

L'impatto emotivo della comunicazione istituzionale durante la pandemia di Covid-19: uno studio di Twitter Sentiment Analysis.

Gloria Gagliardi
Università di Napoli "L'Orientale"
Napoli, Italy
ggagliardi@unior.it

Lorenzo Gregori
Università di Firenze
Firenze, Italy
lorenzo.gregori@unifi.it

Alice Suozzi
Università Ca' Foscari Venezia
Venezia, Italy
956549@stud.unive.it

Abstract

Italiano. L'articolo si propone di investigare l'impatto delle comunicazioni istituzionali durante la crisi sanitaria causata dal Covid-19 in Italia, attraverso l'analisi computazionale dell'attività di micro-blogging su Twitter. A partire dal corpus TWITA, abbiamo realizzato una Sentiment Analysis per osservare eventuali cambiamenti nella polarità delle scritte degli utenti, in reazione a undici discorsi istituzionali tenuti durante l'emergenza sanitaria. Variazioni sono state osservate in reazione a tre discorsi tenuti dal Presidente del Consiglio rispettivamente il 4 marzo, il 9 marzo e il 26 aprile 2020.

English. *This paper aims at investigating the impact of institutional communications during the health crisis due to Covid-19 pandemic in Italy, through the analysis of micro-blogging activities on Twitter by means of NLP techniques. We performed a Sentiment Analysis on the TWITA corpus, to pinpoint potential correlations between opinion polarity (positive or negative) of the users and public speeches during the outbreak. Our findings show changes in sentiment polarity related to three institutional speeches delivered by the Italian Prime Minister Giuseppe Conte on March, 4th, March, 9th, and April, 26th 2020.*

1 Introduzione

L'epidemia Covid-19 si è rapidamente diffusa da Wuhan, in Cina, a numerose nazioni in tutto il mondo e il 28 marzo 2020 l'Organizzazione

Mondiale della sanità (OMS) ha ufficialmente dichiarato lo stato di pandemia del Covid-19. In Italia, il primo caso è stato riportato il 21 febbraio 2020; i contagi si sono poi diffusi a diverse regioni e l'8 marzo 2020 l'intera nazione è stata dichiarata zona rossa. È stato inoltre stabilito un lockdown nazionale, che è durato circa due mesi, fino al 4 maggio 2020. La situazione di emergenza è stata drammatica e completamente nuova: le comunicazioni istituzionali, finalizzate ad illustrare le misure di contenimento e ad indirizzare il comportamento della popolazione, hanno dunque giocato un ruolo determinante. La loro diffusione è avvenuta sia attraverso la televisione pubblica, sia attraverso i social network, come Facebook, Instagram e Twitter; di conseguenza, questi sono diventati i principali canali di diffusione di informazioni e di condivisione di opinioni.

La Sentiment Analysis è un campo di studi multidisciplinare finalizzato ad investigare e valutare le opinioni espresse nei testi (Beigi et al., 2016; Mejova et al., 2013; Zimbardo et al., 2018) cioè l'orientamento (positivo o negativo) che il parlante esprime verso un oggetto (Jurafsky and Martin, 2019) attraverso dispositivi lessicali. Twitter è una fonte privilegiata di dati per l'analisi di emozioni e opinioni, tanto che oggi si parla di Twitter Sentiment Analysis (TSA) per identificare la specifica branca di ricerca basata su dati estratti da questo social network. Oggi la TSA è usata in vari ambiti e per diversi scopi, come, ad esempio, monitorare le opinioni degli utenti su prodotti commerciali (Ghiassi et al., 2013; Jansen et al., 2009), studiare gli orientamenti politici (Mejova and Srinivasan, 2012; Mejova et al., 2013; Garcia and Thelwall, 2013; Wang et al., 2012; Wang et al., 2014), analizzare i livelli e le cause di stress tra gli adolescenti (Basili et al., 2017) o l'opinione pubblica riguardo i vaccini (Tavoschi et al., 2020).

Questo studio utilizza la TSA per investigare l'impatto che le comunicazioni istituzionali hanno

Data	Parlante	Argomento	Durata	Parole
04/03/20	G. Conte	Prime azioni del governo per limitare la diffusione del virus (es. chiusura di scuole e università, sospensione degli eventi sportivi)	4'58"	831
05/03/20	S. Mattarella	Primo discorso sull'emergenza coronavirus	3'40"	455
08/03/20	G. Conte	La Lombardia e altre 14 province del Nord sono dichiarate "zona rossa"	15'44"	2357
09/03/20	G. Conte	Decreto "Io resto a casa"	6'37"	995
11/03/20	G. Conte	L'OMS dichiara la pandemia; tutte le attività vengono chiuse (ad eccezione di supermercati e farmacie)	8'59"	1392
21/03/20	G. Conte	Decreto "Chiudi Italia": vengono imposte ulteriori limitazioni alle attività produttive e agli spostamenti individuali	7'06"	942
27/03/20	S. Mattarella	Secondo discorso sull'emergenza coronavirus	7'28"	1105
01/04/20	G. Conte	Il governo estende il lockdown fino al 13 aprile	4'50"	777
10/04/20	G. Conte	Il governo estende il lockdown fino al 3 maggio	19'47"	3124
26/04/20	G. Conte	Annuncio dell'allentamento delle restrizioni a partire dal 4 maggio ("Fase 2")	28'34"	4538
16/05/20	G. Conte	Annuncio di un ulteriore allentamento delle restrizioni	12'04"	1767

Table 1: Corpus Italic-Covid19

avuto durante la crisi sanitaria di Covid-19 in Italia. Nello specifico, è stato analizzato il sentiment di tweet italiani sul coronavirus per identificare correlazioni tra la polarità delle opinioni degli utenti (positiva o negativa) e i discorsi tenuti dal Presidente del Consiglio e dal Presidente della Repubblica durante la pandemia.

2 Materiali

2.1 Il corpus Italic-Covid19

Preliminarmente allo studio, è stato costruito un piccolo corpus di dichiarazioni e conferenze stampa pronunciate dal Presidente del Consiglio Giuseppe Conte e dal Presidente della Repubblica Sergio Mattarella, chiamato Italic-Covid19 (Gagliardi and Suozzi, 2020).¹ Tutti i dati provengono dal canale YouTube ufficiale di Palazzo Chigi² e del Quirinale³.

Attualmente, il corpus include undici testi orali, pronunciati tra il 3 marzo e il 16 maggio 2020. Tab. 1 riassume il loro contenuto. Per ogni video è stata realizzata una trascrizione ortografica, conforme al formato L-AcT (Cresti and Moneglia, 2018), una versione del formato standardizzato CHAT (MacWhinney, 2000).

2.2 Il corpus Twitter Covid-19

I tweet utilizzati per questo studio sono derivati dal corpus 40wita,⁴ un dataset di tweet italiani

raccolto da Basile e Caselli (2020) durante il periodo di emergenza della pandemia, come parte dell'iniziativa *Computational Linguistics and the Covid-19 Outbreak*⁵ promossa dall'Associazione Italiana di Linguistica Computazionale (AILC). Il dataset è stato costruito giornalmente dal 1 febbraio 2020, filtrando TWITA (Basile and Nissim, 2013; Basile et al., 2018), la più grande raccolta di messaggi pubblicati su social media in lingua italiana, con le parole-chiave relative al Covid-19,⁶.

2.3 Il dataset

Al fine di ottenere delle statistiche generali, è stato esaminato un lasso di tempo di 88 giorni, dal 26 febbraio al 23 maggio 2020. Il periodo considerato inizia una settimana prima della prima comunicazione istituzionale, e finisce una settimana dopo l'ultima. Il numero di tweet giornalieri sul coronavirus durante questo periodo va da un minimo di 9.245 a un massimo di 51.490, per un totale di 2.298.566 post. In fig. 1 è illustrata la distribuzione di questi numeri in funzione del tempo. Si può chiaramente osservare una crescita significativa nei primi giorni di marzo, quando, in alcune regioni, il lockdown è iniziato; il picco viene raggiunto il 10 marzo 2020, in corrispondenza

⁵<http://www.ai-lc.it/covid-19/>

⁶*covid, covid19, covid-19 corona virus, coronavirus, quarantena, autoisolamento, auto-isolamento, iorestoacasa, stateacasa, COVID19Italia, redditodicitadinaza, eurobond, coronabond, restiamoacasa, preghiamoinsieme, NoMes, milanononsiferma, bergamononsiferma, l'italianonsiferma, abbracciauncinese, iononsonouvirus, iononmiferma, aperisera, covidunustria, italiazonarossa, bergamoisrunning, quarantena, chiudetutto, aprietutto, Curaltalia, circordiamotutto, oggisciopero, chiudiamolefabbriche, iononrinuncioallegtrazioni, andràtuttobene, INPSdown, percheQuando, cercareDi, ringraziarevoglio, 600euro, CineINPS, COVID19Pandemic.*

¹<https://euicccovid19.altervista.org/>

²<https://www.youtube.com/user/governoit>

³<https://www.youtube.com/user/presidenzarepubblica>

⁴<http://twita.di.unito.it/dataset/40wita>

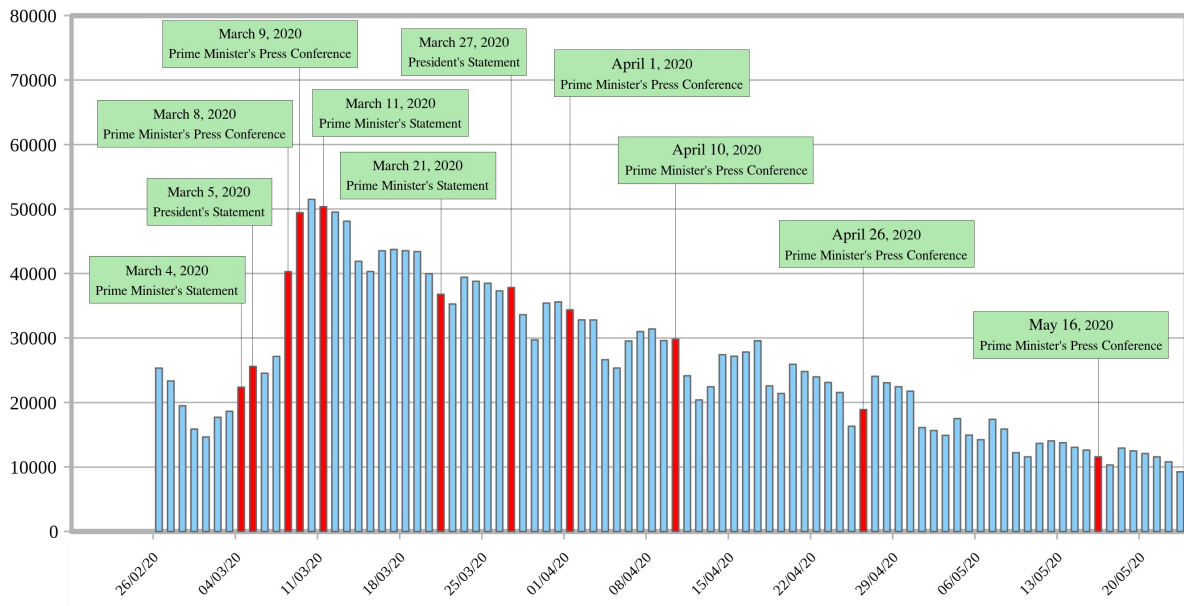


Figure 1: Numero di tweet sul Covid-19 in Italia durante il lockdown.

dell'estensione del lockdown a tutto il territorio nazionale. Segue una diminuzione graduale fino alla fine di maggio, quando la maggior parte delle attività viene riaperta.

La Sentiment Analysis è stata realizzata con Sentita (Nicola, 2018).⁷ Lo strumento, specifico per i tweet italiani, fornisce due valori indipendenti: un valore tra 0 e 1 per la polarità positiva (0 = neutro, 1 = completamente positivo) e uno tra 0 e 1 per la polarità negativa (0 = neutro, 1 = completamente negativo). Per ridurre la complessità del compito, sono stati inclusi nell'analisi solo i tweet con una polarità chiaramente orientata, escludendo quelli con bassa polarità (positivi e negativi sotto la soglia di 0.5) e quelli polarizzati sia in positivo che in negativo (con entrambi i valori superiori a 0.5). Due esempi di tweet analizzati con Sentita sono riportati in tabella 2.

3 Analisi della *sentiment polarity*

Il dataset appena descritto è stato utilizzato per misurare se e in che misura le reazioni dei cittadini alle comunicazioni istituzionali siano riflesse nei loro tweet.

A tal fine, è necessario considerare che la polarità dei tweet non è uniforme durante il giorno (Larsen et al., 2015): i tweet pubblicati al mattino sono solitamente più neutri di quelli della sera. Per avere una panoramica di questa vari-

Tweet	pos.	neg.
Che palle questo coronavirus mi sta rovinando tutti i piani e mi sta togliendo troppe gioie mi uccido	0.046	0.963
So che non si può abbracciare per il Covid-19 ma io voglio mandare un enorme abbraccio virtuale a medici, infermieri, operatori ospedalieri che in queste settimane lavorano senza sosta per garantire la salute di tutti. Grazie, grazie, grazie.	0.985	0.093

Table 2: Due esempi di tweet con polarità negativa e positiva.

azione all'interno del corpus, è stato misurato l'andamento di tweet positivi e negativi nell'arco delle 24 ore e la media è stata calcolata sull'intero periodo di osservazione (88 giorni). Questo ha permesso di comparare l'andamento di tweet positivi e negativi intorno agli eventi comunicativi con la media giornaliera, per determinare eventuali deviazioni. Per l'analisi è stata considerata una finestra temporale che va da quattro ore prima a quattro ore dopo ogni comunicazione: all'interno di essa è stata misurata la percentuale di tweet positivi e negativi per ogni ora. Per evitare ogni potenziale fluttuazione oraria, è stata misurata la deviazione

⁷<https://nicgian.github.io/Sentita/>

dist	hour	pos	neg	pos_dev	neg_dev
-4	15	1.71%	10.91%	-1.44%	1.56%
-3	16	1.80%	11.73%	-1.20%	2.74%
-2	17	2.26%	9.93%	-0.55%	1.07%
-1	18	2.22%	10.17%	-0.53%	1.30%
0	19	1.99%	11.26%	-1.16%	1.08%
1	20	1.79%	12.72%	-1.66%	0.77%
2	21	2.44%	13.01%	-1.70%	-0.01%
3	22	1.14%	10.13%	-2.98%	-3.35%
4	23	2.15%	10.39%	-1.76%	-3.45%

Table 3: Analisi dei tweet nella finestra temporale [-4,+4] rispetto alla comunicazione istituzionale del Presidente del Consiglio del 4 marzo 2020.

di polarità rispetto alla media calcolata su tutti gli 88 giorni. In tab. 3 sono riportati, a titolo di esempio, i dati relativi alla prima comunicazione, la dichiarazione del Presidente del Consiglio del 4 marzo 2020, trasmessa tra le 19:00 e le 20:00.

I dati rappresentati sono i seguenti: distanza in ore dall'evento comunicativo (*dist*), ora del giorno (*hour*), percentuale di tweet positivi (*pos*), differenza tra *pos* e la media percentuale di tweet positivi per la stessa ora del giorno (*pos_dev*); analogamente sono calcolati *neg* e *neg_dev* sui tweet negativi.

I valori delle ultime due colonne (*pos_dev* e *neg_dev*) sono stati divisi in due gruppi: le percentuali relative ai tweet precedenti all'evento (A) e quelle relative ai tweet successivi all'evento (B). Numericamente i tweet considerati per l'analisi (nella finestra temporale [-4,+4] ore da ogni evento) sono 161.197, di cui 27.057 con polarità positiva o negativa. Abbiamo applicato un modello di regressione lineare con una variabile *dummy* per ogni evento, assegnando un valore di 1.0 agli elementi di (A) e di 0.0 a quelli di (B). I *p*-value risultanti sono stati utilizzati per identificare gli eventi dopo i quali la polarità dei tweet è cambiata significativamente.

4 Risultati e discussione

I risultati della regressione lineare sono riportati nella tab. 4.

I dati mostrano che per tre comunicazioni istituzionali si registrano variazioni significative: sono quelle rilasciate il 4 marzo, il 9 marzo e il 26 aprile 2020. Tali risultati sono coerenti con il contenuto delle comunicazioni: le prime due (4 e 9 marzo) hanno annunciato, rispettivamente, l'istituzione delle prime zone rosse e l'estensione del lockdown a tutto il territorio nazionale. L'ultima, al contrario, ha annunciato

Date	POS <i>p</i>	NEG <i>p</i>
03/04/20	0.03152	0.03406
03/05/20	0.9412	0.3768
03/08/20	0.1195	0.721
03/09/20	0.006432	0.001612
03/11/20	0.07579	0.371
03/21/20	0.6074	0.2086
03/27/20	0.9202	0.08326
04/01/20	0.07515	0.8863
04/10/20	0.09524	0.7186
04/26/20	0.449	0.0006091
05/16/20	0.7202	0.4424

Table 4: *p*-value della Regressione Lineare con variabile dummy applicata ai dati riferiti alla finestra temporale [-4,+4].

le misure contenitive relative alla cosiddetta Fase 2 (cioè la parziale riapertura delle attività, e la fine del lockdown). È importante sottolineare che i tre eventi hanno suscitato tre reazioni diverse (fig. 2):

- 4 marzo 2020: significativa diminuzione sia di tweet positivi sia di negativi;
- 9 marzo 2020: significativa crescita sia di tweet positivi sia di tweet negativi;
- 26 aprile 2020: significativa crescita esclusivamente di tweet negativi, con percentuale di tweet positivi invariata.

La prima dichiarazione è stata seguita da un forte calo sia di tweet positivi sia di negativi: questo è coerente con il clima di grande incertezza che ha caratterizzato i primi giorni di marzo 2020. Il Covid-19 si stava, infatti, ancora diffondendo prevalentemente nel Nord Italia, e la dichiarazione analizzata ha stabilito le prime misure contro un'ulteriore diffusione del virus. La percezione di pericolo crescente, da un lato, e l'attesa generale di altre misure contenitive, dall'altro, si traduce in opinioni meno chiaramente orientate.

La seconda dichiarazione è servita ad annunciare il decreto "Io resto a casa": da quel momento, tutta la nazione è diventata "zona protetta". Le reazioni emotive a questa comunicazione sono ambivalenti, con una crescita importante di tweet sia positivi, che negativi. Questa "doppiezza" emotiva mostra da un lato la paura causata dall'epidemia in corso, e il lutto per le vittime, in quel momento particolarmente numerose; dall'altro, l'approvazione che il lockdown nazionale ha suscitato nella popolazione, e la tendenza a rimanere ottimisti e soprattutto uniti, esemplificata dall'hashtag *andràtuttobene*.

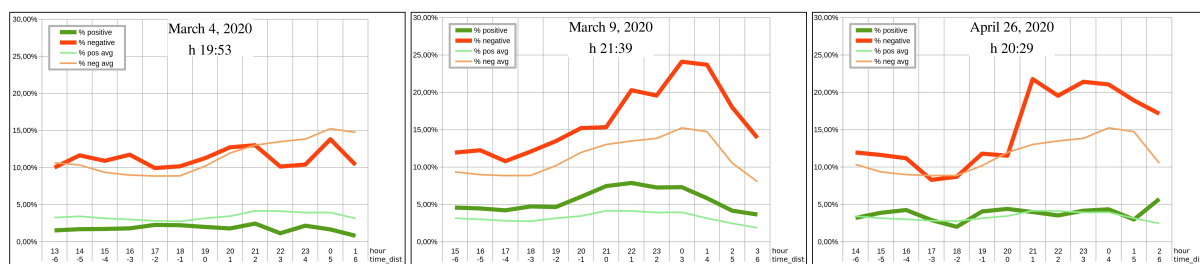


Figure 2: Percentuale di tweet con sentiment positivo e negativo intorno alle comunicazioni del 4 marzo, 9 marzo e 26 aprile 2020 (linee spesse); sono riportate anche le percentuali delle medie orarie (linee fini).

La terza comunicazione ha, invece, provocato solo una significativa crescita di tweet negativi: il 26 aprile sono state annunciate le linee guida per la progressiva riapertura, e l'ingresso nella cosiddetta Fase 2. Una reazione così negativa è imputabile prevalentemente alla delusione: le misure annunciate sono state percepite come ancora troppo restrittive, tanto che molti tweet ironizzavano sulla sostanziale identità di Fase 1 e Fase 2. La delusione è derivata da una precedente aspettativa: un rapido ritorno alla normalità. Questo desiderio, a sua volta, è stato provocato da effetti più generali che la pandemia ha avuto sul benessere mentale ed emotivo della popolazione italiana. La pandemia di Covid-19, analogamente a quanto osservato per altre epidemie, ha suscitato reazioni psicologiche negative, come un'incidenza maggiore di depressione, stress (Shultz et al., 2016), preoccupazione (Thompson et al., 2017), e ansia di essere contagiati (Horney et al., 2010); il tutto è stato esacerbato dall'impossibilità di accedere a servizi di supporto e di dedicarsi ad attività come hobbies e sport.

5 Conclusioni

Per concludere, questo studio ha lo scopo di investigare le reazioni dei cittadini italiani alle comunicazioni istituzionali durante l'epidemia di Covid-19. Grazie al crescente utilizzo dei social network, è infatti possibile esplorare le reazioni psicologiche ad eventi traumatici, sia individuali sia collettivi: secondo i nostri dati, le comunicazioni istituzionali che hanno provocato reazioni psicologico-emotive più forti sono quelle pronunciate il 4 e 9 marzo e il 26 aprile 2020, che hanno annunciato le misure più drastiche di contenimento dell'epidemia. Ulteriori studi potrebbero espandere questo lavoro, considerando contesti nazionali differenti (sia dove sono state attuate

misure di contenimento del virus simili a quelle italiane, sia dove sono state attuate misure molto diverse). Inoltre, potrebbero essere oggetti di analisi eventi socio-politici diversi, come suggerito dallo studio EmoItaly⁸ (es. elezioni politiche, Giornata Internazionale dei Lavoratori, omicidio di George Floyd).

References

- Valerio Basile and Tommaso Caselli. 2020. 40twita 1.0: An collection of Italian Tweets during the COVID-19 Pandemic. <http://twita.di.unito.it/dataset/40twita>.
- Valerio Basile and Malvina Nissim. 2013. Sentiment analysis on Italian tweets. In Alexandra Balahur, Erik van der Goot, and Andres Montoyo, editors, *Proceedings of the 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, page 100–107. Association for Computational Linguistics.
- Valerio Basile, Mirko Lai, and Manuela Sanguinetti. 2018. Long-term Social Media Data Collection at the University of Turin. In Elena Cabrio, Alessandro Mazzei, and Fabio Tamburini, editors, *Proceedings of the Fifth Italian Conference on Computational Linguistics (CLiC-it 2018)*, volume 2253 of *CEUR Workshop Proceedings*, pages 40–45. CEUR-WS.org.
- Roberto Basili, Valentina Bellomaria, Niels J. Bugge, Danilo Croce, Francesco De Michele, Federico Fiori Nastro, Paolo Fiori Nastro, Chantal Michel, Stefanie J. Schmidt, and Frauke Schultze-Lutter. 2017. Monitoring Adolescents' Distress using Social Web data as a Source: the InsideOut Project. In Roberto Basili, Malvina Nissim, and Giorgio Satta, editors, *Proceedings of the Fourth Italian Conference on Computational Linguistics CLiC-it*, Rome.
- Ghazaleh Beigi, Xia (Ben) Hu, Maciejewski Ross, and Huan Liu, 2016. *An Overview of Sentiment Analy-*

⁸<http://corpora.ficlit.unibo.it/EmoItaly/>

- sis in Social Media and Its Applications in Disaster Relief*. Springer, New York.
- Emanuela Cresti and Massimo Moneglia. 2018. Chapter 13. The illocutionary basis of information structure: The Language into Act Theory (L-AcT). In Evangelia Adamou, Katharina Haude, and Martine Vanhove, editors, *Information Structure in Lesser-described Languages. Studies in prosody and syntax*, pages 360–402. John Benjamins, Amsterdam - Philadelphia.
- Gloria Gagliardi and Alice Suozzi. 2020. ItalC-Covid19: Institutional Communication in Times of COVID-19. <http://gloriagagliardi.altervista.org/psychological-reactions-to-the-pandemic/>.
- David Garcia and Mike Thelwall. 2013. Political alignment and emotional expression in Spanish Tweets. In *Proceedings of the TASS Workshop at SEPLN*.
- Manoochehr Ghiassi, J. Skinner, and David Zimbra. 2013. Twitter brand Sentiment Analysis: A hybrid system using n-gram analysis and Dynamic Artificial Neural Network. *Expert Systems with Applications*, 40(16).
- Jennifer A. Horney, Zack Moore, Meredith Davis, and Pia D. M. MacDonald. 2010. Intent to receive pandemic influenza A (H1N1) vaccine, compliance with social distancing and sources of information in NC, 2009. *PloS One*, 5(5).
- Jim Jansen, Mimi Zang, Kate Sobel, and Abdur Chowdury. 2009. Twitter power: Tweets as electronic word of mouth. *Journal of the American Society for Information Science and Technology*, 60(11).
- Dan Jurafsky and H. Martin, James. 2019. *Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, Third Edition Draft*. Pearson, New York.
- Mark E Larsen, Tjeerd W Boonstra, Philip J Batterham, Bridianne O’Dea, Cecile Paris, and Helen Christensen. 2015. We feel: mapping emotion on Twitter. *IEEE journal of biomedical and health informatics*, 19(4):1246–1252.
- Brian MacWhinney. 2000. *The CHILDES Project: Tools for analyzing talk*. Lawrence Erlbaum Associates, Mahwah.
- Yelena Mejova and Padimni Srinivasan. 2012. Political Speech in Social Media Streams: YouTube Comments and Twitter Posts. In Noshir S Contractor, Brian Uzzi, Michael Macy, and W. Nejdl, editors, *Proceedings of the 3rd Annual ACM Web Science Conference*, New York. Association for Computing Machinery.
- Yelena Mejova, Padmini Srinivasan, and Bob Boynton. 2013. GOP primary season on Twitter: popular political sentiment in social media. In *Proceedings of the ACM Web Search and Data Mining Conference (WSDM’13)*, Rome. ACM.
- Giancarlo Nicola. 2018. Bidirectional Attentional LSTM for Aspect Based Sentiment Analysis on Italian. In Tommaso Caselli, Nicole Novielli, Viviana Patti, and Paolo Rosso, editors, *Proceedings of the Sixth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian. Final Workshop (EVALITA 2018)*, volume 2263 of *CEUR Workshop Proceedings*. CEUR-WS.org.
- M. James Shultz, Janice L. Cooper, Florence Bain-gana, Zelde Oquendo, Maria A. and Espinel, Benjamin M. Althouse, Louis HERNs Marcelin, Sherry Towers, Maria Espinola, Laurie McCoy, Clyde B. and Mazurik, Milton L. Wainberg, Yuval Neria, and Andreas Rechkemmer. 2016. The role of fear-related behaviors in the 2013–2016 West Africa Ebola virus disease outbreak. *Current Psychiatry Reports*, 18(104).
- Lara Tavoschi, Filippo Quattrone, Eleonora D’Andrea, Pietro Ducange, Marco Vabanesi, Francesco Marcelloni, and Pier Luigi Lopalco. 2020. Twitter as a sentinel tool to monitor public opinion on vaccination: an opinion mining analysis from September 2016 to August 2017 in Italy. *Human Vaccines & Immunotherapeutics*, 16(5).
- Rebecca R. Thompson, Dana Rose Garfin, E. Alison Holman, and Roxane Cohen Silver. 2017. Distress, worry, and functioning following a global health crisis: A national study of Americans’ responses to Ebola. *Clinical Psychological Science*, 5(3).
- Hao Wang, Dogan Can, Abe Kazemzadeh, François Bar, and Shrikanth Narayanan. 2012. A System for Real-time Twitter Sentiment Analysis of 2012 U.S. Presidential Election Cycle. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, Jeju. Association for Computational Linguistics.
- Yu Wang, Tom Clark, Eugene Agichtein, and Jeffrey Staton. 2014. Towards Tracking Political Sentiment through Microblog Data. In *Proceedings of the Joint Workshop on Social Dynamics and Personal Attributes in Social Media*, pages 88–93, Baltimore.
- David Zimbra, Ahmed Abbasi, Daniel Zeng, and Hsinchun Chen. 2018. The State-of-the-Art in Twitter Sentiment Analysis: A Review and Benchmark Evaluation. *ACM Transactions on Management Information Systems*, 9(5).