

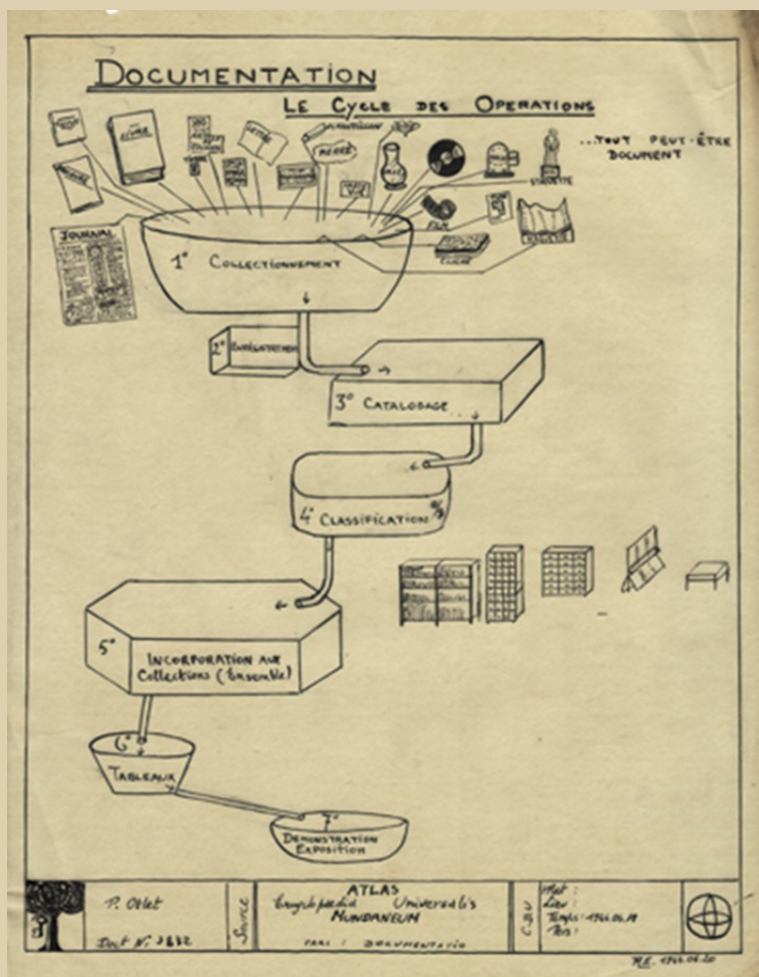
# AIDa informazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

NUMERO 1-2

ANNO 40

GENNAIO-GIUGNO 2022



# **AIDA**informazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

Fondata nel 1983 da Paolo Bisogno

**Proprietario della rivista:**

Università della Calabria

**Direttore Scientifico:**

Roberto Guarasci, *Università della Calabria*

**Direttore Responsabile:**

Fabrizia Flavia Sernia

**Comitato scientifico:**

Anna Rovella, *Università della Calabria*;

Maria Guercio, *Sapienza Università di Roma*;

Giovanni Adamo, *Consiglio Nazionale delle Ricerche* †;

Claudio Gnoli, *Università degli Studi di Pavia*;

Ferruccio Diozzi, *Centro Italiano Ricerche Aerospaziali*;

Gino Roncaglia, *Università della Toscana*;

Laurence Favier, *Université Charles-de-Gaulle Lille 3*;

Madjid Ihadjadene, *Université Vincennes-Saint-Denis Paris 8*;

Maria Mirabelli, *Università della Calabria*;

Agustín Vivas Moreno, *Universidad de Extremadura*;

Douglas Tudhope, *University of South Wales*;

Christian Galinski, *International Information Centre for Terminology*;

Béatrice Daille, *Université de Nantes*;

Alexander Murzaku, *College of Saint Elizabeth, USA*;

Federico Valacchi, *Università di Macerata*.

**Comitato di redazione:**

Antonietta Folino, *Università della Calabria*;

Erika Pasceri, *Università della Calabria*;

Maria Taverniti, *Consiglio Nazionale delle Ricerche*;

Maria Teresa Chiaravallotti, *Consiglio Nazionale delle Ricerche*;

Assunta Caruso, *Università della Calabria*;

Claudia Lanza, *Università della Calabria*.

**Segreteria di Redazione:**

Valeria Rovella, *Università della Calabria*

**Editrice:** Cacucci Editore S.a.s.

Via D. Nicolai, 39 – 70122 Bari (BA)

[www.cacuccieditore.it](http://www.cacuccieditore.it)

e-mail: [riviste@cacuccieditore.it](mailto:riviste@cacuccieditore.it)

Telefono 080/5214220



# AIDAinformazioni

## RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

«AIDAinformazioni» è una rivista scientifica che pubblica articoli inerenti le Scienze dell'Informazione, la Documentazione, la Gestione Documentale e l'Organizzazione della Conoscenza. È stata fondata nel 1983 quale rivista ufficiale dell'Associazione Italiana di Documentazione Avanzata e nel febbraio 2014 è stata acquisita dal Laboratorio di Documentazione dell'Università della Calabria. La rivista si propone di promuovere studi interdisciplinari oltre che la cooperazione e il dialogo tra profili professionali aventi competenze diverse, ma interdipendenti. I contributi possono riguardare topics quali Documentazione, Scienze dell'informazione e della comunicazione, Scienze del testo e del documento, Organizzazione e Gestione della conoscenza, Terminologia, Statistica testuale e Linguistica computazionale e possono illustrare studi sperimentali in domini specialistici, casi di studio, aspetti e risultati metodologici conseguiti in attività di ricerca applicata, presentazioni dello stato dell'arte, ecc.

«AIDAinformazioni» è riconosciuta dall'ANVUR come rivista di Classe A per l'Area 11 – Settore 11/A4 e censita per le Aree 10 – Scienze dell'antichità, filologico-letterarie e storico-artistiche; 11 – Scienze storiche, filosofiche, pedagogiche e psicologiche; 12 – Scienze giuridiche; 14 – Scienze politiche e sociali, così come dall'ARES (Agence d'évaluation de la recherche et de l'enseignement supérieur) che la annovera tra le riviste scientifiche dell'ambito delle Scienze dell'Informazione e della Comunicazione. La rivista è, inoltre, indicizzata in: ACNP – Catalogo Italiano dei Periodici; BASE – Bielefeld Academic Search Engine; ERIH PLUS – European Reference Index for the Humanities and Social Sciences – EZB – Elektronische Zeitschriftenbibliothek – Universitätsbibliothek Regensburg; Gateway Bayern; KVK – Karlsruhe Virtual Catalog; Letteratura Professionale Italiana – Associazione Italiana Biblioteche; The Library Catalog of Georgetown University; SBN – Italian union catalogue; Summon™ – by SerialsSolutions; Ulrich's; UniCat – Union Catalogue of Belgian Libraries; Union Catalog of Canada; LIBRIS – Union Catalogue of Swedish Libraries; Worldcat.

I contributi sono valutati seguendo il sistema del *double blind peer review*: gli articoli ricevuti dal comitato scientifico sono inviati in forma anonima a due referee, selezionati sulla base della loro comprovata esperienza nei topics specifici del contributo in valutazione.

# AIDAinformazioni

Anno 40

N. 1-2 – gennaio-giugno 2022

CACUCCI  EDITORE  
BARI

---

PROPRIETÀ LETTERARIA RISERVATA

---

© 2022 Cacucci Editore – Bari

Via Nicolai, 39 – 70122 Bari – Tel. 080/5214220

<http://www.cacuccieditore.it> e-mail: [info@cacucci.it](mailto:info@cacucci.it)

Ai sensi della legge sui diritti d'Autore e del codice civile è vietata la riproduzione di questo libro o di parte di esso con qualsiasi mezzo, elettronico, meccanico, per mezzo di fotocopie, microfilms, registrazioni o altro, senza il consenso dell'autore e dell'editore.

# Sommario

## Contributi

STEFANO ALLEGREZZA, La conservazione degli archivi di posta elettronica: sviluppi recenti e prospettive future	9
ANDREA BELLANDI, Le Risorse Linguistiche nell'era del Web Semantico. Un insieme di servizi informatici per la gestione di lessici e terminologie	31
GIORGIA DI MARCANTONIO, E se l'archivio non rispecchia l'istituto? Pavone e il rispecchiamento: analisi di una bozza preliminare	51
GERARDO GIARDIELLO, CHIARA FIORAVANTI, FRANCESCO ROMANO, MARIASOLE RINALDI, Strumenti per il reperimento di informazioni rilevanti in Rete. Applicazioni nel dominio della comunicazione pubblica in materia di immigrazione in Italia	69
CLAUDIA LANZA, Termini e testi nella comunicazione dei gruppi No-vax	89
ANNA ROVELLA, La posta elettronica negli archivi di persona: conservazione e accesso	113
GIOVANNI SALUCCI, Il Blog di qualità. Proposta di un nuovo format di pubblicazione nell'editoria scientifica	129

## Note e rubriche

ROBERTO GUARASCI, Paul Otlet e Suzanne Briet. Note in margine al libro di Paola Castellucci e Sara Mori	147
ERIKA PASCERI, L'evoluzione del <i>Nuovo Glossario di biblioteconomia e scienza dell'informazione</i> ai tempi dello <i>Smart Working</i>	151
CLAUDIO GNOLI, Come mi vuoi, canonica o rivoluzionaria?	153
CLAUDIO GRIMALDI, L'universo affascinante del colore: le sociocromie	157





# Contributi



# Termini e testi nella comunicazione dei gruppi No-vax

CLAUDIA LANZA\*

**ABSTRACT:** The study presents an analysis of the recursive linguistic structures characterizing the No-vax movement's messages on the social media environment, in particular on Twitter, where the collection of specific time-framed public posts is feasible to be achieved and they represent a solid documental dataset in input from which to start the semantic analysis tasks. The goal of this paper is to present a case-study overview of the influence the information flow within web documentary typologies exerts on new forms of linguistic standardized models. The results will display some of the main recurrent semantic chains used by outmembers on Twitter to serve as references to demonstrate how prefixed has become the language proper to specific anti-pandemic vaccination and regulations groups.

*Keywords:* COVID-19, No-vax, Semantic chains, Terminology, Twitter.

## 1. Introduzione

Nell'attuale contesto pandemico da COVID-19 la mole di dati e l'immediatezza con la quale questi vengono resi fruibili al largo pubblico sono caratterizzate da una forte ciclicità e livellamento informativo che hanno determinato un alto grado di disinformazione (Furini and Menegoni 2018). Nello specifico, in questo particolare contesto, si parla di "infodemia", termine derivante dalla fusione del concetto di distribuzione su larga scala delle informazioni e del concetto di epidemia come elemento di viralità in un ecosistema (Rovetta 2022). L'Organizzazione Mondiale della Sanità (OMS) definisce il fenomeno dell'infodemia come:

An infodemic is too much information including false or misleading information in digital and physical environments during a disease outbreak. It causes confusion and risk-taking behaviours that can harm health. It also leads to mistrust in health authorities and undermines the public health response. An

---

\* Università della Calabria, Dipartimento di Culture, Eduazione e Società (DiCES), Arvacata di Rende (CS), Italia. [claudia.lanza@unical.it](mailto:claudia.lanza@unical.it).

infodemic can intensify or lengthen outbreaks when people are unsure about what they need to do to protect their health and the health of people around them. With growing digitization – an expansion of social media and internet use – information can spread more rapidly. This can help to more quickly fill information voids but can also amplify harmful messages.

Infodemic management is the systematic use of risk- and evidence-based analysis and approaches to manage the infodemic and reduce its impact on health behaviours during health emergencies<sup>1</sup>.

Si tratta di un fenomeno che provoca la disseminazione di un elevatissimo numero di dati spesso discordanti tra loro che disorienta i processi di analisi critica da parte dei cittadini, i quali perdono di volta in volta fiducia nelle istituzioni garanti della pubblicazione di notizie potenzialmente risolutive. Questa babele di «narrazioni giornalistiche contraddittorie e spettacolarizzanti» (Pedroni 2020) rende difficile la realizzazione di un sistema di credibilità omogeneo che appoggi le scelte governative nazionali e ne sia promulgatore per un convincimento massivo. Al contrario, i messaggi vengono trasmessi attraverso il web in modo ripetitivo e allo stesso tempo contrastante in quanto riflesso del cambiamento legato alla trasformazione dell'infettività del virus COVID-19 e, di conseguenza, delle azioni di contrasto adottate per fronteggiare i processi evolutivi pandemici. Tutto ciò genera un contesto che facilita la creazione delle cosiddette *echo chambers*, ovvero,

“camere dell'eco” (echo chambers), «ambienti chiusi nei quali le parole e le frasi “risuonano” sempre uguali a sé stesse fino a cancellare il dissenso e a non essere più messe in discussione», ciò che «comporta anche una ulteriore assuefazione a parole e concetti che possono assumere la forma di slogan, “parole d'ordine” pressoché desementizzate» (Negri e Tagliani 2021).

Pertanto, in una tale cornice di viralità informativa e creazione di miscredenze basate su un senso comune di frustrazione proveniente dall'incongruenza tra la diffusione di notizie legate alle misure contrastive e la difficile risoluzione concreta e definitiva dello stato pandemico (Valensise 2021), la disposizione delle opinioni pubbliche non risulta univoca e diverse voci basate su un numero elevato di notizie non attendibili circolate come *fake news* prendono forma in gruppi oppositivi sui social media.

Il movimento dell'antivaccinismo è una posizione radicata nel tempo. Se ne tiene traccia a partire dai gruppi oppositivi all'introduzione dell'obbligo vaccinale contro il vaiolo – *Anti Vaccination League* e *Anti-Compulsory Vaccination League*<sup>2</sup> – considerato come strumento mirato a ledere i diritti e le libertà individuali, nonché causa di malattie collaterali. Tra i sostenitori della formazione

<sup>1</sup> World Health Organization (WHO) – Infodemic, [https://www.who.int/health-topics/infodemic#tab=tab\\_3](https://www.who.int/health-topics/infodemic#tab=tab_3) (ultima consultazione: 26/05/2022).

<sup>2</sup> National Library of Medicine, Digital Collections, <https://collections.nlm.nih.gov/catalog.nlm.nlmuid-101235983-bks> (ultima consultazione: 26/05/2022).

e crescita di gruppi di contrasto alle cure vaccinali spiccano anche figure legate al mondo della medicina, come ad esempio Alexander Milton Ross, medico canadese, editore di “The Anti-Vaccinator”<sup>3</sup> e autore del pamphlet (Fig. 1) che nel 1885 ebbe larga diffusione per la sua contestazione ai benefici apportati dal vaccino scoperto da Edward Jenner e per il suo incitamento alla convinzione dell’inutilità di sottoporre la gente all’obbligo della vaccinazione.

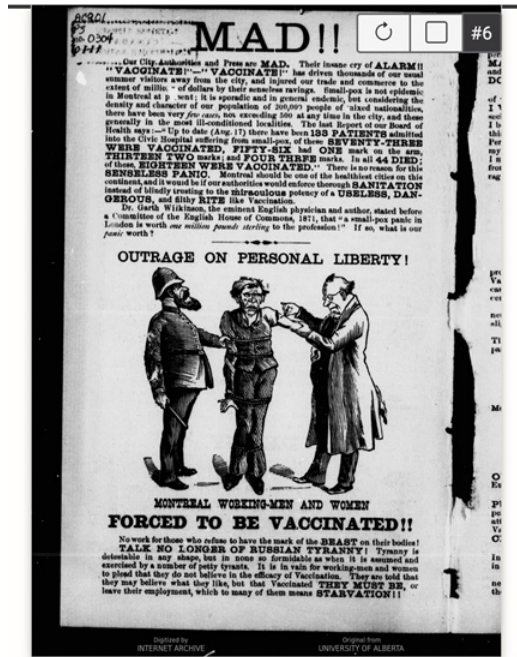


Figura 1: Pamphlet<sup>4</sup>.

*No-vax*<sup>5</sup>, *anti-vaxxers*, *antivaccinatori*, *pro-plague*, *cospirazionisti* sono coloro che, come in passato accadde per le cure contro il vaiolo o morbillo, rimangono scettici nei confronti di scoperte scientifiche che hanno permesso di debellare epidemie e pandemie, basando il loro spirito critico sull’adesione irremovibile a messaggi diffusi dai gruppi oppositivi di appartenenza (Knippenberg, Lossie and Wilke 1994, 289-300).

L’esitazione del singolo e della comunità a sottoporsi alla somministrazione di cure vaccinali – per vari fattori, come ad esempio la mancanza di consapevolezza sui rischi della malattia, il timore dell’ago, l’incertezza sulla disponibilità

<sup>3</sup> The anti-vaccinator, and advocate of cleanliness. Ross, Alexander Milton, 1832-1897.

<sup>4</sup> University of Oxford, Faculty of History, “COVID-19 anti-vaxxers use the same arguments from 135 years ago”, <https://www.history.ox.ac.uk/article/covid-19-anti-vaxxers-use-the-same-arguments-from-135-years-ago> (ultima consultazione: 26/05/2022).

<sup>5</sup> “Chi o che si oppone alla vaccinazione obbligatoria o di massa.” (Trecani, il portale del sapere).

dei vaccini – viene considerata come uno dei dieci rischi globali per la salute pubblica secondo l'OMS (Whitehead et al. 2019).

L'obiettivo del presente articolo consiste in una prima individuazione di un livello di ricorsività nella terminologia e conseguente formulazione scritta delle specifiche asserzioni da parte dei movimenti No-vax in riferimento all'attuale contesto di pandemia da COVID-19. Nel dettaglio, si presenterà un'analisi introduttiva sulle modalità espressive che caratterizzano i messaggi diffusi nei mesi di settembre e ottobre 2021 da un set di profili identificabili come contrari alla campagna vaccinale e alle conseguenti ordinanze nazionali ministeriali come l'obbligo del Green pass. L'indagine ha previsto uno studio analitico dei messaggi divulgati tramite social network (nello specifico, Twitter) attraverso la verifica delle strutture linguistiche più ricorrenti all'interno degli stessi al fine di isolare delle batterie semantiche, sotto una prospettiva di analisi delle nuove forme linguistiche sui social network (Caleffi 2015), ripetute all'interno di gruppi social che guidano alla formazione di messaggi "manifesto" identificativi dell'appartenenza a un'ideologia oppositiva. In questo modo è stato possibile avviare un'indagine preliminare per la rilevazione di sequenze di espressioni che possano rappresentare un livello di prototipicità semantica (Knippenberg 2011) basata sull'osservazione della concatenazione di elementi linguistici.

In particolare, l'analisi che verrà presentata si concentra sulle nuove modalità comunicative impiegate dagli utenti tramite la diffusione di messaggi su social network e, nello specifico, mediante la pratica di costruzione di contenuti informativi basati su hashtag. L'hashtag, come definito dall'Accademica della Crusca, è una

parola (o sequenza continua di parole) preceduta dal simbolo #, usata nell'ambito dei social network per categorizzare e rendere ricercabili contenuti correlati.

È dal luglio 2009 che il social network Twitter ha introdotto ufficialmente gli hashtag come chiavi di ricerca per catalogare e rintracciare commenti attinenti a uno stesso argomento. Un commento con un determinato hashtag viene categorizzato insieme a tutti gli altri, scritti da utenti diversi, a cui sia stato associato il medesimo hashtag. La funzione è stata incorporata negli ultimissimi anni da altri popolari social network come Instagram, Google+ e Facebook. [...]

è una parola inglese composta da hash, uno dei tanti termini che indicano il simbolo # (in italiano detto cancelletto), e tag, 'etichetta', 'marcatore'. Coniata per la prima volta nel 2007 in un blog statunitense e poi proposta su Twitter dagli stessi utenti, la parola all'inizio indicava solamente il simbolo del cancelletto impiegato per creare chiavi di ricerca, non la chiave di ricerca stessa, e questo significato si è mantenuto nella terminologia ufficiale di Twitter: "il

simbolo #, denominato hashtag, viene utilizzato per contrassegnare parole o argomenti chiave in un Tweet. È stato concepito dagli utenti Twitter come metodo per categorizzare i messaggi”. Nell’uso comune, tuttavia, il significato si è esteso a tutta la sequenza, che nella terminologia ufficiale è chiamata *hashtagged word*.

In letteratura sono presenti molteplici studi sui processi di adattamento e nuove formazioni di significato nelle pratiche linguistiche all’interno dei nuovi contesti virtuali (Budnik 2019; Caleffi 2015; Zappavigna 2015), orientando le indagini ai meccanismi di associazione delle intenzioni comunicative agli hashtag. Ad esempio, (Caleffi 2015) individua tra gli usi e le funzioni più comuni degli hashtag i seguenti: denominazioni, espressioni di sentimenti ed emozioni personali, hashtag come mezzi comunicativi per sostenere iniziative o eventi, hashtag come veicoli comunicativi di autoironia o di promozione di brand. L’autrice analizza ancora più precisamente la combinazione dei significati a seconda della composizione delle parole contenute all’interno degli hashtag: se si tratta di una sola parola che segue il simbolo ‘#’ le possibili valutazioni sono relative o a un sostantivo singolare che non presenta un valore pragmatico o semantico rilevante, o a un aggettivo che normalmente esprime un sentimento o un commento agendo da meta-commento sull’intero messaggio (Caleffi 2015, 54); quando un hashtag è costituito dalla combinazione di due parole, generalmente questa tipologia è formata da una frase nominale di tipo attributivo (aggettivo + sostantivo) (Caleffi 2015, 55) o sostantivo (aggettivo+sostantivo) e queste espressioni vengono utilizzate per esprimere un sentimento di protesta, incoraggiamento, emozione o per condividere un commento; la struttura di un hashtag formato da tre parole, nel contesto della lingua italiana, ha un significato equivalente a espressioni esclamative, sarcastiche, di avvertimenti o emozioni, mentre gli hashtag composti da quattro parole hanno una funzione di commento o suggerimento.

(Zappavigna 2015) ha esplorato le nuove funzioni linguistiche messe a disposizione dall’uso degli hashtag: indicatori dell’*aboutness* dei testi presenti nei messaggi pubblicati sui social media, veicoli di formazione di tribù virtuali, mezzi comunicativi di costruzione di tematiche emergenti e, non meno importante, tecniche di formazione di ‘meta-discorsi’. È su quest’ultimo aspetto che l’autrice si sofferma in modo più dettagliato indicando come questa combinazione di parole sui social network precedute dal simbolo ‘#’ «support communication that reflexively points back at itself (and is ‘meta’ in this sense). Like traditional metadata, hashtags afford abstraction, enabling information at higher and more complex orders of experiences to be appended to the main content of a post». Infatti, l’hashtag ha, secondo l’autrice, una funzione sincretica di istanze conversazionali tra gli utenti: in questo senso parla di “metadati sociali” creati nel tempo dalle comunità di utenti, le quali restituiscono così un sistema di annotazione descrittiva ai propri messaggi agevolando un processo

meta-discorsivo nella condivisione delle informazioni nonché di aggregazione di contenuti per tematiche.

(Budnik, Gaputina and Boguslavskaya 2019) suddividono le funzioni degli hashtag nelle seguenti tipologie: mezzi di formazione di nuove tematiche e strumenti di navigazione, strumenti di marketing costituendosi come strutture ipertestuali che riescono a focalizzare gli elementi da promuovere ricorrendo a un vocabolario caratterizzato da espressioni superlative e divenendo, pertanto, degli slogan vincenti; strumenti di condivisione dei giudizi e delle emozioni condensando in brevi parole una visione del mondo da diffondere; veicoli di significati concettuali, indicando con ciò la possibilità di attribuire un serie di significati concettuali agli hashtag relativi a determinati eventi, persone o situazioni.

È nella direzione delle suddette funzioni associate agli hashtag che l'analisi del presente studio si orienterà. Nello specifico, verrà verificato l'uso linguistico-veicolare degli hashtag nei tweet estrapolati considerandoli come strutture ipertestuali di condivisione di stati emotivi e giudizi durante i mesi successivi alla prima campagna vaccinale di massa nel corso della pandemia da COVID-19 e i risultati di questa analisi saranno integrati dall'identificazione di espressioni tipizzate di movimenti No-vax.

## 2. Infodemia da COVID-19

La continua trasmissione delle notizie riferite alle misure contrastive per ridurre il livello di infettività della malattia da COVID-19 appare inarginabile, e gli spazi del web in cui vengono condensati questi dati contraddittori e incessantemente pervasivi offrono una dimensione virtuale a chi contrasta con teorie inventive le evidenze scientifiche sull'efficacia dei vaccini. Come riportano (Negri e Tagliani 2021, 45) discutendo della nuova infodemia da COVID-19:

L'effetto è evidente: non sapendo interpretare i dati, ciascuno di noi può giungere a conclusioni errate, e generare fake news. Gli epidemiologi hanno modelli matematici che, se applicati ai dati grezzi, restituiscono una proiezione dell'andamento dell'epidemia e del tasso di letalità; sono modelli che noi non possediamo e non possiamo applicare. L'interpretazione dei dati "fai da te" non è dunque attendibile, e si somma all'alea d'incertezza della proiezione elaborata dai modelli scientifici, che non può restituire un'interpretazione univoca ma solo stimata, essendo il processo ancora in atto e i dati non definitivi. In questo contesto la "infodemia" a cui, nella società postmoderna, siamo esposti, ci porta alla generazione di fake news molto verosimili, non solo perché non sono pensate per ingannare, ma perché sorgono per sovra-interpretazione dei dati e, cosa ancor più diabolica, si basano sui dati veri.

In questo contesto pandemico, seguendo le parole dell'autore, le motivazioni che spingono i gruppi No-vax a rimanere riluttanti nell'accettazione della



campagna vaccinale per combattere la circolazione del COVID-19 si riferiscono principalmente all'incapacità di interpretare correttamente una fitta serie di informazioni che circolano in modo frenetico sul web. Accanto a questo senso di disorientamento informativo, il meccanismo di disillusione innescato nei confronti degli strumenti di difesa vaccinali è altrettanto legato al fatto che questi stessi vengano considerati, come in passato, strumenti che ledono le libertà di scelta individuali (Jolley 2014). A cambiare sono gli spazi pubblici in cui chi propaganda sommosse intellettuali contro l'obbligatorietà vaccinale trova confronti perpetui, ora confinati alle piattaforme web dei social network e dei blog. Infatti, i social network rappresentano il luogo della concretizzazione comunicativa dei movimenti No-vax e di teorie cospirazioniste su *Big Pharma* e sulla dittatura sanitaria (Cossard 2020; Jolley 2014).

### 2.1. *Messaggi tipizzati dei gruppi anti-vaccino*

La gestione e l'acquisizione dei dati utili per gli scopi di ricerca sulla rilevazione di un livello di ricorsività linguistica dei gruppi No-vax sono state impostate su una metodologia di recupero dell'informazione basata sulla consultazione di contenuti provenienti da fonti del web. La documentazione disponibile all'interno delle piattaforme virtuali permette, infatti, una maggiore espandibilità in tempo reale del tessuto informativo oggetto di studio. Il web, seguendo le parole di (Montagnani 2007, 74) che riflette sul lavoro di Baldazzi, è connotato da un forte tratto di transdisciplinarietà e ipertestualità in cui di volta in volta «Si è andati verso una dissoluzione dei supporti, dei quadri scientifici e tassonomici di classificazione, di possibilità di contenimento in un solo luogo dell'informazione, del carattere enciclopedico del sapere, dirigendosi verso una pluralità e equivocità di quadri ontologici instabili». Ed è in questo contesto virtuale altamente multiforme che il ruolo del documentalista si è trasformato in quello che Baldazzi chiama *cybrarian* a supporto dell'estrazione di contenuti informativi da piattaforme web e della loro giusta interpretazione attraverso un processo di "metacognizione".

Per avviare l'indagine oggetto di studio è stato preso in considerazione il social network Twitter. Una maggiore fluidità del linguaggio, intendendo con ciò l'assenza di una struttura nella disposizione degli elementi lessicali nei messaggi, e la sempre crescente rapidità negli aggiornamenti dei messaggi da parte degli utenti sono alcuni dei tratti specifici di questa piattaforma (Bernard et al. 2009; Sankaranarayanan et al. 2009) che la rendono un portale in cui sono presenti contenuti eterogenei di supporto al recupero dell'informazione per poter avviare analisi diacroniche su nuovi usi terminologici. Considerando l'obiettivo di ricerca, ovvero uno studio preliminare dei processi di condivisione di espressioni comunicative dei gruppi No-vax per mezzo dei messaggi pubblicati tramite i tweet, l'estrazione di dati utili è stata effettuata per mezzo

di parole chiave di riferimento per la tematica in oggetto, hashtag. In primo luogo è stato definito il periodo per l'estrazione dei tweet. Sono stati isolati alcuni messaggi pubblicati online tra settembre e ottobre 2021, arco temporale che coincide con i mesi successivi alla prima campagna di vaccinazione, al fine di mostrare quanto il contenuto dei post sia diventato, in seguito alle varie ondate della pandemia da COVID-19 e con il passare dei mesi, ripetitivo e standardizzato, quasi a costituire sia nel mondo virtuale che reale un vocabolario di riferimento contenente delle espressioni fisse riutilizzabili per il rafforzamento dell'identificazione comunitaria No-vax nel web (Furini 2021).

Una volta delimitato il range temporale, il secondo passaggio ha previsto l'identificazione di un set di hashtag per il recupero dei tweet inerenti ai suddetti scopi di ricerca al fine di effettuare un'indagine conoscitiva dei principali usi e convenzioni semantiche legati al contenuto dei messaggi estratti. Nel dettaglio, gli hashtag utilizzati per gli scopi di questa ricerca sono stati estratti dallo studio su lavori presenti in letteratura per l'identificazione delle principali parole chiave associate ai movimenti anti-vaccino relativi alla pandemia da COVID-19 (Di Giovanni et al. 2021; Monaci 2021; Quintana et al. 2022). Attraverso ognuno di questi hashtag è stato possibile ottenere un'estrazione di un elenco di tweet corrispondenti.

### 3. Metodologia

La metodologia con cui si è avviata l'analisi semantica dei messaggi estrapolati dalla ricerca per hashtag sulla piattaforma Twitter, identificativi dell'appartenenza di utenti a gruppi anti-vaccino, è stata articolata in quattro fasi, così come mostrato dal seguente schema (Fig. 2):

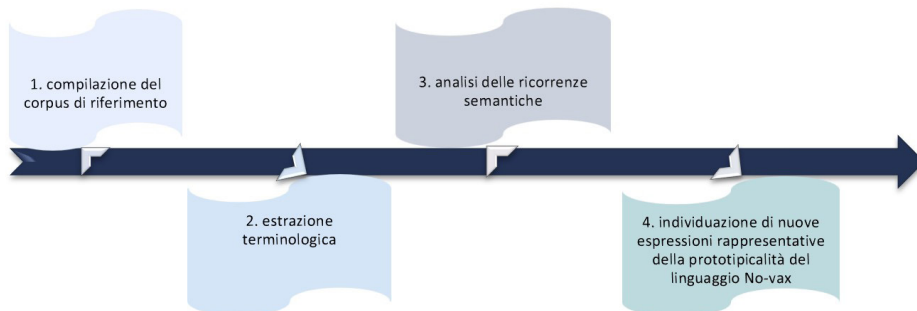


Figura 2: Fasi metodologiche.

### *Compilazione del corpus*

Una volta stabiliti i parametri di ricerca nella prospettiva di avviare uno studio conoscitivo delle modalità espressive di determinati tipi di utenti, la prima fase dello studio si è incentrata sulla realizzazione del corpus di partenza.

Nel processo di estrazione della conoscenza, per ottenere i migliori risultati è importante comprendere al meglio il dominio applicativo di riferimento ed individuare chiaramente gli obiettivi che si intendono raggiungere. [...] Il secondo passo del processo consiste nella selezione dei dati su cui sarà attivato il processo di scoperta. (Lancia e Lapicciarella 2008, 281)

La definizione di un dominio applicativo è, infatti, una delle precondizioni necessarie per una corretta individuazione della popolazione di riferimento con cui far combaciare gli obiettivi di ricerca sulla rappresentazione della conoscenza di settore (Biber 1993).

Un corpus è una raccolta di testi o, in alcuni casi, di parti di testi che vengono scelti in base a criteri prestabiliti perché siano rappresentativi, per quanto possibile, di una lingua, di una varietà linguistica, di un genere, di un dominio di discorso, di un autore o di un argomento (Bowker e Pearson 2002) che possono essere analizzati in un particolare momento o in un certo periodo di tempo (analisi sincronica o diacronica di un corpus) (Torsello 2009, 26).

Le tecniche di compilazione dei corpora seguono alcuni punti nodali che portano alla collazione di testi di riferimento inerenti alla tematica di studio. (Pearson 1998) identifica una serie di requisiti che i testi contenuti nei corpora dovrebbero possedere per configurarsi come strutture di conoscenza aderenti ai campi di studio di cui si cerca di investigare la terminologia:

- **attualità:** bisognerebbe privilegiare testi che possano riflettere l'informazione corrente del settore di studio in modo da impostare un asset documentale in grado di fornire in output informazioni concettuali aggiornate;
- **eterogeneità autoriale:** è preferibile inserire nei corpora di partenza testi scritti da esperti provenienti da diversi settori ai fini di un potenziamento della varietà terminologica;
- **originalità:** i testi non dovrebbero essere traduzioni in modo da garantire l'originalità della terminologia;
- **completezza:** i testi non devono essere frammentari in quanto si rischierebbe di perdere un certo numero di informazioni utili ai fini di un'analisi terminologica completa.

Per (Barrière 2006) la costruzione di un corpus rappresenta il primo passaggio per la realizzazione di una base di conoscenza terminologica, i.e., *terminolo-*

*logical knowledge bases (TKBs)*, da cui verranno sviluppate risorse semantiche orientate all'organizzazione e rappresentazione della conoscenza.

Nel caso specifico, il recupero dell'informazione relativa alle modalità espressive di gruppi No-vax è stato impostato sull'estrazione dei contenuti testuali presenti nella piattaforma di microblogging Twitter che ha rappresentato la fonte principale per la configurazione del corpus di partenza. Nel dettaglio, i testi presenti nel corpus sono tweet pubblicati su questa piattaforma, estratti utilizzando un accesso tramite API, inserendo come chiavi di ricerca nelle queries un gruppo di hashtag, ognuno dei quali ha generato una lista di messaggi temporalmente circoscritti ai mesi di settembre e ottobre del 2021 divenendo i documenti costitutivi del corpus di tweet da analizzare per avviare un'indagine sulla terminologia dei gruppi No-vax. In questo modo il set documentale finale è risultato composto da otto documenti che racchiudono il raggruppamento dei post relativi agli hashtag ricercati così come mostrato dal seguente schema (Tab. 1).

hashtag	n. tweet	documenti	Tot
#DittaturaSanitaria	4.625	Doc1	21604 tweet e 8 documenti
#nomask	250	Doc2	
#nogreenpassobbligatorio	11.871	Doc3	
#NoVaccinoObbligatorio	575	Doc4	
#novaccinoCOVID	34	Doc5	
#novax	3196	Doc6	
#nocavie	50	Doc7	
#noobbligovaccinale	1.013	Doc8	

Tabella 1: Numero di tweet per hashtag.

In totale, il corpus compilato sulla base di un raggruppamento di documenti per hashtag, concettualmente correlati ai principi guida dei movimenti anti-vaccino, è composto da 21.604 tweet per un totale di 81.023 token<sup>6</sup>.

I risultati ottenuti sono parte integrante di un'analisi terminologica realizzata attraverso il processamento automatico dei testi, in questo caso i tweet estrapolati tramite API, eseguito da software di estrazione terminologica pre-addestrati come Text To Knowledge (T2K) (Dell'Orletta et al. 2014), SketchEngine (Kilgarrieff 2014) e VoyantTools (Miller 2018), al fine di poter rilevare le informazioni tipizzate sugli usi terminologici più frequenti, in questo caso di contrasto alle misure di prevenzione alla propagazione dell'infezione da COVID-19. Questi tool consentono di disporre al termine della fase di

<sup>6</sup> Totale delle occorrenze nel corpus di partenza calcolato con il tool di analisi dei corpora T2K.

acquisizione dei documenti e della loro analisi linguistica di una serie di misure statistiche che forniscono informazioni sul grado di varietà terminologica. In particolare, T2K riporta le seguenti informazioni in relazione al corpus in input, in questo caso costituito da otto documenti, ciascuno dei quali corrispondente al gruppo dei tweet estratti in base all'hashtag:

Documenti	Totale token	Totale di lemmi diversi	Densità Lessicale	Rapporto tipo/unità
8	81.023	7.802	0,621	0,096

Sulla base della definizione della misura Type Token Ratio (TTR):

il valore percentuale che indica la varietà di parole diverse contenute in un testo. Più alto è il valore percentuale maggiore è la varietà del vocabolario usato in un testo. Il TTR è calcolato dividendo il numero delle parole diverse usate in un testo (types) per il numero delle parole complessive di quel testo (tokens) moltiplicato per 100<sup>7</sup>.

si evince come il risultato ottenuto attraverso l'analisi del profilo lessicale mediante il tool T2K non raggiunga un valore percentuale alto, e questo è un dato che può confermare la ripetitività dal punto di vista di usi del linguaggio all'interno dei documenti del corpus in input, ovvero i tweet dei gruppi anti-vaccino.

### *Estrazione terminologica*

La seconda fase è mirata all'isolamento dei termini di riferimento nel corpus di partenza ai fini di un'analisi semantica che possa far emergere delle osservazioni circa le modalità espressive dei suddetti gruppi sociali. Questa fase è successiva alla definizione del corpus per l'elaborazione testuale e si riferisce alla configurazione di un sistema di processamento terminologico all'interno dei testi presi in esame. In particolare, si è proceduto con l'estrazione dei termini rappresentativi dei testi contenuti nel corpus di partenza ai fini di portare avanti uno studio esplorativo delle sequenze espressive più ricorrenti nei messaggi No-vax. Per l'esecuzione di processi analitici terminologici ci si avvale generalmente del supporto di software semi-automatici (Loginova et al. 2012) sviluppati per l'implementazione di tecniche di analisi linguistico-statistiche che mirano a rilevare quali siano gli usi terminologici più frequenti ed emergenti e, al contempo, fare inferenza su quali siano le informazioni tipizzate entro contesti linguistici specifici.

<sup>7</sup> Paisà, <https://www.corpusitaliano.it/it/help/readability.html> (ultima consultazione: 26/05/2022).

Dopo una prima valutazione di sistemi di analisi terminologica esistenti<sup>8</sup> il processo di estrazione dei dati è stato eseguito attraverso i software T2K e SketchEngine, tool di analisi dei corpora utilizzati nel campo della lessicografia, il secondo dei quali più flessibile nelle fasi di inclusione di *stopwords* esterne e per la possibilità di ottenere una rappresentazione grafica delle co-occorrenze morfosintattiche dei lemmi utile ai fini delle successive valutazioni semantiche. Questo passaggio ha previsto l'applicazione di alcuni parametri statistici di partenza al fine di ottenere in output una lista di termini rappresentativi che potessero garantire un buon livello quantitativo di copertura semantica rispetto alle informazioni contenute nei post oggetto di studio. Ai fini di una corretta esecuzione di analisi testuale è stata effettuata una pulitura preliminare dei documenti attraverso la rimozione di alcuni dati non essenziali come simboli (più precisamente, emoticon) e riferimenti ipertestuali (Gupta and Rattikorn 2018, 341) e un elenco di *stopwords* non funzionali ai processi di recupero delle informazioni di contesto. Il numero di unità linguistiche singole estratte è stato pari a 11.988, mentre il numero di unità linguistiche composte pari a 15.840.

Nella Tabella 2 si riportano i primi risultati ottenuti in seguito all'analisi linguistico-statistica del corpus di partenza dal punto di vista dell'individuazione dei termini più rappresentativi, eseguita dal software T2K applicando una misurazione di contrasto con un secondo corpus al fine di individuare il livello di specificità terminologica. Il contrasto è stato effettuato attraverso l'attivazione di un confronto tra il corpus di partenza costituito dai tweet estratti e un corpus in lingua italiana giornalistico inserito come documentazione direttamente all'interno del tool T2K (Bonin et al. 2010). L'approccio contrastivo descritto da (Bonin et al. 2012, 209) prevede l'identificazione di unità terminologiche rilevanti ai fini di un'indagine conoscitiva dell'informazione specialistica in corpora di dominio attraverso un confronto tra raggruppamenti documentali che presentano caratteristiche differenti nella distribuzione delle unità terminologiche rappresentative:

L'estrazione di unità terminologiche monorematiche e polirematiche è cioè condotta a partire dal confronto della distribuzione delle unità terminologiche monorematiche e polirematiche nel corpus di acquisizione rispetto a un corpus di riferimento (detto anche 'corpus di contrasto'). In questo modo, la lista finale di unità terminologiche estratte conterrà quelle unità che sono maggiormente rilevanti nel corpus di acquisizione rispetto (ovvero 'per contrasto') al corpus di riferimento. A questo scopo sono state sviluppate una serie di metodologie in grado di computare la misura della diversa rilevanza di unità terminologiche all'interno dei due corpora che vengono confrontati. La possibilità di discriminare termini e non-termini è così empiricamente realizzata sulla base di un'analisi 'contrastiva' della loro distribuzione in un corpus di

<sup>8</sup> Si vedano i tool Sketch Engine, Text to Knowledge (T2K), Voyant Tools, Wordsmith.

dominio (il corpus di acquisizione) rispetto a un corpus rappresentativo della lingua comune (usato come 'corpus di contrasto').

estrazione terminologica		applicazione corpus di contrasto	
termini	frequenza	termini	frequenza
#NoGreenPassObbligatorio	308	#NoGreenPassObbligatorio	308
cose	272	#quartarepubblica	176
vaccino	213	#greenpass	129
massa	184	conto corrente bancario	174
conto	182	#novax	63
soldi	181	#novaccinoobbligatorio	55
#quartarepubblica	176	#vaccino	51
conto corrente bancario	174	#dittatura	45
#greenpass	129	#vaccini	43
persone	110	Vaccinati	40
libertà	102	RT	37
obbligo	89	Covid	36
dose	78	#nogreenpass	28
anni	76	#noobbligovaccinale	27
governo	75	#nogreenpassobbligatorio	25
diritto	74	#BASTADITTATURA	25
pass	64	#libertà	22
#novax	63	Pass	64
emergenza	57	#BASTADITTATURA	19
scuola	57	#DITTATURASANITARIA	19
parte	55	Pandemia	19
#novaccinoobbligatorio	55	#NoGreenPassObbligatorio	18
#vaccino	51	#NoVaccinoObbligatorio	18
vita	48	Grazie	18
volta	48	Cure	30
tampone	47	reazioni avverse	17
giorno	47	#nessunacorrelazione	15
#dittatura	45	Greenpass	14
paese	45	Mascherina	31
#vaccini	43	dittatura sanitaria	13
vaccinati	40	#Dittatura	13
lavoro	40	#Draghi	13
		#draghi	13
		@borghi_claudio	13

figli	40	#greenpass.	13
diritto costituzionale	24	#Fuoridalcoro	13
RT	37	Vaccino	213
via	37	#NoGreenPassObbligatorio #noobbligovaccinale	12
virus	37	#NoGreenPassObbligatorio #NoObbligoVaccinale	12
legge	36	#drittoeroveschio	12
covid	36	#governodeipeggiori	12
rischio	36	#libertadiscegliere	12
casa	35	#dimartedi	12
decessi	34	#DittaturaSanitaria	12
tempo	34	#mattarelladimettiti	12
lavoratori	34	@matteosalvinimi	12
problema	33	Contagi	25
vaccinazione	32	immunità di gregge	11
verità	32	#staseraitalia	11
paura	32	#terzadose	11
mascherina	31	#novax #NoGreenPassObbligatorio	10
punto	31	#nogreepass	10
caso	31	#vaccinoobbligatorio	10
#BASTADITTATURA #DITTATURASANITARIA	19	Mail	10

Tabella 2. Estrazione terminologica dai tweet elaborata con T2K e contrasto con un corpus generale.

Da questa tabella si possono evincere alcune osservazioni. In primo luogo, dal momento che il corpus di partenza risulta costituito da otto documenti contenenti i tweet estratti tramite ricerca per hashtag, i termini che sono considerati più rappresentativi applicando la misura di contrasto sono gli hashtag utilizzati dagli utenti per esprimere le proprie posizioni su Twitter. Come già osservato nei precedenti paragrafi, le funzioni che possiedono gli hashtag nei contesti comunicativi dei social network sono riferite soprattutto alla possibilità di veicolare tramite meta-discorsi pratiche discorsive degli utenti. Infatti, in rosso si può notare il cambiamento di posizione dei termini a seconda della lista di appartenenza e del livello di specificità acquisito nel momento in cui si confronta la terminologia raccolta nel corpus di riferimento con uno generale. e.g., come *greenpass*, *RT*, *novaccinoobbligatorio*, viceversa, termini come *mascherine* e *reazioni avverse* risultano rappresentativi in un corpus italiano



generale rispetto alla frequenza osservata nel raggruppamento documentale specifico della terminologia No-vax su Twitter, ma sono termini che vengono riportati in posizioni superiori per la rilevanza che assumono nel corpus di partenza.

Questa prima fase di estrazione terminologica risulta utile ai fini di un recupero di espressioni che presentano un alto grado di specificità rispetto alle tematiche trattate in un contesto d'uso oggetto di studio e che possono essere prese in considerazione per ulteriori indagini semantiche mirate alla scoperta di strutture linguistiche prefissate ricorsive di questi gruppi oppositivi.

### *Analisi delle ricorrenze semantiche*

Gli elementi informativi del linguaggio sono inseriti in

sistemi di relazioni di ordine lineare, di incassamento gerarchico tra sintagmi e di dipendenze a lungo raggio che nel loro insieme vengono a comporre la struttura sintattica delle frasi. Queste ultime si articolano infine in complesse strutture retoriche che formano la tessitura del discorso. La struttura linguistica è la porta di accesso al contenuto del testo (Lenci, Montemagni e Pirrelli 2005, 241).

I testi che hanno costituito il gruppo documentale di partenza presentavano al loro interno molteplici informazioni linguistiche che hanno fatto emergere dati interessanti per l'avvio della terza fase metodologica. L'analisi dei termini rappresentativi contenuti all'interno del complesso documentale dei tweet ha permesso di mettere in evidenza una struttura di concatenazioni semantiche prototipiche per la rappresentazione concettuale dell'ideologia dissidente dei gruppi No-vax. Infatti, all'interno dei messaggi analizzati come testo non-strutturato, ad esempio, la maggior parte delle volte il termine *vaccino* compare associato a verbi e aggettivi negativi, quali 'inutile', 'pseudo', 'mortifero', 'morire', 'funzionicchiano', 'alterare', come si può osservare dalle seguenti figure (Fig. 3, Fig. 4, Fig. 5).



Figura 3: Aggettivi correlati al termine “vaccino”.

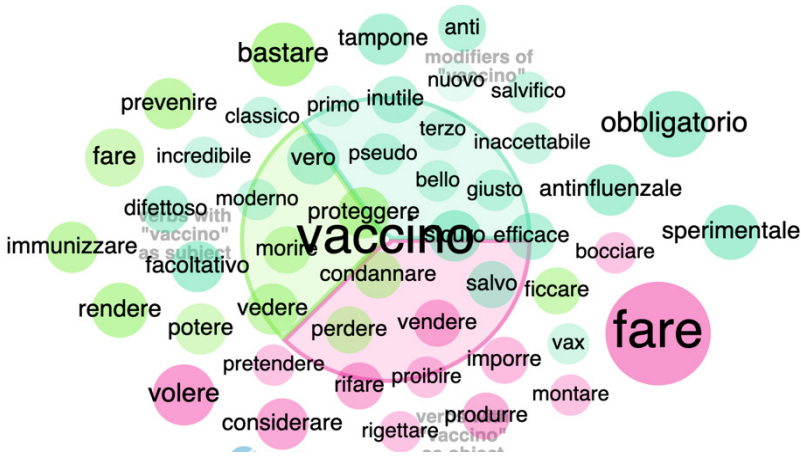


Figura 4: Aggettivi e verbi correlati al termine “vaccino”.

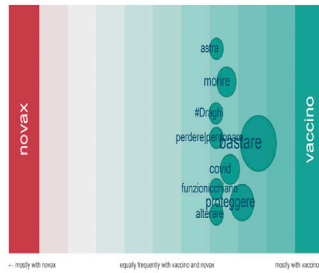


Figura 5: Verbi correlati al termine “vaccino”.

Si ravvisano altresì all'interno dei tweet analizzati associazioni semantiche negative relative alla descrizione dei gruppi No-vax – per il fatto che la ricerca dei tweet è stata effettuata attraverso l'impiego di parole chiave/hashtag che potevano comprendere diversi tipi di contenuti potenzialmente riferibili a contesti differenti – i.e., accostamenti con aggettivi aventi polarità negativa, come 'terrorista', 'pericoloso' o 'minaccioso' (Fig. 6). A loro volta, questi stessi gruppi oppositivi tendono a considerare le misure attuative di emergenza per combattere la diffusione del COVID-19 in modo altamente dispregiativo: infatti, il greenpass presenta diverse co-occorrenze con aggettivi a polarità negativa come 'falso', 'obbligatorio', 'incostituzionale' e viene considerato come uno strumento 'abominio' contro cui 'protestare' (Fig. 7); la mascherina viene associata a una visione ostile, è infatti percepita come uno 'scandalo' (Fig. 8), così come la gestione della situazione pandemica attuale recepita come una 'dittatura sanitaria' 'imposta' (Fig. 9).

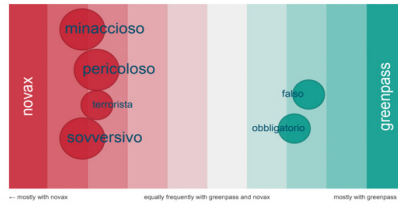


Figura 6: Confronto tra le co-occorrenze dei termini “novax” e “greenpass”.

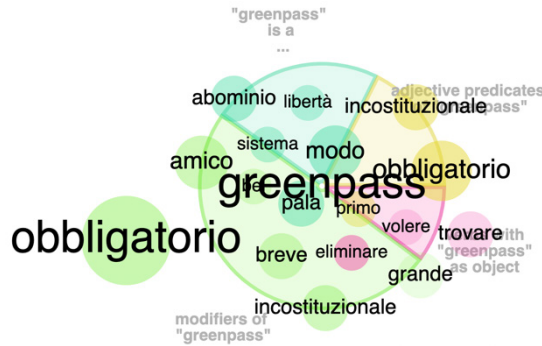


Figura 7: Concatenazioni semantiche riferite al termine “green pass”.



Figura 8: Esempio di relazione associativa per il termine “mascherina”.



Figura 9: Rete semantica per il termine “dittatura sanitaria”.

### *Identificazione di nuove catene semantiche del linguaggio No-vax*

Twitter è una piattaforma di microblogging (Qui et al. 2012, 710-18) grazie alla quale si riescono a seguire in modo costante processi evolutivi dell'uso del linguaggio da parte degli utenti attraverso l'osservazione di nuove formazioni terminologiche in hashtag (Mehmood 2013). Questi ultimi sono caratterizzati dal fatto di non essere dipendenti da un contesto specifico, di contenere una serie di slang e abbreviazioni e espressioni non standardizzate (Gutba 2018). Per la rilevazione di ulteriori sequenze semantiche attribuibili a un uso ricorsivo e stereotipato da parte di utenti Twitter No-vax l'ultima fase del presente studio ha riguardato l'isolamento di tutti i termini preceduti dal simbolo '#', che tipicamente precede la presenza di hashtag sui social network, all'interno della lista terminologica ottenuta in seguito al processamento dei testi contenuti nel corpus di partenza. Da questi è stato possibile identificare nuovi trend semantici (Lau, Collier and Baldwin 2012) associabili alle modalità espressive dei gruppi No-vax.

Pertanto, attraverso l'individuazione delle entità hashtag all'interno dei risultati dell'estrazione terminologica, rappresentativa del corpus documentale dei tweet presi in esame, è stato possibile verificare la presenza di associazioni di elementi semantici ricorsivi all'interno dei messaggi pubblicati da profili Twitter contrari alle cure vaccinali nel contesto della pandemia da COVID-19.

Di seguito viene presentata una lista esemplificativa delle co-occorrenze più frequenti rintracciate nella lista dei termini riferiti al contenuto dei tweet dei gruppi No-vax oggetto di studio:

- #fascisti
  - #fascisti #nazisti
- #guerrieri
  - #guerrieri #novax
- #nessunacorrelazione
  - #nessunacorrelazione #vaccinibrodaglia
- #provax
  - #provax contro #novax
- #terroristi
  - #terroristi #vaccino
- #terzadose
  - #terzadose di #Terrorismo
- #vaccino
  - #vaccino #giocopubblico
  - #vaccino #novax #complotto
- #vaccinoobbligatorio
  - #depopolazione

#nogreenpassobbligatorio

- #governodeipeggiori
- #NoVaccinePassport

#pandemia

- #falsapandemia #neonazisti
- #variantisenzafine

Si può evincere un uso terminologico identificativo dell'appartenenza ideologica di profili No-vax in cui si ricorre spesso a formule stereotipate negative nei confronti del controllo della situazione pandemica attuale considerando le soluzioni adottate come 'complotto' o forme di 'terrorismo'.

#### 4. Conclusioni

Dai risultati ottenuti attraverso l'analisi del corpus di partenza costituito dal raggruppamento di tweet estratti grazie all'uso di determinati hashtag in riferimento ai movimenti antivaccinismo, è stato possibile osservare l'uso di nuove modalità espressive adottate dai profili iscritti a piattaforme sociali per esprimere proprie posizioni. Il processo di costruzione fraseologica, caratterizzata dall'uso di termini specifici al contesto pandemico di riferimento, entro i confini virtuali trova nell'utilizzo massivo degli hashtag, come strutture meta discorsive, la configurazione di nuovi modelli espressivi. In particolare, in questo studio preliminare sono state analizzate una serie di concatenazioni espressive estrapolate dal corpus di partenza e isolate tramite l'applicazione di misure statistiche in seguito al processamento terminologico, molte volte caratterizzate dalla presenza di hashtag come strumenti di supporto alla rappresentazione sintetica di concetti più complessi diffusi da gruppi No-vax all'interno della piattaforma di microblogging Twitter. L'obiettivo è stato basato sull'intenzione di isolare un set di messaggi ricorrenti utilizzati per manifestare una forma di appartenenza ideologica e atti a diventare un vocabolario di espressioni linguistiche tipizzate dei suddetti gruppi sociali. Infatti, automatizzando dei processi di identificazione di associazioni semantiche su un gruppo di documenti rappresentativi di tematiche specifiche è possibile osservare e isolare delle costruzioni linguistiche ricorsive. Lo studio della ricorsività di strutture morfo-sintattiche presenti all'interno dei tweet oggetto di questo studio ha permesso di verificare come l'uso congiunto di alcuni termini che rientrano nell'uso del linguaggio proprio di un gruppo di utenti specifico abbia supportato la definizione di espressioni regolarmente usate nel linguaggio comune delle comunità virtuali No-vax per esprimere determinati concetti e opinioni. Queste ultime divengono nel corso del tempo delle espressioni manifesto prefabbricate da utilizzare in modo da rispecchiare un'ideologia di appartenenza. Nello

specifico, costituiscono strutture di mediazione e interpretazione semantica orientate all'individuazione di informazioni emblematiche e alla loro integrazione con schemi di pensiero omogenei. Il linguaggio dei social network offre la possibilità di seguire a ritmi costanti l'inserimento di nuove parole che nel tempo formano veicoli espressivi di contenuti semantici per la descrizione di nuove realtà come quella pandemica, diventando altamente emblematiche di determinati contesti descrittivi. Appare pertanto utile disporre di analisi e metodologie terminologiche come quelle qui presentate che possano guidare alla scoperta di espressioni linguistiche contrassegnate da valori iterativi per l'assimilazione di trend costitutivi di fenomeni storici.

### Riferimenti bibliografici

- Baldazzi, Anna. 2004. "Documentazione: saperi e professioni in evoluzione." *Rassegna bibliografica infanzia e Adolescenza* 5, no. 2: 5–25.
- Barrière, Caroline. 2006. "Semi-automatic corpus construction from informative texts." In *Lexicography, Terminology and Translation: Text-Based Studies in honour of Ingrid Meyer*, edited by Lynne Bowker. University of Ottawa Press.
- Biber, Douglas. 1993. "Representativeness in Corpus Design." In *Literary and Linguistic Computing* 8, no. 4: 243–57.
- Bonin, Francesca, Felice Dell'Orletta, Giulia Venturi, and Simonetta Montemagni et al. 2010. "A contrastive approach to multi-word term extraction from domain corpora." In *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC 2010)*.
- Bonin, Francesca, Felice Dell'Orletta, Giulia Venturi, and Simonetta Montemagni. 2012. "Lessico settoriale e lessico comune dell'estrazione di terminologia specialistica da corpora di dominio." In *XLIV congresso internazionale di studi della società di linguistica italiana*, 207–20, Viterbo, 27–29 settembre 2010.
- Budnik Ekaterina, Violetta Gaputina, and Vera Boguslavskaya. 2019. "Dynamic of hashtag functions development in new media: Hashtag as an identificational mark of digital communication in social networks." In *Proceedings of the XI International Scientific Conference Communicative Strategies of the Information Society (CSIS'2019)*. Association for Computing Machinery, New York, NY, USA, Article 28, 1–5. <https://doi.org/10.1145/3373722.3373795>.
- Caleffi, Paola-Maria. 2015. "The 'hashtag': A new word or a new rule?" *SKASE journal of theoretical linguistics* 12, no. 2.

- Cossard, Alessandro, Gianmarco De Francisci Morales, Kyriaki Kalimeri, Yelena Mejova, Daniela Paolotti, and Michele Starnini. 2020. "Falling into the Echo Chamber: The Italian Vaccination Debate on Twitter". In *Proceedings of the International AAAI Conference on Web and Social Media* 14, no. 1: 130–40. <https://ojs.aaai.org/index.php/ICWSM/article/view/7285>.
- Dell'Orletta, Felice, Giulia Venturi, Andrea Cimino, and Simonetta Montemagni. 2014. "T2K: A system for automatically extracting and organizing knowledge from texts." In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland. European Language Resources Association (ELRA).
- Di Giovanni, Marco, Lorenzo Corti, Silvio Pavanetto, Francesco Pierri, Andrea Tocchetti, and Marco Brambilla. 2021. "A Content-based Approach for the Analysis and Classification of Vaccine-related Stances on Twitter: The Italian Scenario." In *Information Credibility and Alternative Realities in Troubled Democracies @ ICWSM 2021*, 1–6.
- Furini, Marco and Gabriele Menegoni. 2018. "Public Health and Social Media: Language Analysis of Vaccine Conversations." In *2018 International Workshop on Social Sensing (SocialSens)*, 50–55. 10.1109/SocialSens.2018.00022.
- Furini, Marco. 2021. "Identifying the features of ProVax and NoVax groups from social media conversations", *Computers in Human Behavior*, 120, 106751. <https://doi.org/10.1016/j.chb.2021.106751>.
- Gupta, Vibhuti, and Hewett Rattikorn. 2020. "Real-Time Tweet Analytics Using Hybrid Hashtags on Twitter Big Data Streams." *Information*, 11(7), 341. <https://doi.org/10.3390/info11070341>.
- Jansen, Bernard J., Mimi Zhang, 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, no. 11: 2169-88.
- Kalima, Juha, and Mikael Matti. 2013. "The effect of topic on word formation and the frequency of neologism use on Internet forums." Master's Thesis, University of Jyväskylä.
- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. "The Sketch Engine: ten years on." In *Lexicography ASIALEX* 1, 7–36. <https://doi.org/10.1007/s40607-014-0009-9>.
- Knippenberg, Daan van. 2011. "Embodying who we are: Leader group prototypicality and leadership effectiveness." *The Leadership Quarterly* 22, no. 6: 1078-91. <https://doi.org/10.1016/j.leaqua.2011.09.004>.

- Knippenberg, Daan van, Nathalie Lossie and Henk A. M. Wilke. 1994. "In-group prototypicality and persuasion: Determinants of heuristic and systematic message processing." *British Journal of Social Psychology* 33 , n. 3 (settembre): 289-300.
- Lancia, Maurizio, e Andrea Lapicciarella. 2008. "Data Mining e Text Mining". In *Dal Documento all'Informazione*, a cura di Roberto Guarasci, Milano: ITER Srl.
- Lau Jey Han, Nigel Collier, and Timothy Baldwin. 2012. "On-line Trend Analysis with Topic Models: #twitter Trends Detection Topic Model Online." In *Proceedings of COLING 2012*, 1519-34, Mumbai, India. The COLING 2012 Organizing Committee.
- Lenci, Alessandro, Simonetta Montemagni, e Vito Pirrelli. 2005. *Testo e computer. Elementi di linguistica computazionale*. Roma: Carrocci Editore.
- Loginova Clouet E., Elizaveta, Anita Gojun, Helena Blancafort, Marie Guegan, Tatiana Gornostay, and Ulrich Heid. 2012. "Reference Lists for the Evaluation of Term Extraction Tools." In *Terminology and Knowledge Engineering Conference (TKE)*, Madrid, Spain.
- Mehmood, Rizwan Hermann Maurer, and Muhammad Tanveer Afzal. 2013. "Knowledge discovery in hashtags#." In 2013 *IEEE 9th International Conference on Emerging Technologies (ICET)*, 1-6, doi: 10.1109/ICET.2013.6743538.
- Miller, Alissa. 2018. "Text mining digital humanities projects: Assessing content analysis capabilities of voyant tools." *Journal of Web Librarianship* 12(3):169-197.
- Monaci, Sara. 2021. "The Pandemic of Conspiracies in the COVID-19 Age: How Twitter Reinforces Online Infodemic." *Online Journal of Communication and Media Technologies*, 11(4): e202120.
- Montagnani, Daniele. 2007. "Anna Baldazzi Documentazione: Saperi e professioni in evoluzione." *Biblioteche oggi* 3: 74-7.
- Negri, Antonella, e Roberto Tagliani, a cura di. 2021. *Fictio, falso, fake: sul buon uso della filologia*. Milano: Ledizioni Ledipublishing.
- Pearson, Jennifer. 1998. "Terms in context." *Studies in Corpus Linguistics* 1, no. 12: 246.
- Pedroni, Marco. 2020. "Narrazioni virali. Decostruire (e ricostruire) il racconto dell'emergenza coronavirus." *Mediascapes Journal* 15: 24-43.
- Quintana, Ignacio Ojea, Ignacio Ojea, Marc Cheong, Mark Alfano, Ritsart Reimann, Colin Klein. 2022. "Automated clustering of COVID-19 anti-vaccine discourse on Twitter." arXiv <https://doi.org/10.48550/arXiv.2203.01549>.



- Qiu, Lin, Han Lin, Jonathan E. Ramsay, and Fang Yang. 2012. "You are what you tweet: Personality expression and perception on Twitter." *Journal of Research in Personality* 46, n. 6 (december): 710-18. <https://doi.org/10.1016/j.jrp.2012.08.008>.
- Rovetta, Alessandro, e Lucia Castaldo. 2022. *Siamo sicuri di aver compreso appieno cosa sia un'infodemia? Una prospettiva globale sui problemi infodemiologici*. 10.6084/m9.figshare.16863439.
- Sankaranarayanan, Jagan, Hanan Samet, Benjamin E. Teitler, Michael D. Lieberman, and Jon Sperling. 2009. "Twitterstand: News in tweets." In the *17th ACM SIGSPATIAL International Conference*. New York, New York, USA: Association for Computing Machinery Press. <https://doi.org/10.1145/1653771.1653781>.
- Torsello, Carol Taylor, Katherine Ackerley, Erik Castello, Fiona Dalziel, Sara Gesuato, Francesca Helm, e Lisa Lena Opas-Hänninen. 2009. "Lavorare con un corpus, in EXPLICS – (EXPL)ointing (I)nternet (C)ase Studies and (S)imulation Templates for Language Teaching and Learning." In *L'uso di studi di un caso e di simulazioni su internet per l'insegnamento e l'apprendimento delle lingue*, a cura di Johann Fischer Maria, Teresa Musacchio, Alison Standring, 25-34.
- Valensise, Carlo M., et al. 2021. "Lack of evidence for correlation between COVID-19 infodemic and vaccine acceptance." arXiv preprint arXiv:2107.07946.
- Whitehead, Martin, Niall Taylor, Alex Gough, Danny Chambers, Mike Jessop, and Phil Hyde. 2019. "The anti-vax phenomenon." *The Veterinary record*. 184, no. 24 (jun): 744. <https://doi.org/10.1136/vr.l4027>.
- Zappavigna, Michele. 2015. "Searchable talk: The linguistic functions of hashtags." *Social Semiotics* 25, no. 3: 274-91.





# AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

Anno 40

N. 1-2 – gennaio-giugno 2022

## Contributi

STEFANO ALLEGREZZA

*La conservazione degli archivi di posta elettronica: sviluppi recenti e prospettive future*

ANDREA BELLANDI

*Le Risorse Linguistiche nell'era del Web Semantico. Un insieme di servizi informatici per la gestione di lessici e terminologie*

GIORGIA DI MARCANTONIO

*E se l'archivio non rispecchia l'istituto? Pavone e il rispecchiamento: analisi di una bozza preliminare*

GERARDO GIARDIELLO, CHIARA FIORAVANTI,  
FRANCESCO ROMANO, MARIASOLE RINALDI

*Strumenti per il reperimento di informazioni rilevanti in Rete. Applicazioni nel dominio della comunicazione pubblica in materia di immigrazione in Italia*

CLAUDIA LANZA

*Termini e testi nella comunicazione dei gruppi No-vax*

ANNA ROVELLA

*La posta elettronica negli archivi di persona: conservazione e accesso*

GIOVANNI SALUCCI

*Il Blog di qualità. Proposta di un nuovo format di pubblicazione nell'editoria scientifica*

## Note e Rubriche

ROBERTO GUARASCI

*Paul Otlet e Suzanne Briet. Note in margine al libro di Paola Castellucci e Sara Mori*

ERIKA PASCERI

*L'evoluzione del Nuovo Glossario di biblioteconomia e scienza dell'informazione ai tempi dello Smart Working*

CLAUDIO GNOLI

*Come mi vuoi, canonica o rivoluzionaria?*

CLAUDIO GRIMALDI

*L'universo affascinante del colore: le sociocromie*



mundaneum

In copertina

Disegno di Paul Otlet, Collections Mundaneum, centre d'Archives, Mons (Belgique).

ISBN 979-12-5965-163-1



9 791259 651631

ISSN 1121-0095



9 770112 100950