari. Quest'ultimi sono solitamente soggetti a numerosi bias (Grandcolas, Rettie, et Marusenko, 2003; Olson, 2006; Warner, 1965), risentono delle dimensioni del campione di intervistati e raccolgono risposte che diventano meno spontanee dal momento in cui ci si sottopone ad intervista. Al contrario, un insieme di testi può rappresentare le espressioni di un'intera popolazione in senso statistico: per esempio, tutti gli articoli che menzionano Greta Thunberg e le questioni ecologiche. Il SBS può essere calcolato su qualsiasi collezione di testi, inclusi quelli raccolti da e-mail, tweet e post sui social media. L'obiettivo è quello di cogliere le espressioni delle persone (giornalisti, consumatori, amministratori delegati, politici, cittadini, ecc....) nei luoghi in cui queste appaiono normalmente.

Il concetto di "brand" è molto flessibile e il SBS può essere calcolato per qualsiasi parola, o insieme di parole, in un corpus. Per 'brand' si può intendere il nome di un politico, o un insieme di parole chiave che rappresentano un concetto (ad esempio, il concetto di "innovazione" o un core value aziendale). La misura è stata già utilizzata per: valutare le dinamiche di transizione che si verificano quando un marchio nuovo sostituisce uno vecchio (A Fronzetti Colladon, 2018); valutare il posizionamento di aziende concorrenti; prevedere i risultati delle elezioni politiche (Fronzetti Colladon, 2020); stimare i trend di visitatori nei musei, analizzando i forum online su cui dialogano i turisti (Fronzetti Colladon, Grippa, et Innarella, 2020).

Il SBS ha tre dimensioni: prevalenza, diversità e connettività. La prevalenza misura la frequenza d'uso del nome di un marchio, cioè il numero di volte in cui un brand viene menzionato direttamente. Può essere considerata una misura che cattura, almeno parzialmente, la brand awareness, intesa come il grado di conoscenza del brand da parte del pubblico. La diversità misura l'eterogeneità delle parole associate a un marchio, cioè la ricchezza della suo intorno lessicale. In ultimo, la connettività rappresenta la capacità di un brand di collegare concetti, parole e temi differenti all'interno di un discorso. La somma di questi tre indicatori costituisce il SBS, una metrica estensivamente descritta nel lavoro di Fronzetti Colladon (2018).

A livello tecnico, gli step di calcolo del SBS prevedono che vi sia un pre-processamento dei testi da analizzare, che include operazioni quali la trasformazione delle maiuscole in minuscole, la rimozione di caratteri speciali e link web, la rimozione delle stop-word, la tokenizzazione dei testi e la rimozione dei suffissi delle parole, nota come stemming (Jivani, 2011). A valle del pre-processamento, i testi vengono trasformati in reti di parole, ovvero grafi in cui ogni parola è rappresentata da un nodo e i link tra nodi sono dati dalle co-occorrenze lessicali che si verificano all'interno di un range di 7 parole di distanza dal brand (Fronzetti Colladon, 2018). Questo livello soglia può essere aggiustato in funzione dell'analisi. Le reti vengono altresì filtrate al fine di rimuovere le co-occorrenze trascurabili, ovvero i link di minor peso. La generazione di queste reti consente il successivo calcolo delle misure di diversità e connettività, calcolate rispettivamente attraverso le metriche di degree centrality e betweenness centrality pesata (Brandes, 2001; Freeman, 1979). La prevalenza, invece, è data dal conteggio delle frequenze delle parole. Ogni misura viene poi standardizzata, considerando i valori di tutti i nodi nella rete, sottraendo la media ai singoli punteggi e dividendo per la deviazione standard. I punteggi standardizzati vengono sommati per calcolare il SBS. Le analisi vengono condotte utilizzando la SBS BI web app, disponibile sul sito internet <https://bi.semanticbrandscore.com> (Fronzetti Colladon et Grippa, 2020).

## 4. Risultati

Dai risultati del SBS emerge come il cluster *Ritorno alle Urne* mostri un trend di importanza crescente all'interno del discorso (Figura 3). Più variabile, invece, l'andamento del cluster *Politici Inaffidabili*, che ha picchi e ricadute a seconda dei giorni (picchi il 20 e il 24 agosto). Meno importanti le rappresentazioni della scelta del Leader come *Mossa Disperata* o come *Suicidio Politico*.

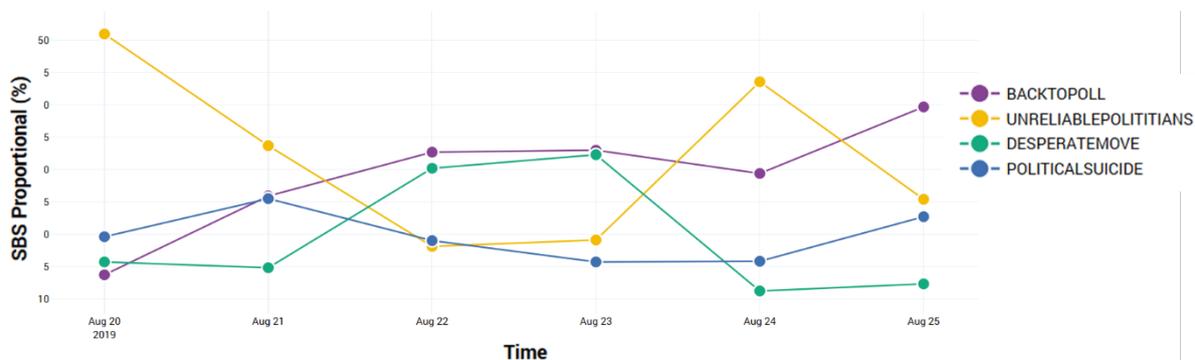


Figura 3 - Valori proporzionali del SBS

In generale, considerando l'intero periodo, vediamo come il cluster *Politici Inaffidabili* sia il tema di maggiore rilievo, seguito da *Ritorno alle Urne* e *Suicidio Politico* (Figura 4). I primi

due sono le rappresentazioni che maggiormente tengono unito il discorso, agendo da catalizzatori di argomenti differenti. Benché il tema della *Mossa Disperata* venga menzionato più spesso di quello dei *Politici Inaffidabili*, il discorso intorno a questo tema è più piatto, ovvero meno diversificato, e con scarso potere connettivo, inteso come la capacità di fare da ponte tra i diversi argomenti del discorso complessivo.

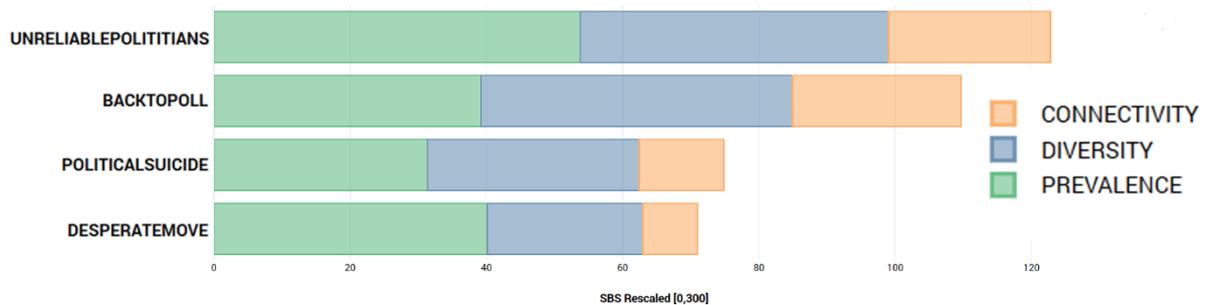


Figura 4 - SBS medio, uno spaccato delle tre dimensioni

In termini di sentiment (Figura 5), calcolato utilizzando la libreria *pattern* scritta per il linguaggio Python (Malik, 2019), vediamo che quasi tutti gli argomenti appaiono vicini al neutro, con l'alternarsi di positività e negatività che quasi si annullano in media. In questo approccio (diverso dall'ETM) il calcolo del sentiment è automatizzato e non 'interpretato'; si limita inoltre all'analisi delle porzioni di testo che includono le parole chiave caratterizzanti ogni tema. La *Mossa Disperata* è il tema con il sentiment maggiormente negativo. Il più positivo, invece, è quello dei *Politici inaffidabili*. Probabilmente questa differenza è connessa con la generalizzazione dell'inaffidabilità di tutta la classe politica, a differenza di quanto espresso nello specifico, circa le azioni di Salvini.

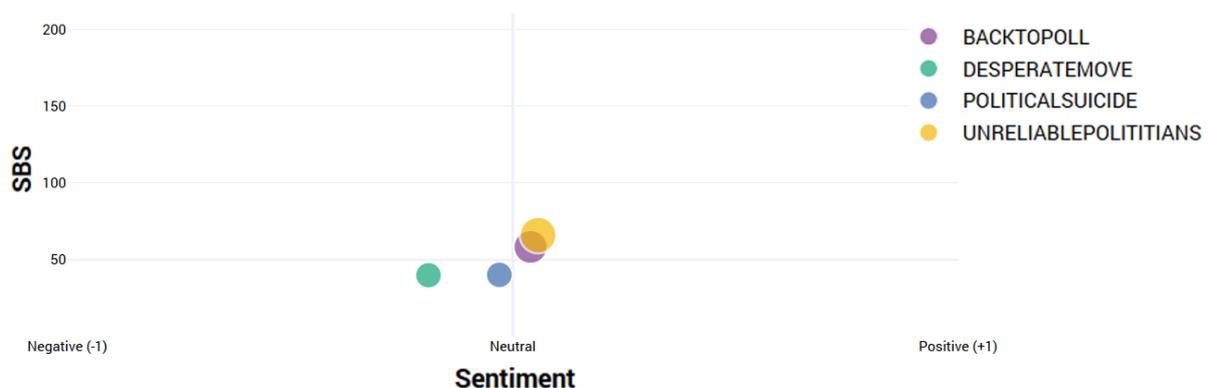


Figura 5 - Posizionamento e sentiment dei cluster

In ogni caso, si può notare come gli algoritmi più comuni per il calcolo del sentiment, o almeno quelli qui utilizzati, si rivelano meno esplicativi rispetto alla visione e alle informazioni che emergono dall'ETM.

#### 4. Discussione e conclusioni

Questo studio integra tecniche di text mining (ETM) e social network analysis (SBS) allo scopo di misurare il sentiment espresso dagli utenti di Twitter a fronte della crisi di governo e di comprendere lo sviluppo dinamico dei temi del discorso nel periodo immediatamente successivo alla presentazione delle dimissioni da parte del Presidente del Consiglio.

Dal punto di vista metodologico, l'integrazione fra SBS e ETM ha consentito di esplorare la percezione pubblica della crisi, attraverso l'identificazione della matrice simbolico-culturale, delle rappresentazioni dei cittadini, del sentiment relativo al leader politico e degli elementi di priorità all'interno del discorso. Il risultato appare interessante, in quanto l'applicazione combinata delle due procedure consente di analizzare il fenomeno da diverse prospettive. Infatti, l'ETM sembra cogliere meglio la dimensione del sentiment, individuandone le determinanti più significative. Questa tecnica viene utilizzata per una preliminare individuazione dei temi caratterizzanti il discorso e per comprendere l'orientamento degli utenti di Twitter circa questi temi. Il SBS, e l'app SBS BI (Fronzetti Colladon et Grippa, 2020), tornano utili in un secondo step, per un'analisi più approfondita delle dinamiche di interesse e per capire, nel tempo, come evolve l'importanza dei diversi temi, studiandone anche l'immagine e il posizionamento.

Dall'analisi dei messaggi pubblicati emerge come la scelta del leader della *Lega* di uscire dal governo suscita sentimenti prevalentemente negativi (73%). Tale risultato non è completamente inatteso. La principale fonte di legittimazione di un'istituzione politica è indissolubilmente legata alla capacità dell'istituzione stessa di soddisfare i bisogni e le aspettative dei cittadini, contribuendo al contempo a migliorare la percezione pubblica e individuale della sicurezza. La sicurezza è un concetto multidimensionale, che non si esaurisce nella mera incolumità fisica, ma al contrario si declina in molteplici ambiti. Non a caso la forte ascesa della *Lega* nei risultati elettorali è motivata anche dalla proposta di un programma politico in cui la sicurezza – intesa non soltanto nella sua dimensione geografico-strutturale, ma anche economica – riveste un'importanza centrale.

La decisione di Salvini è giudicata sostanzialmente avventata nel 47,5% dei messaggi analizzati (*Mossa Disperata* e *Suicidio Politico*) e dimostra la generale inaffidabilità dei leader politici per un ulteriore 25,4% (*Politici Inaffidabili*). Tale decisione avviene in un contesto in cui, per motivi sostanzialmente noti (la globalizzazione, le incertezze dello scenario geopolitico, il dissolvimento delle reti di protezione del Welfare State, ecc.), l'esistenza dei cittadini è caratterizzata da crescenti elementi di precarietà. Parallelamente, la responsabilità di aver provocato una crisi di governo – e quindi di aver messo a repentaglio quel particolare aspetto della sicurezza politica che è la stabilità delle istituzioni – è stata addebitata al leader della *Lega*. Di qui, verosimilmente, il sentimento prevalentemente negativo espresso nei discorsi condivisi sui social media.

Al di là delle caratterizzazioni emozionali, il tema dell'inaffidabilità dei leader politici è quello che gli utenti di Twitter hanno più a cuore. Si tratta del tema più dibattuto, anche in modo più diversificato, e con il Semantic Brand Score più alto. Al secondo posto troviamo il discorso del ritorno alle urne, che pure assume un peso rilevante e mostra un forte potere connettivo tra i vari argomenti. In effetti, sono questi i due temi che hanno maggiore impatto sul futuro della nazione e dei cittadini. Gli altri due argomenti che giudicano le scelte dei vari politici (*Mossa Disperata* e *Suicidio Politico*), hanno minore rilevanza, per quanto le espressioni di sentiment possano essere più accentuate.

Quanto al restante 27,1% di Tweet classificati come positivi, tale sentimento non è necessariamente espresso da simpatizzanti per la proposta politica della *Lega* e, nello specifico, quella del suo leader. In tali opinioni si rispecchia un'ampia gamma di posizioni individuali, da quanti ammirano la spregiudicatezza del leader a quanti, pur non mostrandosi vicini alle sue posizioni, concordano sulla necessità – e spesso auspicano – di tornare alle urne, al fine di liberarsi da una coalizione di governo percepita come lontana dai canoni della convenzionale offerta politica espressa dai partiti «storici» della seconda Repubblica.

## Riferimenti Bibliografici

- Brandes U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2): 163–177. <https://doi.org/10.1080/0022250X.2001.9990249>
- Carli R. (1990). Il processo di collusione nelle rappresentazioni sociali. *Rivista di Psicologia Clinica*, 4: 282-296.
- Castells M. (2007). Communication, Power and Counter-power in the Network Society. *International Journal of Communication*, 1: 238-266
- Cepernich, C., & Novelli, E. (2018). Sfumature del razionale. La comunicazione politica emozionale nell'ecosistema ibrido dei media. *Comunicazione politica*, 1: 13-30.
- Ceron A., Curini L. & Iacus S.. (2013). *Social Media e Sentiment Analysis. L'evoluzione dei fenomeni sociali attraverso la Rete*. Milano: Springer.
- Coluccia A., Ferretti F., Lorenzi L. & Buracchi T. (2008). Media e percezione della sicurezza. Analisi e riflessioni. *Rassegna italiana di criminologia*, 2(2): 326-336.
- Cordella B., Greco F., Carlini K., Greco A., & Tambelli R. (2018). Infertilità e procreazione assistita: evoluzione legislativa e culturale in Italia. *Rassegna di Psicologia*, 35(3), 45-56.
- Fornari F. (1976). *Simbolo e codice: Dal processo psicoanalitico all'analisi istituzionale*. Milano: Feltrinelli.
- Freeman L.C. (1979). Centrality in social networks conceptual clarification. *Social Networks*, 1: 215–239.
- Fronzetti Colladon A. (2018). The Semantic Brand Score. *Journal of Business Research*, 88, 150–160. <https://doi.org/10.1016/j.jbusres.2018.03.026>
- Fronzetti Colladon A. (2020). Forecasting election results by studying brand importance in online news. *International Journal of Forecasting*, 36(2), 414-427. <https://doi.org/10.1016/j.ijforecast.2019.05.013>
- Fronzetti Colladon, A., & Grippa, F. (2020). Brand Intelligence Analytics. In A. Przegalinska, F. Grippa, & P. Gloor (Eds.), *Digital Transformation of Collaboration. Proceedings of the 9th International Conference on Collaborative Innovation Networks*, in press. Cham, Switzerland: Springer International Publishing.
- Fronzetti Colladon A., Grippa F. & Innarella R. (2020). Studying the association of online brand importance with museum visitors: An application of the semantic brand score. *Tourism Management Perspectives*, 33, 100588. <https://doi.org/10.1016/j.tmp.2019.100588>
- Garland D. (2001). *The Culture of Control: Crime and Social Order in Contemporary Society*. Oxford: Oxford University Press
- Gentry J. (2016). R Based Twitter Client. R package version 1.1.9.
- Giordano V. (2016). Squarciare i veli? *Jura Gentium*, 13(2): 6-21.
- Gloor P.A. (2017). *Sociometrics and Human Relationships: Analyzing Social Networks to Manage Brands, Predict Trends, and Improve Organizational Performance*. London, UK: Emerald Publishing Limited.
- Grandcolas U., Rettie R. & Marusenko K. (2003). Web Survey Bias: Sample or Mode Effect? *Journal of Marketing Management*, 19(5–6): 541–561. <https://doi.org/10.1080/0267257X.2003.9728225>
- Greco F. (2016). *Integrare la disabilità. Una metodologia interdisciplinare per leggere il cambiamento culturale*. Milano: Franco Angeli.
- Greco F., Alaimo L. & Celardo L. (2018). Brexit and Twitter: The voice of people. In D.F. Iezzi, L. Celardo & M. Misuraca (Eds.), *JADT' 18. Proceedings of the 14th International Conference on Statistical Analysis of Textual Data* (pp. 327-334). Roma: Universitalia.

- Greco F., Maschietti D. & Polli A. (2017). Emotional text mining of social networks: The French pre-electoral sentiment on migration. *Rivista Italiana di Economia Demografia e Statistica*, 71(2), 125-136.
- Greco F. & Polli A. (2019a). Emotional Text Mining: Customer profiling in brand management. *International Journal of Information Management*. <https://doi.org/10.1016/j.ijinfomgt.2019.04.007>
- Greco F. & Polli A. (2019b). Vaccines In Italy: The Emotional Text Mining of Social Media. *Rivista Italiana di Economia Demografia e Statistica*, 73(1), 89-98.
- Greco F. & Polli A. (2019c). Anatomy of a government crisis. Political institutions, security, and consensus. In L.S. Alaimo, A. Arcagni, E. di Bella, F. Maggino & M. Trapani (Eds.), *Libro dei Contributi Brevi: AIQUAV 2019, VI Convegno Nazionale dell'Associazione Italiana, per gli Studi sulla Qualità della Vita, Benessere Collettivo e Scelte Individuali, Fiesole (FI), 12-14 Dicembre 2019* (pp. 177-183). Genova: Genova University Press. ISBN: 978-88-94943-76-4
- Hopkins D. & King G. (2010). A method of automated nonparametric content analysis for social science. *American J. Pol. Sci.*, 54(1): 229-247.
- Jivani A.G. (2011). A Comparative Study of Stemming Algorithms. *International Journal of Computer Technology and Applications*, 2(6): 1930–1938. <https://doi.org/10.1.1.642.7100>
- Keller K.L. (1993). Conceptualizing, Measuring, and Managing Customer-Based Brand Equity. *Journal of Marketing*, 57(1): 1–22.
- Lancia F. (2017). *User's Manual: Tools for text analysis*. T-Lab version Plus 2017.
- Lebart L. & Salem A. (1994). *Statistique Textuelle*. Paris: Dunod.
- Liu B. (2012). *Sentiment analysis: Mining opinions, sentiments, and emotions*. *Sentiment analysis: Mining opinions, sentiments, and emotions*. Morgan & Claypool.
- Malik U. (2019). Python for NLP: Introduction to the Pattern Library. <https://stackabuse.com/python-for-nlp-introduction-to-the-pattern-library/>.
- Matte Blanco I. (1981). *L'inconscio come insieme infiniti: Saggio sulla bi-logica*. Torino: Einaudi.
- Moscovici, S. (2005). *Le rappresentazioni sociali*. Bologna: Il Mulino.
- Olson K. (2006). Survey Participation, Nonresponse Bias, Measurement Error Bias, and Total Bias. *Public Opinion Quarterly*, 70(5): 737–758. <https://doi.org/10.1093/poq/nfl038>
- Poulantzas N. (1973). *Political Power and Social Classes*. London: NLB.
- Salvatore S. & Freda M.F. (2011). Affect, unconscious and sensemaking: A psychodynamic, semiotic and dialogic model. *New Ideas in Psychology*, 29(2): 119–135.
- Savaresi S.M. & Boley D.L. (2004). A comparative analysis on the bisecting K-means and the PDDP clustering algorithms. *Intelligent Data Analysis*, 8(4): 345-362.
- Schoen H., Gayo-Avello D., Metaxas P., Mustafaraj E., Strohmaier M. & Gloor P. (2013). The power of prediction with social media. *Internet Res.*, 23(5): 528-543.
- Strange S. (1998). *Chi governa l'economia mondiale*, Bologna: Il Mulino.
- Warner S.L. (1965). Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association*, 60(309): 63–69. <https://doi.org/10.1080/01621459.1965.10480775>
- Wasserman S. & Faust K. (1994). *Social Network Analysis: Methods and Applications*. *American Ethnology*. <https://doi.org/10.1525/ae.1997.24.1.219>