

Pubblicato il: ottobre 2024

©Tutti i diritti riservati. Tutti gli articoli possono essere riprodotti con l'unica condizione di mettere in evidenza che il testo riprodotto è tratto da www.qtimes.it
Registrazione Tribunale di Frosinone N. 564/09 VG

Analyzing Hate speech against Pope Francis on Twitter and Telegram through the Online Speech Hate Spectrum

Analizzare l'Hate speech verso Papa Francesco su Twitter e Telegram attraverso lo Spettro dell'Odio Online

di

Matteo Mancini

matteo.mancini2@unicatt.it

Università Cattolica del Sacro Cuore di Milano

Abstract:

The article examines the integration of machine learning techniques and socio-educational analysis to classify online hate speech targeting Pope Francis, focusing on Twitter and Telegram from January 2022 to April 2024. The aim is to understand how hate is articulated concerning specific themes addressed by the pontificate, employing the "Online Hate Spectrum" model, previously applied to other forms of hate speech such as anti-Romani sentiment, Islamophobia, and antisemitism. On Twitter, the analysis was conducted across eight thematic categories, selecting tweets using machine learning techniques based on keywords and research criteria defined by the Mediavox Observatory of the Catholic University. After identifying monthly peaks in hate speech, a manual classification determined the presence of discriminatory content, analyzing the rhetoric used and specific targets. In parallel, the study on Telegram examined 92 channels disseminating hostile and conspiratorial narratives, mapping content, interactions, and the use of images. The analysis revealed a centralized network of shares and comments, characterized by an informational bubble with few origin points. Channels were classified based on their tone, distinguishing between neutral dissemination, constructive criticism, and provocative attitudes. The results provide a detailed picture of online hate dynamics against Pope Francis, highlighting the association between conversational flows and

©Anicia Editore

QTimes – webmagazine

Anno XVI - n. 4, 2024

www.qtimes.it

doi: 10.14668/QTimes_16427

processes of target selection, and contribute to a better understanding of the strategies needed to effectively counter online hate speech.

Keywords: hate studies, Pope Francis, Twitter, Telegram, Online Hate Spectrum.

Abstract:

L'articolo esamina l'integrazione di tecniche di machine learning e analisi socioeducativa per classificare le espressioni di odio online contro Papa Francesco, focalizzandosi su Twitter e Telegram tra gennaio 2022 e aprile 2024. L'obiettivo è comprendere come l'odio si articoli rispetto a tematiche specifiche affrontate dal pontificato, utilizzando il modello dello "Spettro dell'Odio Online", già applicato ad altre forme di hate speech come l'antiziganismo, l'islamofobia e l'antisemitismo. Su Twitter, l'analisi è stata condotta su otto categorie tematiche, selezionando tweet attraverso tecniche di machine learning basate su parole chiave e criteri di ricerca definiti dall'Osservatorio Mediavox dell'Università Cattolica. Una volta identificati i picchi mensili di incitamento all'odio, una classificazione manuale ha determinato la presenza di contenuti discriminatori, analizzando le retoriche utilizzate e i bersagli specifici. Parallelamente, lo studio su Telegram ha esaminato 92 canali che diffondono narrazioni ostili e complottiste, mappando contenuti, interazioni e l'uso di immagini. L'analisi ha rivelato una rete centralizzata di condivisioni e commenti, caratterizzata da una bolla informativa con pochi centri di origine. I canali sono stati classificati in base al loro tono, distinguendo tra divulgazione neutrale, critica costruttiva e atteggiamenti provocatori. I risultati forniscono un quadro dettagliato delle dinamiche d'odio online contro Papa Francesco, evidenziando l'associazione tra flussi conversazionali e processi di elezione a bersaglio e contribuiscono a comprendere meglio le strategie necessarie per contrastare efficacemente l'hate speech.

Parole chiave: odio online, hate studies, Papa Francesco, Twitter, Telegram, Spettro dell'Odio Online.

Introduzione

Questo articolo si colloca negli *Hate Studies*, quel campo di ricerca interdisciplinare internazionale (Schweppe - Perry, 2022; James - McBride, 2022; Guillén-Nieto, 2023) che, attraverso ricerche e progetti, ha da anni unito gli interessi di ricercatori, politici, esperti di comunicazione, attivisti dei diritti umani, responsabili di ong; da alcuni anni gli studi in tale ambito si sono focalizzati sul web sociale, ossia lo spazio caratterizzato dalla coautorialità degli spettatori (*prosumer*) nella produzione e fruizione di flussi comunicativi e sociali, chiedendosi quali aspetti di novità e quali di continuità il digitale comporta rispetto ai processi di elezione a bersaglio.

Sia dall'ambito educativo delle istituzioni europee (De Latour, Del Felice, Ettema, 2017; Keen, Georgescu, 2016), sia dalla più recente riflessione pedagogica italiana (Pasta, 2018; 2023; Santerini, 2019; 2021; Annacontini, 2019; Fiorucci, 2019; Crescenza, 2024) è maturata la convinzione che una maggiore conoscenza delle modalità con cui si manifesti l'odio – inteso come “categoria ambigua ma utile” (Pasta, 2024a, p. 113) – sia necessario, specialmente nelle *flame wars* del web sociale, poiché online forme diverse di odio e nuclei tematici diversi necessitano risposte educative differenti che sappiano conoscere e successivamente co-costruire modelli di intervento calati nei contesti di vita quotidiana, privilegiando reti comunitarie che propongano percorsi di welfare socioculturale e di

©Anicia Editore

QTimes – webmagazine

Anno XVI - n. 4, 2024

www.qtimes.it

doi: 10.14668/QTimes_16427

crescita collettiva (Santerini, 2023; Pasta, Santerini, 2021; Marangi, 2021; Mancini, 2021). Non è un caso che in Italia, pur con apporti interdisciplinari, sia proprio negli studi pedagogici che sia stato teorizzato lo "Spettro dell'odio online" (Pasta, 2024b) e sperimentato in termini di ricerca educativa (Pasta, 2023; Forzinetti, Della Vedova, Pasta, Santerini, 2024).

Collocandosi in questo scenario, l'articolo riporta un caso studio sull'analisi dell'odio online, integrando tecniche di machine learning e analisi socioeducativa per classificare le retoriche d'odio online contro la figura pubblica di Papa Francesco per la sua rilevanza e per le sue prese di posizione, oggetto di polarizzazioni (Quattrococchi, 2016) su Twitter e Telegram dal gennaio 2022 all'aprile 2024. Analizza l'odio diffuso tramite tweets e canali Telegram, evidenziando le strategie comunicative ostili. Lo studio identifica tipologie di odio, bersagli e intensità tramite il modello dello "Spettro dell'Odio Online". Fornisce un quadro delle dinamiche d'odio e delle bolle informative su Telegram, contribuendo alla comprensione dell'hate speech online e delle strategie per contrastarlo (Santerini, 2021).

1. La cultura dell'esclusione, la disinformazione e l'odio antireligioso online

Come nota Maria Ranieri (2016) nell'analisi dei fenomeni di othering sottesi alle flame wars, le forme di odio online sono da collegare alla cultura dell'esclusione, i meccanismi di elezione a bersaglio e i suoi effetti sulle persone esposte. Ronson (2015) offre una riflessione stimolante dove il potere delle masse online influenzando il pensiero collettivo, potrebbe comportare il rischio di esclusione per chi non si conforma, alimentando la vergogna pubblica. Interessante notare quanto giochi, dunque, l'essere umano e quanto la tecnologia nella cultura dell'esclusione. Si potrebbe pensare che una concatenazione dei due fattori produca quello che viene definito "populismo digitale" (Palano, 2020) che spinge "all'atomismo di milioni di persone convinte di aver ragione non insieme (come credevano, sbagliando, le chiese ideologiche del secolo scorso), ma da sole o, meglio, con il solo riscontro del web" (Ferraris, 2017, p. 23). Nell'esclusione e nel proprio atomo, bolla in cui l'essere umano si troverebbe, si andrebbe a coltivare un'educazione all'indifferenza nei confronti di cosa è vero e cosa non lo è affatto; nella proliferazione di un atteggiamento che amplia il narcisismo ed il soggettivismo. In questa postura si crea un ambiente autoreferenziale, dove la percezione della realtà si plasma ad immagine dello spettatore che osservando agisce sulla sua prospettiva di sguardo. In questa nuova educazione digitale, la tecnica e gli algoritmi lavorano sugli interessi e sulle possibilità costruendo l'impalcatura della *filter bubble*¹ (Parisier, 2012), mentre l'essere umano attraverso comportamenti di stimolo e ricompensa mantiene stabile la struttura. I principali pericoli derivanti dai fenomeni sopra menzionati includono la diffusione della disinformazione, la proliferazione delle teorie del complotto e la divisione dei gruppi, che può portare all'estremismo.

"Le scienze dell'educazione sono chiamate ad affrontare questa costellazione di fenomeni e comportamenti, che possiamo sinteticamente individuare come una sorta di pedagogia nera in rete". Così Milena Santerini (2019, p. 52) descrive il lavoro pedagogico ed educativo che cerca di accompagnare e contrastare il discorso d'odio che spesso sfocia in aggressività, rabbia, violenza e frustrazione (Cohen Almagor, 2015, p. 205), tutte dimensioni biologiche e psichiche degli esseri umani che non nascono certamente con internet, sottolinea, ma che con l'espansione rapidissima delle

¹ Filter Bubble, loc. s., in Enciclopedia Treccani, Neologismi, 2017. "L'ambiente virtuale che ciascun utente costruisce in Internet tramite le sue selezioni preferenziali, caratterizzato da scarsa permeabilità alla novità e alto livello di autoreferenzialità".

tecnologie hanno assunto nuove caratteristiche. I contenuti che siano essi immagini, testo, video o un'ibridazione dei precedenti, sono tutti legati a doppio filo con le emozioni e proprio per questo riescono facilmente a catturare l'attenzione.

“La carica emotiva messa nelle comunicazioni serve a personalizzare i consumi e ad agganciare – i consumatori individuandone gusti e propensioni. Ci si trova così di fronte a ciò che è stato definito – contagio morale – amplificato dai like, shares, retweet e algoritmi. Il - mi piace – online, infatti, è meno ambiguo e più netto di quello che riceviamo faccia a faccia. La risposta ai contenuti emozionali diffusi rende ancor più gratificante la loro espressione, come afferma Brady (2019) riportando numerosi studi neuroscientifici sull'argomento, che dimostrano come varie aree del cervello associate alla gratificazione si attivino alla condivisione di contenuti emozionali. Le emozioni morali, inoltre, sono motivate soprattutto dalla volontà di mantenere un'identità positiva del proprio gruppo. Ci si scambia più facilmente storie e informazioni da condividere perché si cerca una convergenza in-group (contro l'*outgroup*)” (Santerini, 2019, p. 58). Appaiono definiti i contorni dell'hate speech che nella sua nascita e diffusione è favorito dal sistema stesso dei media, costruito sulla logica di condivisione all'interno delle proprie “tribù” di appartenenza. In tutto ciò recenti studi evidenziano che maggiore è l'esposizione e l'uso dei dispositivi mobili, minore è la capacità di generare empatia e risposte empatiche (Terry, Cain, 2016) proprio perché i social media inibiscono le possibilità di conversazioni in presenza, faccia a faccia, di cui necessitiamo come esseri umani. Dunque, un cuore nascosto all'educazione, privato della possibilità di connettersi alla linfa generativa dei processi partecipativi di comunità è monade isolata, eremita paranoico, in un loop sostenuto dai meccanismi algoritmici e dalla propria incapacità esplorativa, celata sotto una coltre di paura e timore. Istruire, formare ed educare sono i pilastri pedagogici sempre validi anche nel campo dei media digitali, volti a fare dell'essere umano un essere capace di amare l'altro al di fuori di sé.

2. La metodologia applicata alla ricerca su Twitter e su Telegram

In questa ricerca, per procedere, si sono applicate tecniche di machine learning e analisi socioeducativa (Balaji, Chandra, Annushree, 2021). Questo ha permesso un approccio combinato che ha avvicinato più campi disciplinari integrati all'interno del Centro di Ricerca sulle Relazioni Interculturali dell'Università Cattolica del Sacro Cuore di Milano². Da anni attraverso l'Osservatorio Mediavox sull'odio online dell'Università³ (Pasta, 2023a; Pasta, Santerini, Forzinetti, Della Vedova, 2021). si era sentita la necessità di integrare i campi di studio e renderli forti insieme (Pasta, 2023b; Pasta, 2021; Pasta, 2024b). Per poter analizzare e meglio individuare le retoriche d'odio e i bersagli senza perdere di vista la mole di dati da categorizzare e sistematizzare si è fatto riferimento a competenze nel campo dell'informatica e della data analysis. Sono poi stati affiancati strumenti linguistici e statistici per l'analisi di contenuto e di text mining (processamento dei dati). Tanti gli

² Centro di Ricerca sulle Relazioni Interculturali dell'Università Cattolica del Sacro Cuore di Milano: il Centro di Ricerca sulle Relazioni Interculturali dell'Università Cattolica del Sacro Cuore di Milano è stato costituito nel 2009. Nel momento in cui si è svolta la ricerca, il direttore scientifico è Milena Santerini. Per ulteriori informazioni, visitare: https://centridiricerca.unicatt.it/relazioni_interculturali.

³ Mediavox – Osservatorio sull'odio online è un progetto del Centro di Ricerca sulle Relazioni Interculturali dell'Università Cattolica del Sacro Cuore di Milano. I soggetti finanziatori sono l'Università Cattolica del Sacro Cuore, l'Istituto Giuseppe Toniolo in collaborazione con l'Ufficio Comunicazioni Sociali della Conferenza Episcopale Italiana e Unar, l'Ufficio Nazionale Antidiscriminazioni Razziali. L'Osservatorio ha avviato collaborazioni con Retinopera, il quotidiano Avvenire ed altri. Per maggiori dettagli, consultare: <https://www.mediavox.network/>.

strumenti di analisi utilizzabili, da quelli full AI (completamente automatizzati) a quelli ibridi, dove la componente umana aiuta nei procedimenti di interpretazione del big data testuale. La metodologia utilizzata in questa ricerca viene denominata: *Machine Learning Grounded Theory* (Muller, Guha, Baumer, Mimno, Shami, 2016) e indica un percorso in cui osservazione ed elaborazione teorica procedono di pari passo, in un'interazione continua. “È presente e forte non solo una metodologia Grounded, bensì un ampio quadro interpretativo di indicatori e sistemi utili anche in riferimento ad altri target d'odio, costruendo più in generale un processo di analisi con una visione più ampia” (Pasta, 2018).

La ricerca delle parole chiave e delle categorie, utile a reperire i dati, si è mossa dapprima sull'individuazione di Papa Francesco come bersaglio (Pasta, 2018) principale, successivamente si è cercato di ipotizzare intorno alla sua persona fisica altri ambiti di interesse che avrebbero potuto apportare critiche e odio. In una prima ipotesi si è ragionato sulla sua persona in carne ed ossa, immaginando possibili attacchi alla forma fisica, al paese di provenienza e al costume, all'abito che indossa, compresi gli accessori. Si è cercato, dunque, di ipotizzare una sua sfera d'azione di viaggi e spostamenti, di lettere, documenti, scritti, gesti simbolici e di contatto, oltre che azioni direttamente all'interno della Chiesa.

Pensando al suo ruolo si è considerato il papato e la Chiesa Cattolica nello specifico, con i suoi ministri e con le sue regole. In questo emerge evidente il pensiero di Bergoglio, osservato speciale rispetto a tanti ambiti dei quali si sono scelti in particolare: l'ambiente, il dialogo interreligioso, la famiglia e la sessualità, l'immigrazione, la politica e la povertà. Le riflessioni fatte in precedenza sulla figura papale hanno fatto emergere successivamente due categorie: Il dogma - liturgia e l'autorità - istituzione.

La prima che riprendeva le azioni papali interne alla Chiesa e la seconda che vedeva nella sua figura l'incarnazione dell'autorità e istituzione Chiesa. In questo primo lavoro si è affiancato uno studio di sottocategorie che ci avrebbero aiutato in un secondo momento ad individuare le parole chiave. Si è cercato di accostare il nome del Papa a parole che lo avrebbero potuto coinvolgere nelle diverse categorie elencate, come: Papa Francesco – ecologia o Papa Francesco – cambiamento climatico in riferimento alla categoria ambiente. Le parole chiave sono state scelte utilizzando strumenti statistici come *Google Trends* per sondare l'utilizzo quantitativo della parola nel web confrontandola con il bersaglio principale anche se non in correlazione. Questa struttura di partenza ci ha permesso di costruire una base d'indagine utile per indagare l'odio antireligioso con un focus principale su Papa Francesco.

Il gruppo di lavoro di Mediavox ha esplorato la possibilità di rilevare picchi mensili nell'incitamento all'odio tramite analisi temporali di campioni di tweet, valutati manualmente da esperti per identificare le retoriche e le forme predominanti di odio. La metodologia adottata si iscrive nelle tecniche di analisi dei social network. I dati sono stati raccolti utilizzando il programma open source PyCharm nell'edizione “community”⁴, consentendo l'estrazione dei tweet tramite una ricerca per query. Il corpus è stato estratto utilizzando una stringa di ricerca che comprendeva le parole chiave associate alla categoria e le parole in sottocategoria (ad esempio, rispetto alla categoria autorità e

⁴ Disponibile al link: <https://www.jetbrains.com/pycharm/>.

istituzioni: AntiPapa, FalsoPapa, Demon⁵, Pedofil, Emerito...). Per i due anni considerati (2022/2023), sono stati estratti 6.959 tweet. 984 rispetto al periodo specifico dal 24 febbraio al 24 giugno 2022 e 5.975 rispetto al periodo che va dal 1° gennaio 2022 al 31 gennaio 2023. Nella prima serie sono stati campionati circa mille tweet in un periodo specifico, mentre nella seconda serie sono stati campionati 100 tweet al mese per un totale di 13 mesi, il tutto seguendo la tecnica del campionamento casuale semplice senza ripetizione (James et al., 2017). Ogni tweet scaricato ha incluso vari attributi oltre al testo stesso, come l'username dell'autore, menzioni, hashtag, numero di retweet e like, data, ora e geolocalizzazione. (Pasta, 2018). Questo dataset è stato poi valutato manualmente da annotatori specializzati per determinare se i tweet contenessero contenuti discriminatori. In caso affermativo, sono state identificate le retoriche e le forme specifiche di odio antireligioso presente nel tweet. In precedenza, Milena Santerini e Stefano Pasta dell'osservatorio Mediavox hanno proposto una classificazione dei contenuti online basata su sette indicatori denominati "Spettro dell'odio online" (Santerini, 2023; Pasta, 2022; Pasta, 2023a).

Questi indicatori possono essere descritti nel seguente modo:

- *accessibilità al pubblico*: il contenuto è disponibile per la visione generale e non è limitato a un pubblico specifico o privato;
- *gruppo target o individuo colpito*: il contenuto d'odio prende di mira un gruppo specifico o un individuo legato a quel gruppo, spesso mirando alle minoranze o ai loro rappresentanti;
- *assenza di intento redentivo*: il messaggio d'odio non cerca di cambiare le opinioni della vittima, ma mira esclusivamente a insultare o danneggiare;
- *intenzione di causare danno*: il contenuto è creato con l'intenzione di ferire la vittima, manifestando un chiaro intento nocivo;
- *espressione di odio*: il contenuto d'odio contiene parole o discorsi che esprimono odio esplicitamente, come insulti o deumanizzazione della vittima;
- *possibilità di risposta violenta*: il linguaggio utilizzato nel contenuto è così violento o provocatorio che può facilmente scatenare risposte altrettanto violente o incitare all'odio;
- *incitamento all'odio e alla violenza*: il contenuto incoraggia apertamente l'odio e la violenza, cercando di coinvolgere altri nel diffondere tali sentimenti negativi, aumentando così il rischio di azioni odiose offline.

È importante notare che un contenuto può essere considerato d'odio anche se non soddisfa tutti e sette gli indicatori, ma la presenza di tali indicatori può aiutare a valutare l'intensità e le caratteristiche per contrastarlo in modo più efficace. Nel dataset dove annotatori esperti hanno valutato manualmente ogni singolo tweet estrapolato, si è osservata anche la classificazione delle retoriche dell'odio, dei possibili bersagli (non solo Papa Francesco) e della tipologia d'odio.

Retoriche dell'odio: insulti, derisione/ironia, esclusione/separazione, pregiudizio, umiliazione/disprezzo, paura, concorrenza, incitamento/violenza.

Bersagli dell'odio: disumanizzazione, Migranti/stranieri, colore della pelle, seconde generazioni, Rom, Ebrei, Papa Francesco, Chiesa, musulmani, personaggi famosi, casi personali, cristiani, movimenti/ordini religiosi, media cattolici, donne, orientamento sessuale e politici.

⁵ Demon, Pedofil e altri vocaboli usati nella query senza genere, né numero, perdono l'ultima lettera per aumentare la possibilità di captare un maggior numero di commenti e post che comprendano la parola. Inserendo il genere e il numero si rischierebbe di limitare la capacità d'intercettazione.

Tipologie d'odio: odio per motivi religiosi, odio sociale, odio su base di genere, odio tra pari, odio antipolitico, razzismo.

È stata inoltre presa in considerazione la tipologia d'odio che classificasse da una parte l'odio complottista e dall'altra l'odio conservatore; dove per complottista si intendeva una visione distorta, paranoica e distopica della realtà, mentre per conservatrice si è osservata una visione rigida, dogmatica e tradizionalista della realtà, escludente al nuovo e al cambiamento.

3. Sintesi dei risultati principali sui campioni di Tweets tramite Python

La prima query utilizzata per questo caso studio è stata:

(PapaFrancesco OR Papa OR Pontifex OR PapaFrancis OR Francesco OR SantoPadre OR Chiesacattolica OR PapaRomano) AND (Antipapa OR Anticristo OR Masson_ OR NWO OR Davos OR Elite OR FalsoPapa OR Falso OR Emerito OR Ratzinger OR Razinger OR Codice OR Demon_ OR Satan_ OR Pagan_ OR Pedofil_)

Il periodo temporale preso in considerazione è: 1° gennaio 2022 – 31 gennaio 2023 (figura 1).

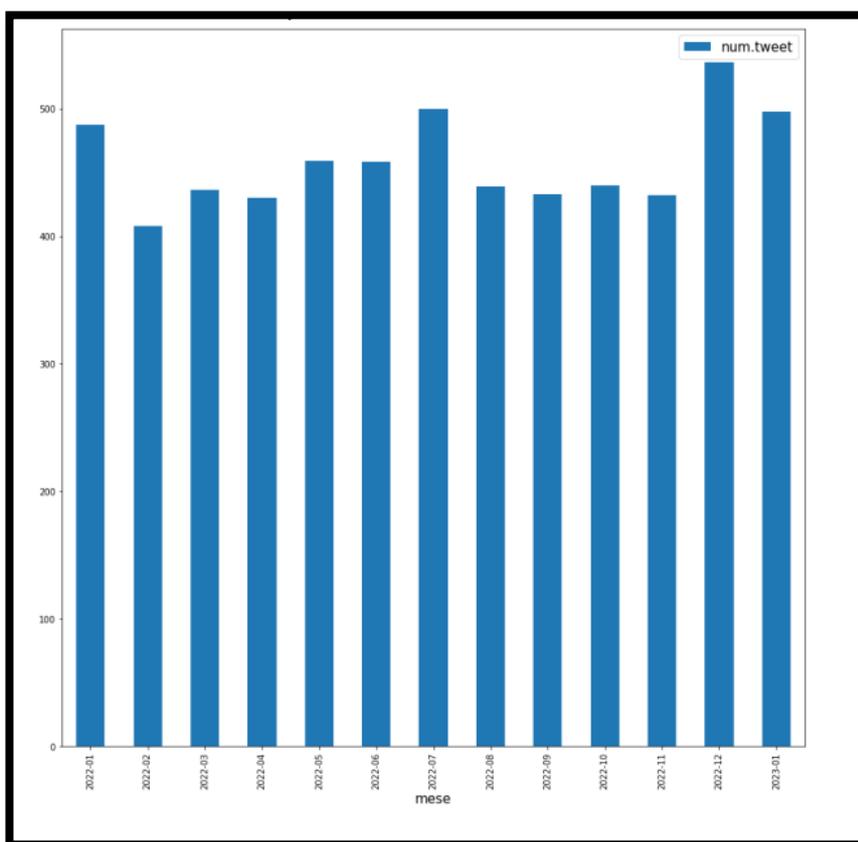


Figura 1 - Il grafico elenca la quantità di tweet processati mese per mese

In totale sono stati tutti annotati manualmente 5975 tweet estratti casualmente ogni settimana, di seguito viene mostrato l'andamento dei 5975 tweet nell'anno preso in considerazione.

A seguito della classificazione fatta manualmente dall'annotatore, sono stati individuati 383 tweet contenenti odio, il 6.41% del totale (figura 2).

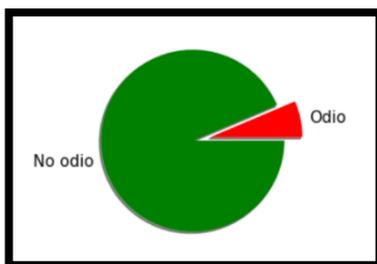


Figura 2 - il grafico evidenzia in rosso la percentuale di odio

Nel grafico in figura 3 si può osservare che su di un totale di 383 tweet, l'81,98% (314 tweet) appartiene al tipo d'odio conservatore, mentre il 21,41% (82 tweet) indica la tipologia d'odio elitaria/complottista.

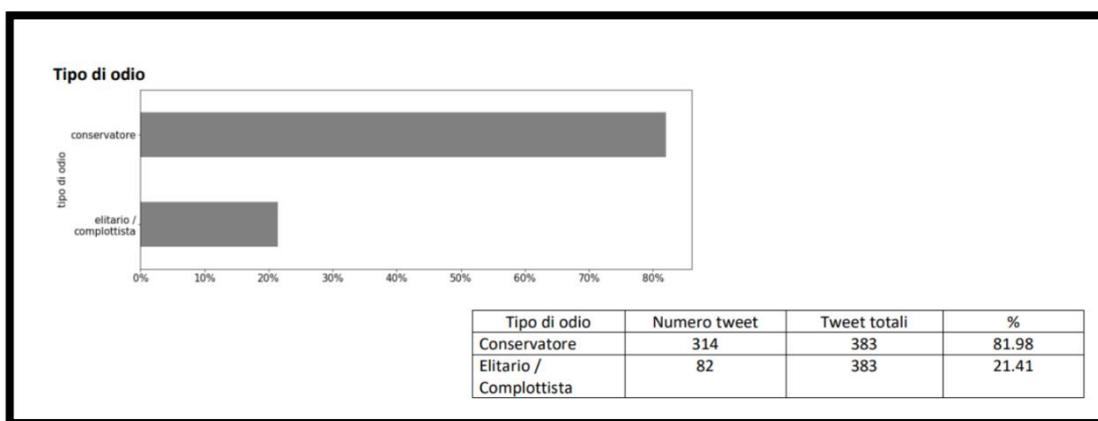


Figura 3 - il grafico evidenzia i tipi di odio

Attraverso gli indicatori, su di un gruppo target di 383 tweet esprimenti odio, quasi la totalità lo esprime esplicitamente, il 96,08% (368 tweet) pubblicamente. Il 3,92% (15 tweet) incita all'odio e alla violenza, il 3,39% (13 tweet) mostra l'intento di fare del male, lo 0,78% (3 tweet) ha potenziale per una risposta violenta e solamente uno non vuol redimere (figura 4).

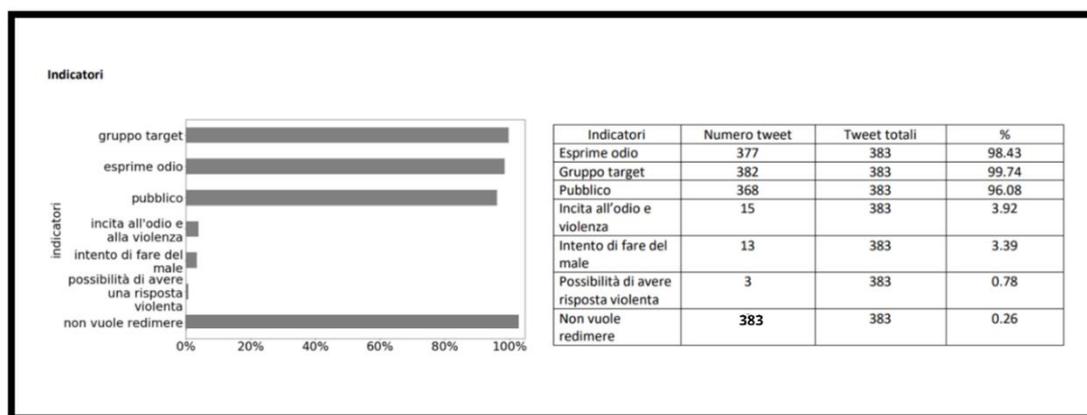


Figura 4 - il grafico evidenzia gli indicatori

Analizzando le categorie (figura 5), sul gruppo dei tweet contenenti odio, quasi la totalità rientra all'interno della categoria di ricerca, mentre da osservare come i tweet contenenti odio spazino tra categoria politica, il 6.53% (25 tweet) e il 3.39% (13 tweet) in riferimento al dogma e alla liturgia della Chiesa.

L' 1.83% (7 tweet) interessa la categoria famiglia/sexualità.

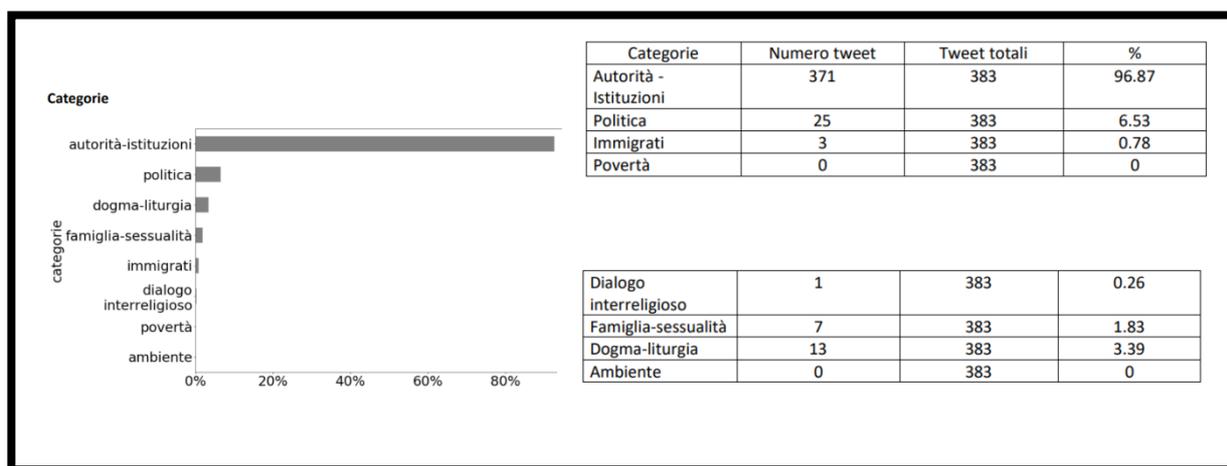


Figura 5 - il grafico evidenzia le categorie

Nel grafico in figura 6, viene mostrato come il 65.27% (250 tweet) contenga umiliazione e disprezzo nei messaggi e nei post, il 55.35% (212 tweet) pregiudizio, il 46.48% (178 tweet) prevalentemente insulti, il 29.24% (112 tweet) esclusione/separazione ed il 20.89% (80 tweet) disumanizzazione.

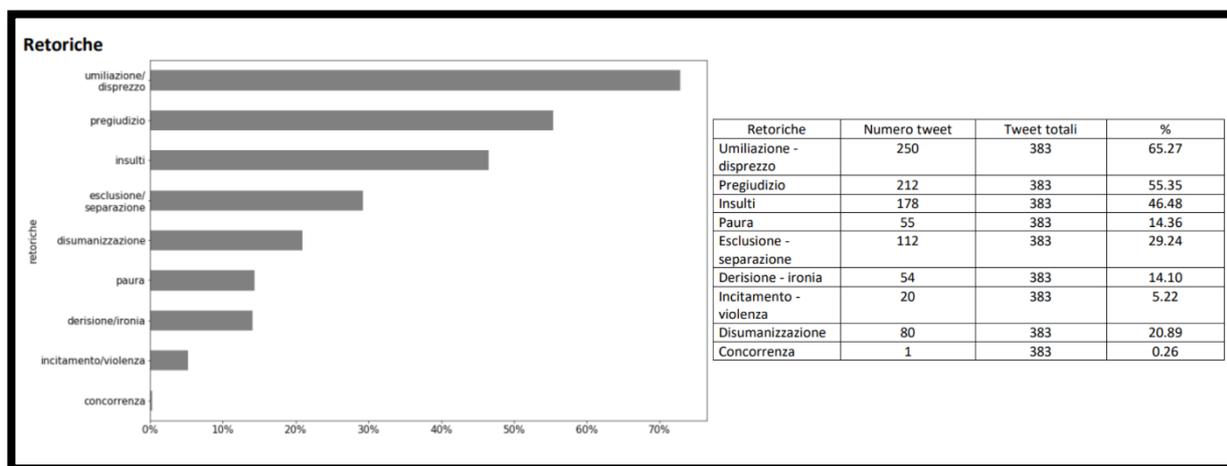


Figura 6 - Il grafico evidenzia le retoriche

Nei bersagli (figura 7) notiamo un 92.69% (355 tweet) verso Papa Francesco, un 28.46% (109 tweet) verso la Chiesa e un 13.84% (53 tweet) verso i cristiani.

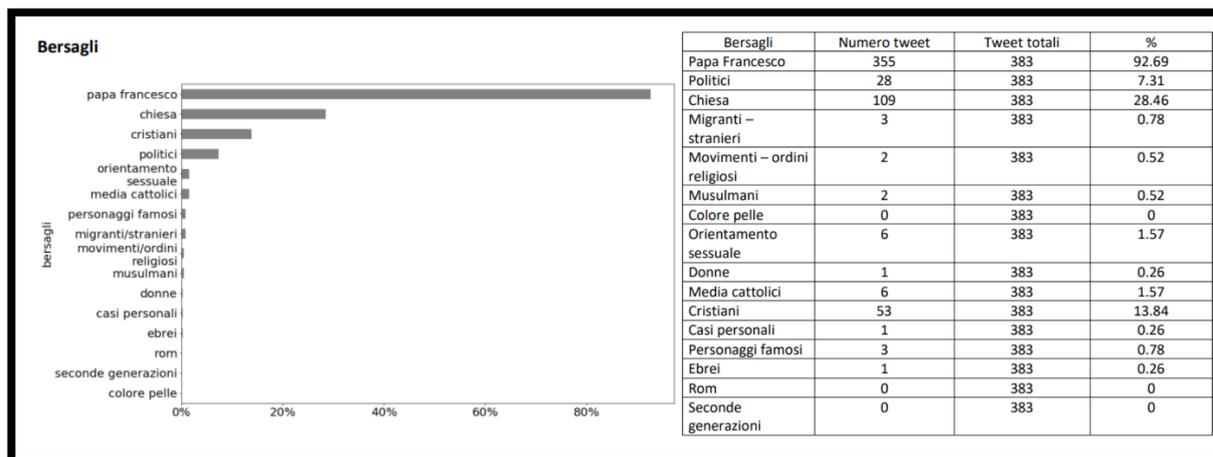


Figura 7 - il grafico evidenzia i bersagli

L'odio è in prevalenza concentrato ed in riferimento a motivi religiosi, ben 364 tweet su 383, per un 95.04%, ma rimane interessante notare l'odio sociale pari al 20.10% (77 tweet) del totale (figura 8).

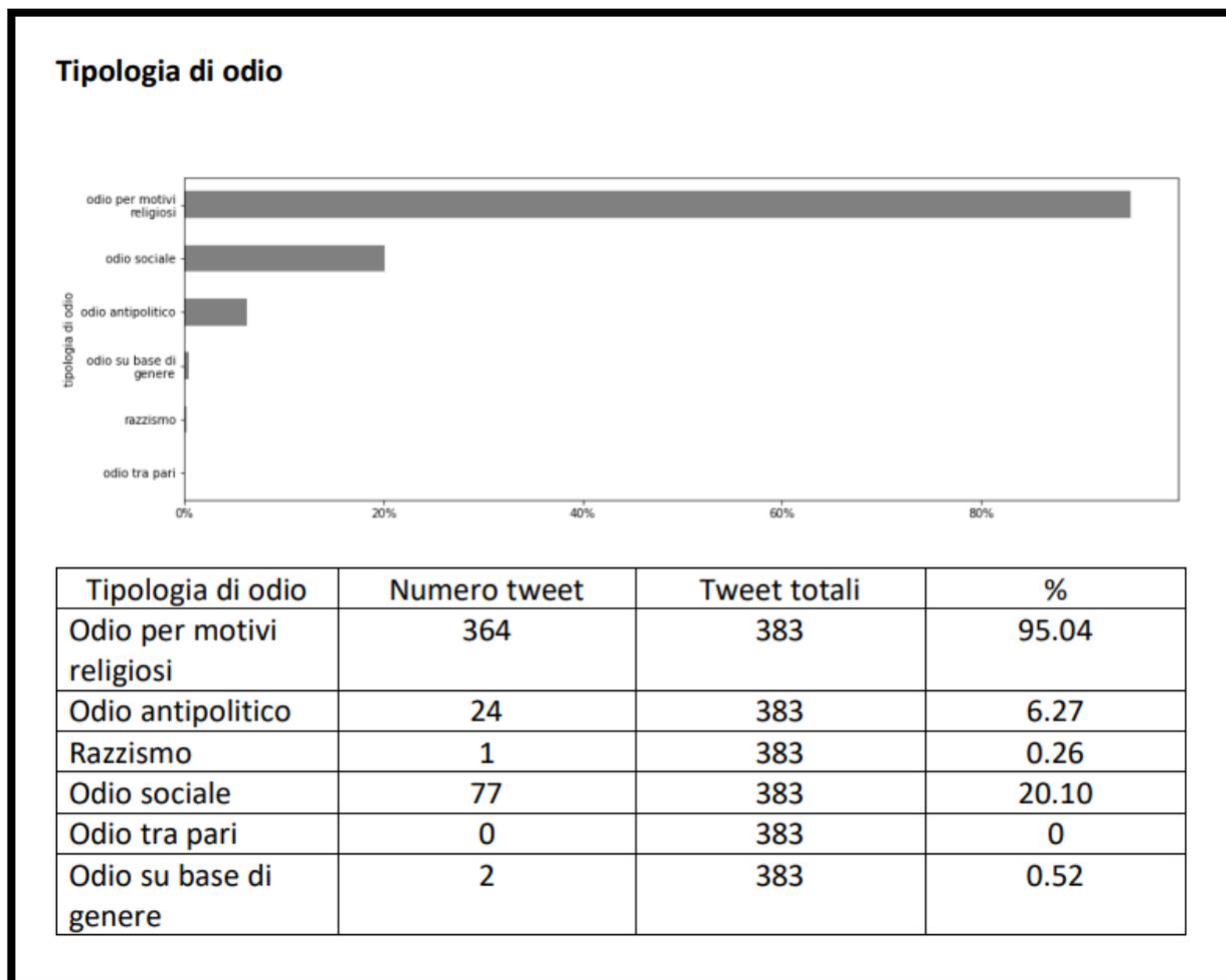


Figura 8 - il grafico evidenzia le tipologie di odio

La seconda query utilizzata per questo caso studio è stata:

(PapaFrancesco OR Papa OR Pontifex OR PapaFrancis OR Francesco OR SantoPadre OR Chiesacattolica OR PapaRomano) AND (Comunist_ OR Badoglio OR Nazista OR Mafioso OR Usurpatore OR Imbroglione OR Pupazzo OR Burattino) AND (Zelensky OR Guerra OR Ucraina OR Russia OR Pace OR Politica OR Putin OR Biden OR Nato OR Ursula OR Leyen OR Vaticano)

Il periodo temporale preso in considerazione è: 24 febbraio 2022 – 24 giugno 2022. In totale sono stati scaricati 984 tweet e sono stati tutti annotati manualmente e di seguito si viene mostrato l'andamento dei 984 tweet nei tre mesi presi in considerazione (figura 9).

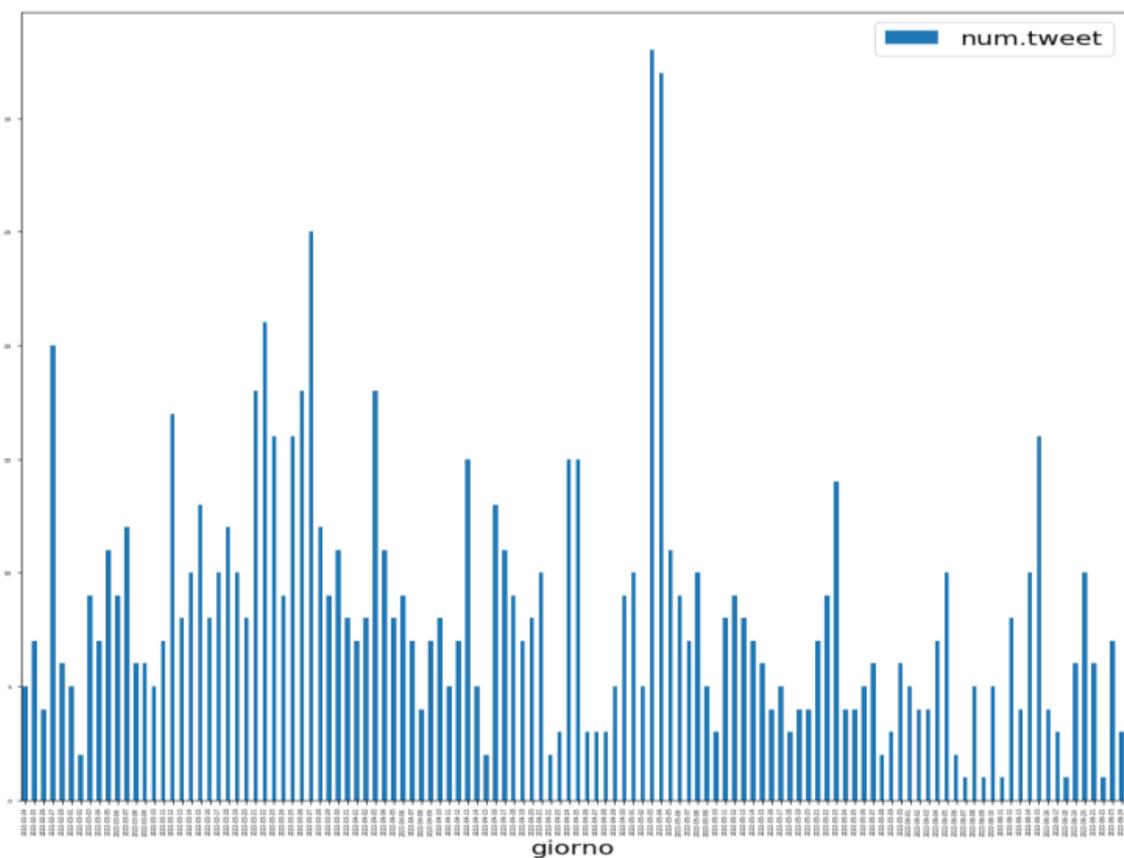


Figura 9 - il grafico elenca la quantità di tweet giornalieri processati nel periodo di indagine

Sono stati individuati 33 tweet contenenti odio, il 3.35% del totale (figura 10).

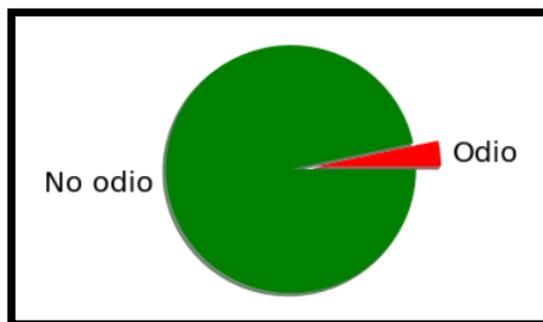


Figura 10 - il grafico evidenzia in rosso la percentuale di odio

Nel grafico in figura 11 si può osservare che su di un totale di 33 tweet, il 78,79% (26 tweet) appartiene al tipo d'odio conservatore, mentre il 21,21% (7 tweet) indica la tipologia d'odio elitaria/complottista.

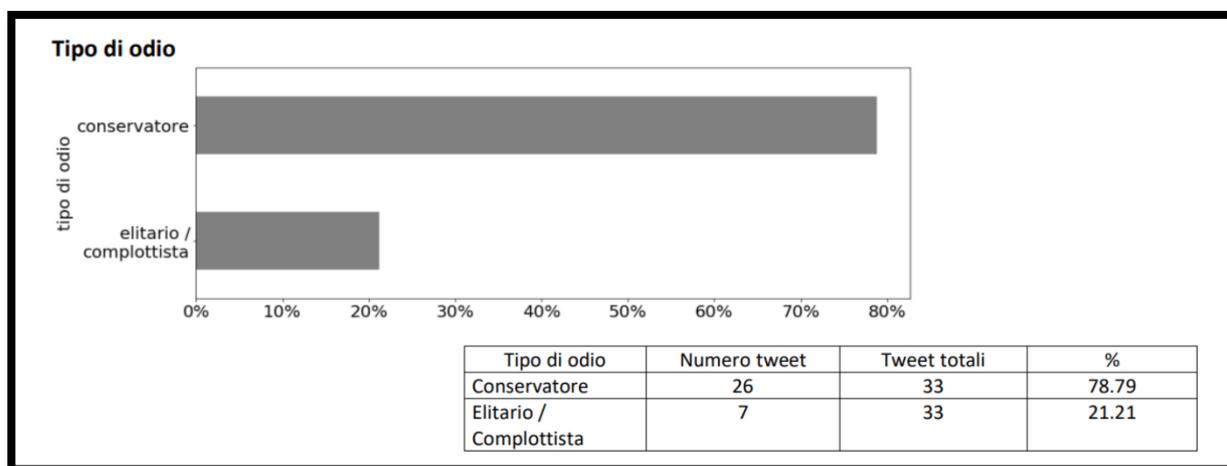


Figura 11- il grafico evidenzia i tipi di odio

Attraverso gli indicatori, su di un gruppo target di 33 tweet esprimenti odio, la totalità lo esprime esplicitamente e pubblicamente. Il 15,15% (5 tweet) incita all'odio e alla violenza, il 3,03% (1 tweet) mostra l'intento di fare del male (figura 12).

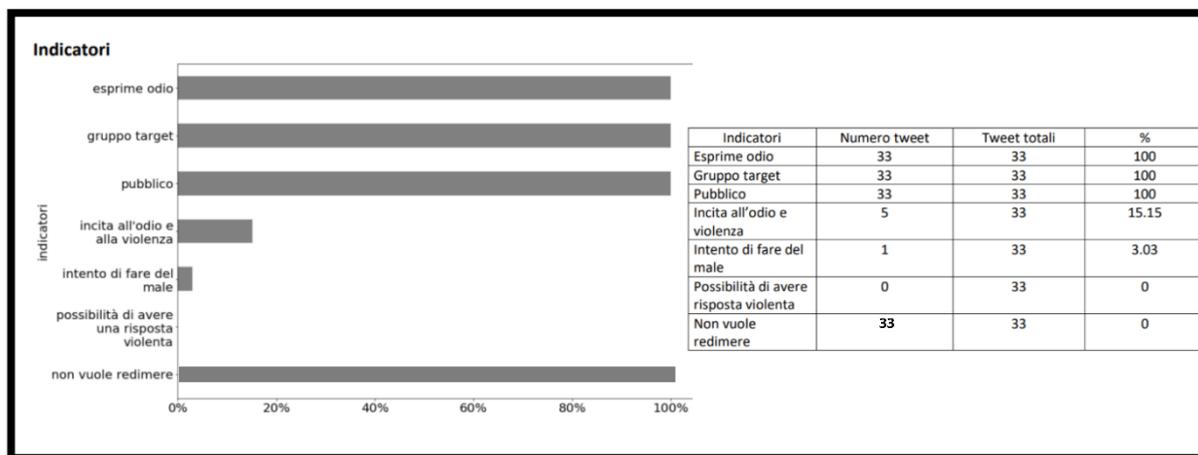


Figura 12 - Il grafico evidenzia gli indicatori

©Anicia Editore

QTimes – webmagazine

Anno XVI - n. 4, 2024

www.qtimes.it

doi: 10.14668/QTimes_16427

Analizzando le categorie (figura 13), sul gruppo dei tweet contenenti odio, il 57,58% (19 tweet) rientra all'interno della categoria di ricerca, mentre da osservare come i tweet contenenti odio spazino anche tra categoria autorità/istituzioni, il 93.94% (31 tweet) e il 12.12% (4 tweet) in riferimento alla categoria immigrati. Il 3.03% (1 tweet) interessa la categoria povertà.

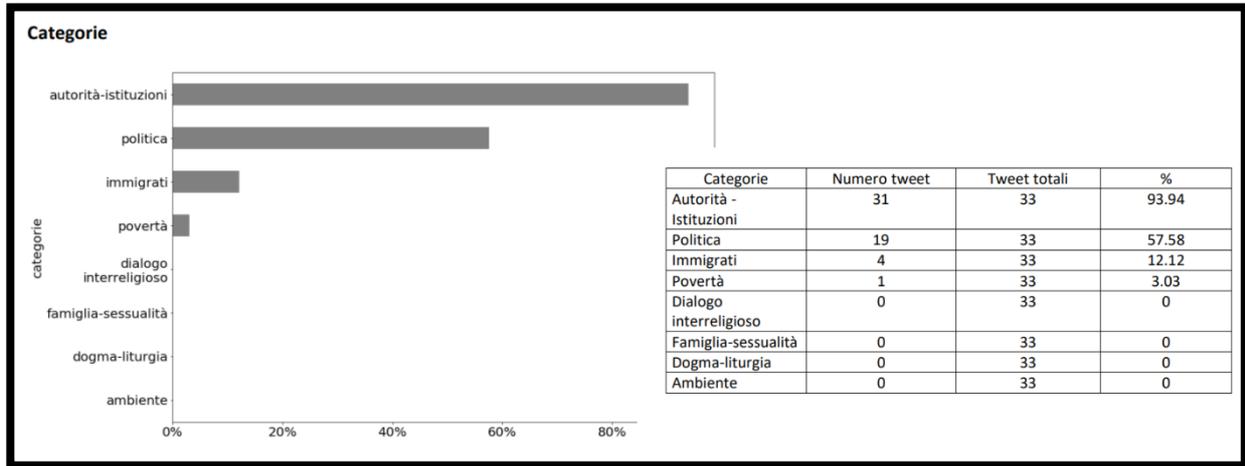


Figura 13 - il grafico evidenzia le categorie

Il grafico in figura 14 mostra come il 90.91% (30 tweet) contenga umiliazione, disprezzo e pregiudizio nei messaggi e nei post, il 54.55% (18 tweet) insulti, il 30.30% (10 tweet) paura, il 21.21% (7 tweet) derisione/ironia ed il 12.12% (4 tweet) disumanizzazione. Da notare come spostando la categoria ed il focus tematico la retorica disumanizzante sembra decrescere incentrandosi di meno sulla figura di Papa Francesco e spostandosi sui politici e sulla politica.

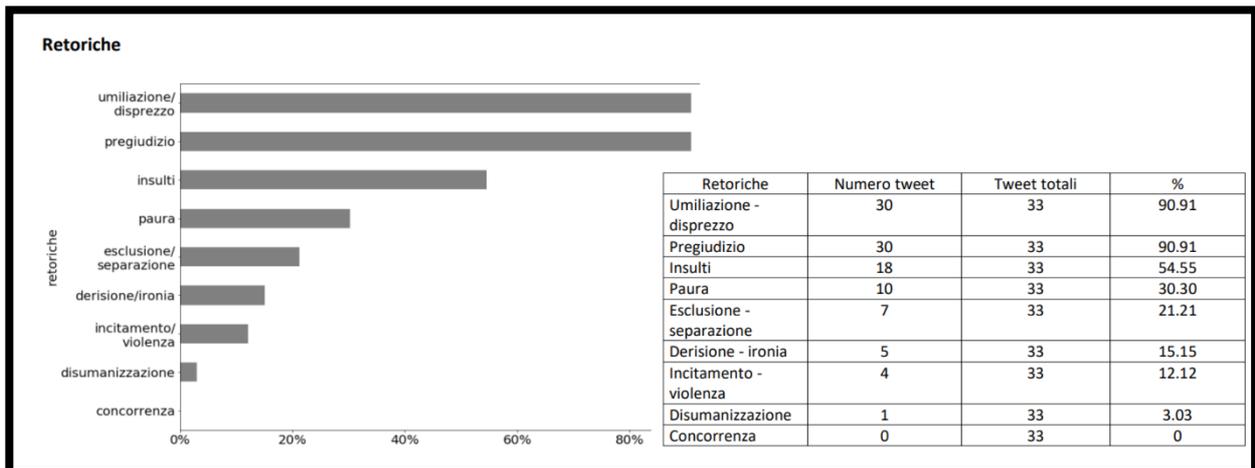


Figura 14 - il grafico evidenzia le retoriche

Nei bersagli (figura 15) notiamo un 90.91% (30 tweet) verso Papa Francesco, un 51.52% (17 tweet) verso i politici e un 21.21% (7 tweet) verso la Chiesa. Compaiono alcuni tweet verso movimenti/ordini religiosi, i migranti/stranieri, i musulmani e in riferimento al colore della pelle.

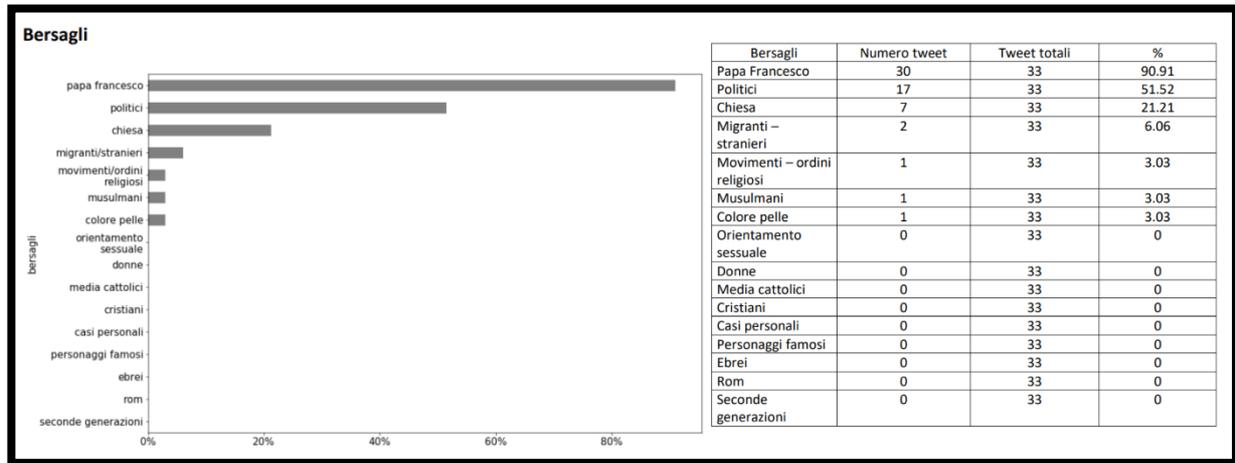


Figura 15 - il grafico evidenzia i bersagli

L'odio è in prevalenza concentrato ed in riferimento a motivi religiosi, ben 29 tweet su 33. Per un 87.88%, è odio politico pari al 51.52% (17 tweet) e per un 12.12% (4 tweet) si rivela sia odio razziale che sociale (figura 16).

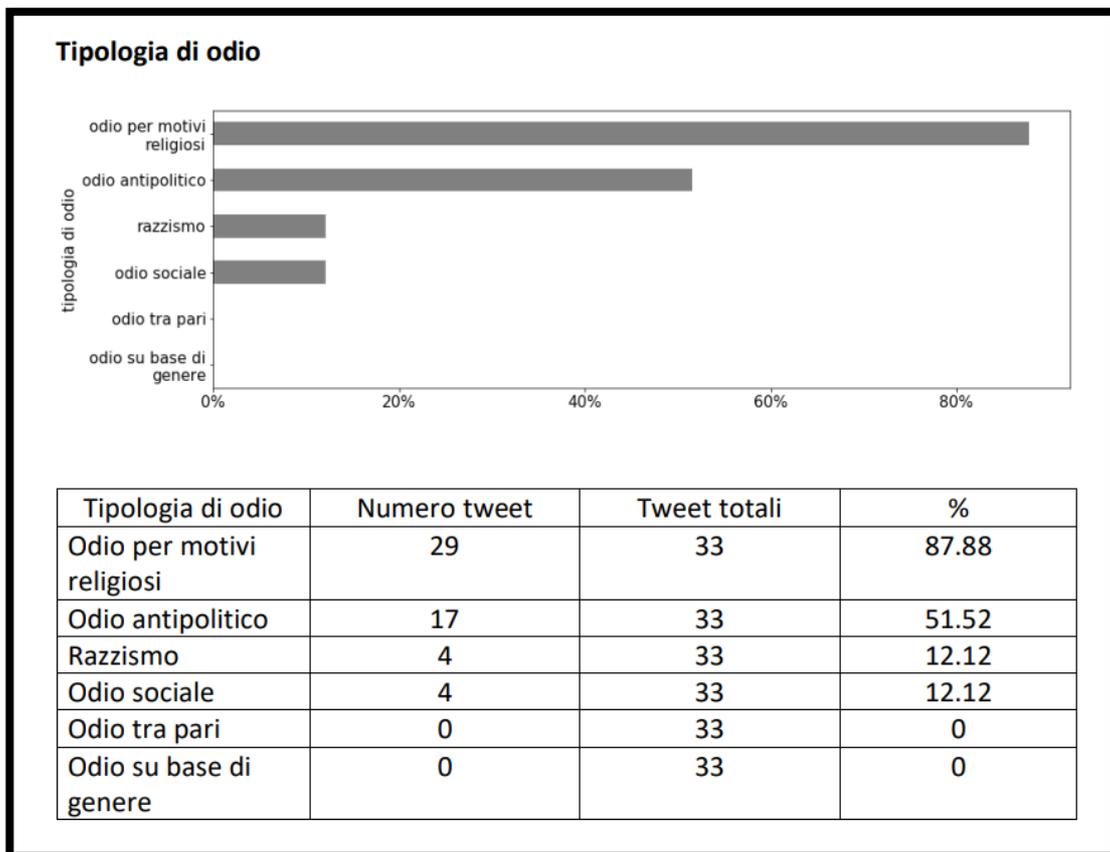


Figura 16 - il grafico evidenzia le tipologie di odio

4. I canali di disinformazione religiosa su Telegram

Telegram è un social network importante da osservare anche sotto la luce e la lente dell'odio e dei messaggi antireligiosi che imperversano, sapendo che addentrarsi all'interno di questo social per recepire dati ed informazioni non è così facile. In questa seconda fase della ricerca prenderò a campione un gruppo di canali e di messaggi che gravitano attorno alle parole "Papa Francesco" e "Papa Ratzinger". Sfrutteremo la ricerca per parole per "navigare" all'interno del social network e spazieremo per i canali che conterranno messaggi e post con particolare attenzione ai commenti sottostanti. Sarà difficile, perché Telegram permette molte funzionalità in grado di mascherare, limitare o bannare automaticamente i contenuti a seconda delle impostazioni private di ogni singolo canale ricercato. Questo vuole dire che potremo trovare canali molto aperti, che abilitano i commenti e le reazioni e altri più limitati dove ci saranno i messaggi caricati dagli amministratori delle pagine, ma non i commenti del pubblico.

In questa direzione si è tracciata una domanda di ricerca, che potesse collegare i diversi canali trovati e ci si è resi successivamente conto che spesso questi canali tra di loro comunicavano.

Attraverso la condivisione dei loro stessi post e messaggi, che spesso non erano tutti di originaria creazione, ma di altri canali che a loro volta condividevano e disseminavano il contenuto di pagina in pagina. Partendo da un primo canale a campione uscito dalla ricerca attraverso le due parole chiave, ho provato a seguire la traccia lasciata dagli amministratori di pagine cliccando sui nomi che uscivano a fonte di ogni messaggio. Questo mi ha permesso di arrivare ad un numero chiuso di canali che si ripostavano e si condividevano tra loro in maniera sistemica, circolare e pressoché infinita. Una vera e propria bolla in un reticolato "tribale" in cui si riconosce l'onda di movimento degli "sciame digitali" (Palano, 2020), delle masse.

Un indice di 92 canali che ad un certo punto tende a condividersi costantemente replicando post di cerchie definite rafforzando legami ed escludendo altri possibili contatti esterni⁶. I canali sono sempre gli stessi e andando a ritroso nel tempo in un periodo spazio-temporale di circa 3 anni, dal 2022 al 2024 si nota come la sua struttura si componga fondamentalmente degli stessi amministratori, con alcuni gruppi di pagine principali che in maniera gerarchica, sia per followers che per contenuti, primeggiano, in forma piramidali sui contenuti di tutte le altre cellule. Emerge dunque un'informazione apparentemente molto diffusa, ma veicolata da pochi centri d'origine. Ho diviso i canali per cinque tipologie di argomento che spesso si manifestavano nell'osservazione dei messaggi e dei post, cercando di suddividere i diversi canali all'interno di un diagramma di Venn a cinque incroci. Questo mi ha permesso di visualizzare i canali che con numeri di pubblico elevato e capacità di risposta tematica tenevano il podio dell'informazione all'interno della bolla. Tengo a precisare che non tutti i canali individuati contengono odio, anzi, alcuni sono di semplice divulgazione di contenuti rispetto alla tematica prescelta e conservano uno spirito critico e di confronto, altri sono molto meno attivi e tendono a caricare qualche contenuto ogni tanto.

La bolla, dunque, costruisce il pubblico di fruitori basandosi sulle loro interazioni e le condivisioni gestite dai detentori delle pagine, ma non tutti i canali che vengono condivisi o utilizzati adottano un approccio d'odio. Si costituisce una specie di gerarchia interna dove i canali con più pubblico, attivi e che spaziano su più argomentazioni, a volte, per interesse personale del singolo amministratore

⁶ Nei confini di questa ricerca non sono state prese in considerazione le ramificazioni internazionali ed estere, che avrebbero sicuramente allargato l'orizzonte.

della pagina, o del gruppo, se ce n'è più di uno, tendano a condividere canali non per forza con attitudine ostile o provocatoria. Un caso può essere riconosciuto nel canale Telegram del noto vaticanista Aldo Maria Valli che più volte rientrerà nelle condivisioni dei gruppi della bolla, pur mantenendo un profilo riflessivo, critico, ma rispettoso. Qui di seguito la struttura della bolla individuata (figura 17) e i numeri di riferimento dei canali che non riporto per esteso, ma che sono stati ulteriormente classificati per tematiche e numero di followers complessivi.

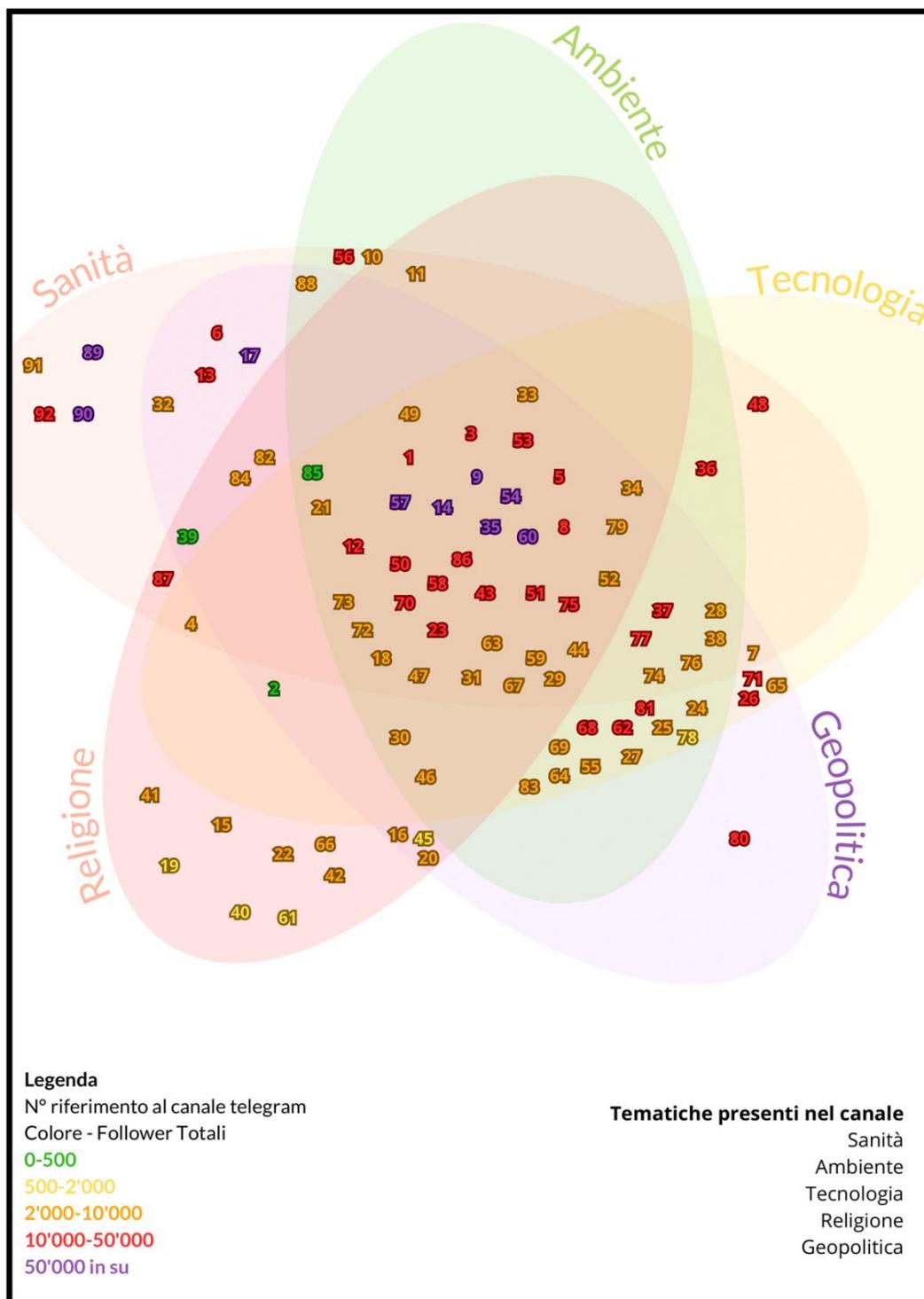


Figura 17 - Il diagramma evidenzia il posizionamento dei canali elencati in seguito, secondo i dati riportati in legenda

Dalla prima fase di ricerca osservativa ci si sposta verso “studi di caso” scelti tra le tante pagine di questa bolla che ha sicuramente delle caratteristiche:

- è definita nei suoi contorni e confini a livello nazionale;
- ogni canale ha un’immagine di profilo riconoscibile;
- ci sono almeno più di dieci messaggi pubblicati dagli amministratori nel periodo di riferimento preso in considerazione;
- ogni canale tocca almeno uno tra i seguenti cinque argomenti tematici presi in considerazione: Sanità, ambiente, tecnologia, religione e geopolitica.

I canali presi ad esempio e che vedremo successivamente hanno tra i loro messaggi una predominanza di post a tematica religiosa per i quali sono stati scelti alcuni a campione per osservarne il contenuto ed i commenti.

Il primo caso riguarda un post su “Piccolo resto cattolico con don Minutella” che riprende una foto di Papa Francesco sulla Papamobile visibilmente stanco, insieme alla frase "Il volto è lo specchio dell’anima". Il secondo caso riguarda un post su "La cruna dell’ago" di Cesare Sacchetti, condiviso il 25 marzo 2024. Il post mostra un'immagine di Papa Francesco visibilmente stanco e include la frase "Non manca molto direi. Il terzo caso è un post su "In Telegram Veritas", condiviso il 7 settembre 2022. Questo post cita un articolo dell'Ansa intitolato "Andate a votare. Il Papa accorcia la visita a Matera" e aggiunge un commento critico dell'autore.

5. Verso una contro narrazione che “costruisce insieme”

Anche la contro narrazione nei confronti dell’hate speech si basa sulle relazioni e su come una comunità si possa attivare insieme per rispondere efficacemente ad episodi singoli o diffusi di odio online.

“Le narrazioni alternative e le contro narrazioni combattono il discorso d’odio screditando e decostruendo le narrazioni su cui si basano. Propongono anche narrazioni (alternative) basate sui diritti umani e sui valori democratici, come l’apertura, il rispetto della differenza, la libertà e l’uguaglianza. Possono farlo fornendo informazioni alternative e accurate, usando l’umorismo e facendo appello alle emozioni sui temi in questione e tenendo conto delle diverse prospettive e punti di vista”⁷.

Un valido esempio è il manuale *We CAN!*⁸ Il manuale è una risorsa sviluppata dal Consiglio d'Europa nell'ambito della campagna "No Hate Speech Movement" (De Latour, Del Felice, Ettma, 2017). È concepito per aiutare giovani e attivisti per i diritti umani a contrastare il discorso d'odio online tramite narrazioni contro e alternative. Il manuale fornisce approcci comunicativi ed educativi basati sui principi dell'educazione ai diritti umani e della partecipazione giovanile. Oltre a "We CAN!", la campagna offre altre risorse educative come il manuale "Bookmarks" (Keen, Georgescu, 2016) che è

⁷ Redazione, Come costruire una contronarrazione, Rete Nazionale per il contrasto ai discorsi e ai fenomeni d’odio, febbraio 2021, <https://www.retecontrolodio.org/2021/02/01/come-si-costruisce-una-contronarrazione/>

⁸ We CAN! Taking action against hate speech through counter and alternative narratives, Council of Europe, 2017, <https://www.coe.int/en/web/no-hate-campaign/we-can-alternatives>.

specificamente progettato per supportare l'educazione ai diritti umani tra i giovani e include una serie di attività per contrastare il discorso d'odio online. Questi strumenti sono disponibili in diverse lingue e possono essere scaricati dal sito del Consiglio d'Europa, il che facilita la loro diffusione ed utilizzo in contesti internazionali.

6. Sintesi e riepilogo dati emersi su Telegram e Twitter (X)

Twitter (X)

Categoria Autorità e istituzioni (gennaio 2022 - gennaio 2023)

In questa categoria sono stati analizzati 5.975 tweet, di cui 383 (6.41%) contenevano espressioni di odio. La maggior parte di questi tweet presentava un tipo di odio conservatore (81.98%), seguito da quello elitario/complottista (21.41%). I principali indicatori di odio rilevati includevano l'espressione esplicita di odio (96.08%), l'incitamento all'odio e alla violenza (3.92%) e l'intenzione di causare danno (3.39%). Le retoriche dell'odio più comuni erano umiliazione/disprezzo (65.27%), pregiudizio (55.35%), e insulti (46.48%). Papa Francesco era il bersaglio principale (92.69%), seguito dalla Chiesa (28.46%) e dai cristiani (13.84%).

Categoria politica (febbraio 2022 - giugno 2022) In questa categoria sono stati analizzati 984 tweet, di cui 33 (3.35%) contenevano odio. Anche in questo caso, l'odio conservatore era predominante (78.79%), seguito dall'odio elitario/complottista (21.21%). Tutti i tweet di odio contenevano espressioni esplicite di odio (100%), mentre l'incitamento all'odio e alla violenza era presente nel (15.15%) dei casi e l'intenzione di causare danno nel 3.03%. Le retoriche dell'odio più comuni erano umiliazione/disprezzo (90.91%), pregiudizio (90.91%) e insulti (54.55%). Papa Francesco era ancora una volta il principale bersaglio (90.91%), seguito dai politici (51.52%) e dalla Chiesa (21.21%).

Fluttuazione dell'odio: Si sono notati picchi di tweet contenenti odio in corrispondenza di eventi specifici e dichiarazioni pubbliche di Papa Francesco, come la sua posizione sulle spese militari e la guerra in Ucraina.

Retoriche dominanti: Umiliazione, disprezzo, pregiudizio, e insulti sono le retoriche più comuni nei tweet d'odio. Tipologie di Odio: L'odio conservatore è predominante, riflettendo una reazione alla percezione di Papa Francesco come destabilizzatore dell'ordine tradizionale.

L'odio elitario/complottista è legato a teorie del complotto e visioni paranoiche. L'analisi ha dimostrato che l'odio online verso Bergoglio è fortemente influenzato da eventi specifici e tematiche politiche e religiose, con un focus significativo su retoriche di umiliazione e pregiudizio.

Telegram

L'analisi dei canali di disinformazione religiosa su Telegram rivela come immagini decontestualizzate e frasi provocatorie possano generare odio e alimentare teorie complottiste. La figura di Bergoglio è spesso bersaglio di questi attacchi, con post che sfruttano la sua immagine per suscitare reazioni negative. L'individuazione di una "bolla" chiusa di 92 canali che tra di loro mantengono legami di contatto attraverso le condivisioni non identifica necessariamente una bolla d'odio, ma un ambiente digitale nel quale è possibile coltivare disinformazione e complottismo. Tutto ciò è dovuto alle poche fonti di informazione, anche se la percezione del fruitore potrebbe essere posizionata nella moltitudine di prospettive e pareri.

Il caso di don Giuseppe Minutella

Nel post osservato, l'immagine di Papa Francesco è utilizzata per suscitare dubbi istintuali sull'osservatore. La frase "Il volto è lo specchio dell'anima" accompagna l'immagine, concentrando l'attenzione del pubblico sulla figura del Papa, che appare visibilmente affaticato. La foto è decontestualizzata, omettendo dettagli come il ricovero ospedaliero recente. Nei commenti, l'immagine del Papa viene interpretata come maligna, con epiteti come "faccia da demonio" e "gesuita schifoso". Una teoria complottista lo associa alla divinità Pacha Mama, basandosi su un evento nei giardini Vaticani decontestualizzato.

Il post sulle condizioni di salute del Papa su "La Cruna dell'Ago" di Cesare Sacchetti

Cesare Sacchetti gestisce il canale Telegram "La cruna dell'ago" con oltre 62.000 followers. Un post del 25 marzo 2024 ritrae il Papa visibilmente stanco con la frase "Non manca molto direi". Il post ha 38.000 visualizzazioni, 161 commenti e 1.176 reazioni, con molte emoticon di approvazione e alcune offensive. Tra i commenti, il 23.6% esprime odio, con auguri di morte e teorie complottiste su un passato oscuro del Papa, sostenute da fonti inaffidabili.

Il post su Papa Francesco, C.e.i. e diritto di voto su "In Telegram Veritas"

"In Telegram Veritas", con più di 33.000 followers, è attivo da novembre 2020. Un post del 7 settembre 2022 tratto da un articolo dell'Ansa aggiunge un commento ironico sul Papa e il diritto di voto. Il post ha 18.200 visualizzazioni, 78 commenti e 266 reazioni, con emoticon di approvazione e insulti. Tra i commenti, il 17.9% esprime odio, criticando il Papa e associandolo a teorie complottiste, come la mafia di San Gallo che avrebbe manipolato la sua elezione. In conclusione, la disinformazione religiosa su Telegram è alimentata da immagini manipolate e commenti provocatori.

Conclusioni

La ricerca offre un contributo alla comprensione delle dinamiche dell'hate speech online, con un focus particolare sull'odio antireligioso e sulle strategie ostili rivolte a figure pubbliche come Bergoglio. L'integrazione di tecniche di machine learning e analisi socioeducativa ha permesso di organizzare sistematicamente i dati, fornendo una rappresentazione dettagliata delle espressioni di odio e delle loro peculiarità.

L'adozione del modello dello "Spettro dell'Odio Online" ha facilitato la valutazione dell'intensità e della pericolosità dei messaggi d'odio, aiutando a individuare le aree più critiche che necessitano di interventi mirati. In questo contesto, le scienze dell'educazione possono giocare un ruolo chiave, promuovendo approcci pedagogici volti a contrastare la cultura dell'esclusione e la diffusione della disinformazione.

Le strategie di intervento suggerite includono un'educazione digitale diffusa e integrata nel welfare di comunità con percorsi educativi che migliorino la consapevolezza delle dinamiche d'odio online e incoraggino l'uso responsabile dei social media, promuovendo l'empatia e la comprensione reciproca. Negli ultimi anni, molti interventi volti a contrastare la diffusione dell'hate speech online si sono concentrati sui giovani⁹, poiché rappresentano un indicatore di come stiano evolvendo il linguaggio

⁹ Un esempio da citare può essere "REASON" - REAct in the Struggle against ONline hate speech, un progetto, promosso e finanziato dalla Commissione europea, ideato e coordinato da Unar (Ufficio Nazionale Antidiscriminazioni Razziali)

e la percezione di certe espressioni. Le parole, infatti, acquisiscono significati e impatti diversi a seconda del periodo storico e del contesto (Moscovici, 1989). Quando si affronta il tema dell'odio online in una classe oggi, spesso è l'educatore a fare una prima scoperta: le soglie di accettabilità delle espressioni offensive sembrano essersi spostate, al punto che ciò che un tempo avrebbe suscitato riprovazione ora viene spesso trattato con leggerezza e ironia. Questa dinamica è ancora più marcata sul web, dove l'effetto è amplificato e normalizzato da meme e pagine di umorismo nero, che i giovani non solo seguono con passione, ma contribuiscono attivamente a creare e diffondere (Vitullo, 2020). Il fenomeno della comunicazione online richiede una revisione critica dei paradigmi educativi, specialmente in relazione alla diffusione di discorsi d'odio. È fondamentale riconoscere che le esperienze cognitive e affettive sono interconnesse nei contesti educativi, come la scuola, per promuovere un approccio integrato che unisca pensiero, emozione e azione. Questa integrazione è essenziale per sviluppare una consapevolezza morale riguardo ai fenomeni di odio, guidando educatori e famiglie verso strategie più efficaci di intervento e prevenzione (Santerini, 2024). Per concludere, il contrasto dell'odio online può essere affrontato attraverso un approccio integrato alla cittadinanza digitale, articolato in tre modalità principali:

1. *Interpretativa*: Analisi dei contenuti mediali che considerano anche le emozioni e le esperienze degli utenti, promuovendo una comprensione più profonda delle dinamiche online.
2. *Scienze sociali*: Studio delle interconnessioni tra media e contesti socio-culturali, economici e politici, per evidenziare come questi influiscano sulla diffusione dell'odio.
3. *Creativa*: Produzione di opere personali che riflettono l'esperienza individuale e incoraggiano risposte creative all'odio, utilizzando diversi formati mediatici.

Queste modalità non sono isolate, ma si intrecciano per favorire un'educazione al pensiero critico, indispensabile per affrontare le sfide della comunicazione digitale (Buckingham, 2020).

Rimangono infine, due azioni specifiche in grado di prevenire e contrastare il fenomeno: agire per opporsi alla polarizzazione e alla disinformazione presenti all'interno delle bolle di informazione, specialmente su piattaforme come Telegram, attraverso il debunking¹⁰ e la promozione di contenuti che incentivino il confronto e il dialogo; oltre a proseguire con l'osservazione e la categorizzazione dei contenuti d'odio per identificare nuove retoriche emergenti e gruppi target, migliorando così le strategie di prevenzione e contrasto.

Riferimenti bibliografici:

Annacontini, G., De Serio, B. (2019). Riconoscere l'odio per progettare un mondo migliore. *MeTis - Mondi educativi. Temi, indagini, suggestioni*: 9(2).

Buckingham, D. (2020). *Un manifesto per la media education*. Milano: Mondadori.

della Presidenza del Consiglio dei ministri, in partenariato con Università Cattolica del Sacro Cuore, IRS - Istituto per la ricerca sociale SCARL e Associazione Carta di Roma. Disponibile al link: <https://reasonproject.eu/il-progetto/>

¹⁰ Debunking, s. m. inv., in Enciclopedia Treccani, Neologismi, 2020. "Opera di demistificazione e confutazione di notizie o affermazioni false o antiscientifiche, spesso frutto di credenze, ipotesi, convinzioni, teorie ricevute e trasmesse in modo acritico".

- Cohen Almagor, R. (2015). *Confronting the Internet's Dark Side. Moral and Social Responsibility on the Free Highway*. New York: Cambridge: University Press.
- Consiglio d'Europa (2022). *Lotta contro il discorso d'odio. Raccomandazione CM/Rec(2022)161 del Comitato dei Ministri agli Stati membri sulla lotta contro i discorsi d'odio (Adottata dal Comitato dei Ministri il 20 maggio 2022 in occasione della 132^a Sessione del Comitato dei Ministri)*. Strasburgo: Cedex.
- Crescenza, G. (ed.) (2024), *Educare in tempi di odio e violenza. Sfide pedagogiche e istituzionali* (pp. 15-24). Bari: Progedit.
- De Latour, A. Del Felice, C., Ettema, M. (eds.) (2017). *We can! Taking Action against Hate Speech through Counter and Alternative Narratives*. Strasbourg: Council of Europe.
- Faloppa, F. (2020). *#Odio. Manuale di resistenza alla violenza delle parole*. Milano: Utet.
- Ferraris, M. (2017). *Postverità e altri enigmi*. Bologna: Il Mulino.
- Fiorucci, M. (2019). Narrazioni tossiche e dialogo interculturale. *MeTis - Mondi educativi. Temi, indagini, suggestioni*: 9(2).
- Forzinetti, E., Della Vedova, M., Pasta, S., & Santerini, M. (2024). *Indicators for Characterising Online Hate Speech and Its Automatic Detection*. Cornell University, arXiv:2402.08462.
- Guillén-Nieto, V. (2023). *Hate Speech, Linguistic Perspectives*. Berlino: DeGruyter.
- James, G., Witten, D., Hastie, T., Tibshirani, R., et al. (2013). *An introduction to statistical learning*. vol. 112. Springer.
- James, Z., McBride, K. (2022). Examining the Contours of Hate: A Critical Hate Studies Analysis. In E. Hall, J. Clayton, & C. Donovan (Eds.), *Landscapes of Hate Tracing Spaces, Relations and Responses* (Vol. 0, pp. 22-37). University of Plymouth.
- Keen E., Georgescu M. (eds.) (2016). *Bookmarks. A manual for combating hate speech online through human rights education*. Revised edition. Strasbourg: Council of Europe.
- Mancini, M. (2021). Meme, video e immagini: comunicare nell'arena di Instagram. In S. Pasta, M. Santerini (eds.), *Nemmeno con un click. Ragazze e odio online* (pp. 125-131). Milano: FrancoAngeli.
- Marangi, M. (2021). Protagonismo giovanile e crossmedialità, per favorire la partecipazione sociale. In S. Pasta, M. Santerini (eds.), *Nemmeno con un click. Ragazze e odio online* (pp. 94-106). Milano: FrancoAngeli.
- Moscovici, S. (1989). *Le rappresentazioni sociali*. Bologna: Il Mulino.
- Muller, M., Guha, S., Baumer, E., Mimno, D., Shami, N. (2016) Machine Learning and Grounded Theory Method: Convergence, Divergence, and Combination. *Mathematics, Statistics and Computer Science Faculty Research and Publications*. 515.
- Palano, D. (2020). *Bubble democracy: la fine del pubblico e la nuova polarizzazione*. Brescia: Scholé Morcelliana.
- Pariser, E. (2012). *The Filter Bubble: What The Internet Is Hiding From You*. Londra: Editore Penguin.
- Pasta, S. (2018). *Razzismi 2.0. Analisi socio-educativa dell'odio online*. Brescia: Scholé Morcelliana.
- Pasta, S. (2021). Detection di odio antimusulmano tra machine learning e valutazione qualitativa. In Polenghi, S., Cereda, F., Zini, P. (eds.), *La responsabilità della pedagogia nelle trasformazioni dei rapporti sociali. Storia, linee di ricerca e prospettive*, Ebook Sessioni Parallele (pp. 1169-1179). Lecce-Rovato (BS): Pensa Multimedia.

- Pasta, S., Santerini, M., Forzinetti, E., Della Vedova, M. (2021). Antisemitism and Covid-19 on Twitter. The search for hatred online between automatism and qualitative evaluation. *Antisemitismo e Covid-19 in Twitter. La ricerca dell'odio online tra automatismi e valutazione qualitativa*. In *Form@re. Open Journal per formazione in rete*, XXI, n. 3.
- Pasta, S. (2021). Uno sguardo intersezionale: femmine e... In Santerini, M., Pasta, S. (eds.), *Nemmeno con un click. Ragazze e odio online* (pp. 38-50). Milano: FrancoAngeli.
- Pasta, S. (2022). L'odio online e il posizionamento della chiesa cattolica. In *Veritas et Jus*, 25(2), 85-105.
- Pasta, S. (2023a). L'antisemitismo 2.0 in Twitter, La detection dell'hatespeech online. In Santerini, M., *L'antisemitismo e le sue metamorfosi* (pp. 185-212). Firenze: Giuntina.
- Pasta, S. (2023b). Hate Speech Research: Algorithmic and Qualitative Evaluations. A Case Study of Anti-Gypsy Hate on Twitter. In *REM. Research on Education and Media*, 15(1), 130-139.
- Pasta, S. (ed.). (2024a). Lo "Spettro dell'Odio Online": Una proposta di classificazione tra valutazioni algoritmiche e qualitative. In Crescenza, G. (ed.). *Educare in tempi di odio e violenza. Sfide pedagogiche e istituzionali* (pp. 15-24). Bari: Progedit.
- Pasta, S. (2024b). Hate Studies tra logica computazionale e classificazione umana. Un caso studio sull'antisemitismo in Twitter. In *Scholè. Rivista di educazione e studi culturali*, LXII (1), 230-252.
- Quattrocioni, W., Vicini A. (2016). *Polarizzazioni*. Milano: FrancoAngeli.
- Ranieri, M. (2016). *Populism, Media and Education: Challenging discrimination in contemporary digital societies*. Londra: Routledge.
- Ronson, J. (2015). *So You've Been Publicly Shamed*. New York: Macmillan.
- Santerini, M. (2019). Discorsi d'odio sul web e strategie di contrasto. In *MeTis. Saggi-Essays. Mondi Educativi. Temi, indagini, suggestioni*, 9(2), 52.
- Santerini, M. (2021). *La mente ostile: forme dell'odio contemporaneo*. Milano: Raffaello Cortina.
- Santerini, M. (ed.). (2023). *L'antisemitismo e le sue metamorfosi* (pp. 185-212). Firenze: Giuntina.
- Santerini, M. (ed.). (2024). L'odio contemporaneo: sfide educative. In Crescenza, G. (ed.), *Educare in tempi di odio e violenza. Sfide pedagogiche e istituzionali* (pp. 15-24). Bari: Progedit.
- Schweppe, J., & Perry, B. (2022). A Continuum of Hate: Delimiting the Field of Hate Studies. *Crime, Law and Social Change*, 77, 503-528.
- Terry, C., Cain, J. (2016). The Emerging Issue of Digital Empathy. In *American Journal of Pharmaceutical Education*, 80(4), 58.
- Vitullo, A. (2020). *Educare onlife. Ridurre le diseguaglianze digitali per prevenire i discorsi d'odio*, Treccani.it, 9/10/2020.
- Volpato, C. (2022). *Deumanizzazione, come si legittima la violenza*. Bari: Editori Laterza.