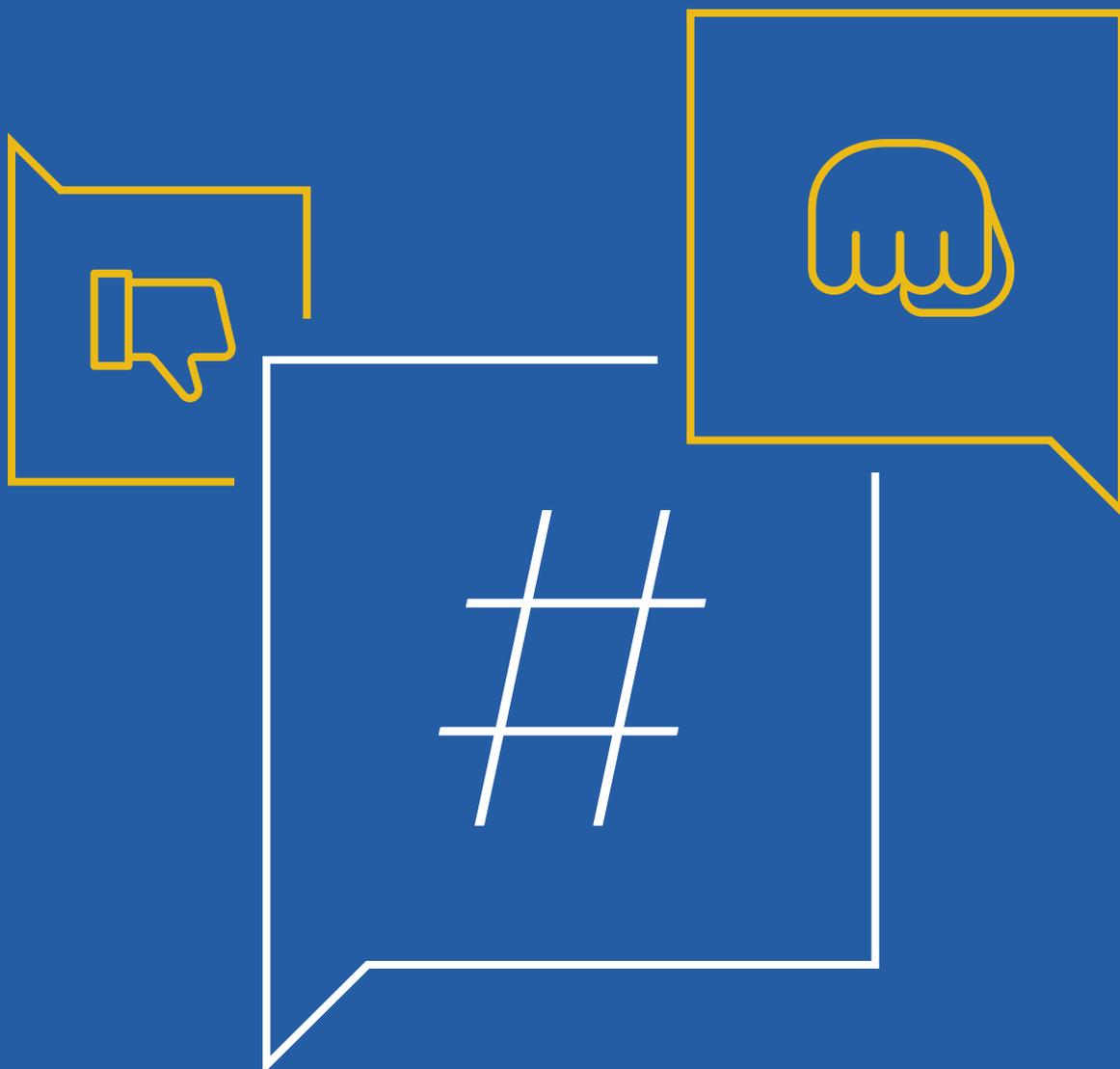


Barometro dell'odio nello Sport



UNIVERSITÀ
DEGLI STUDI
DI TORINO



L'hate speech online
nelle conversazioni
sportive su Facebook
e Twitter

RICERCA REALIZZATA DA

Giuliano Bobba, Moreno Mancosu, Antonella Seddone, Federico Vegetti, con la supervisione di Marinella Belluati, Centro CODER del Dipartimento di Culture, Politica e Società dell'Università di Torino nell'ambito del progetto Odiare non è uno Sport (AID 011797) promosso da CVCS e finanziato dall'Agenzia Italiana di Cooperazione allo Sviluppo

REVISIONE TESTI

Silvia Pochettino e Ilaria Leccardi

DESIGN

Luca Morano

Questo progetto è stato realizzato con il contributo della Agenzia Italiana per la Cooperazione allo Sviluppo all'interno del progetto Odiare non è uno Sport AID 011797.

I contenuti sono di esclusiva responsabilità dell'ente promotore del progetto e non rappresentano necessariamente il punto di vista dell'Agenzia.

AID 011797 COFINANZIATO DA:



PROMOSSO DA:



IN COLLABORAZIONE CON:



INDICE

5	Introduzione
	CAPITOLO 1
8	L'hate speech e i social media
	APPROFONDIMENTO
14	La ricerca scientifica e l'hate speech
	CAPITOLO 2
16	L'hate speech e lo sport
	CAPITOLO 3
19	Come e dove si è svolta la ricerca
	CAPITOLO 4
26	Hate speech nello sport: un fenomeno diffuso
	CAPITOLO 5
31	Gli sport non sono tutti uguali sui social media e in relazione all'hate speech
	CAPITOLO 6
41	Focus sul calcio: le squadre e l'hate speech
	CAPITOLO 7
50	Focus sul calcio: i personaggi sportivi e l'hate speech
	CAPITOLO 8
60	Distribuzione nel tempo dell'hate speech
	CAPITOLO 9
63	Interazioni tra utenti e hate speech
	CAPITOLO 10
65	Conclusioni
67	Riferimenti bibliografici
68	Appendice: metodo e descrizione del dizionario utilizzato



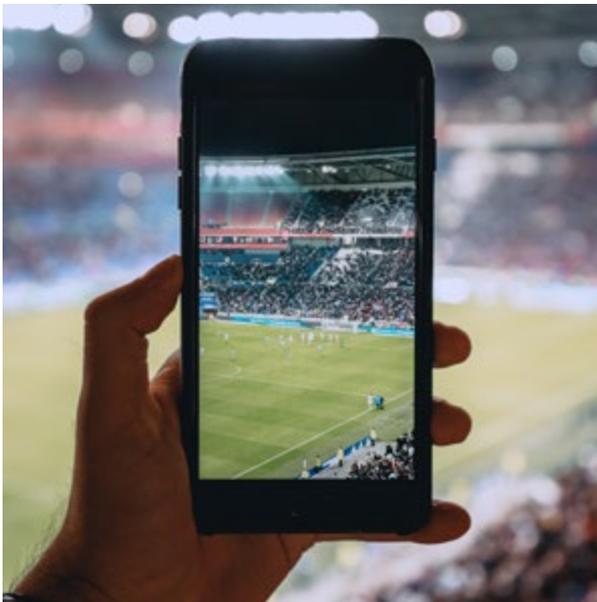
Introduzione

Quanta volgarità, minacce e insulti anche a sfondo razziale o sessista sono presenti nelle discussioni online che parlano di sport?

È ormai un sentimento comune che, malgrado gli innumerevoli pregi, i social media siano diventati uno dei luoghi in cui maggiormente avvengono scontri verbali, anche accesi, forme più o meno velate di discriminazione e attacchi che possono generare vere e proprie tempeste di offese (le cosiddette *shitstorm*). Le regole e i codici sociali appaiono meno rigidi online e dunque comportamenti, espressioni, parole e modalità che in ambiente offline sarebbero biasimati (o comunque comporterebbero una sanzione sociale) divengono legittimi. Va sottolineato però che **anche se i discorsi d'odio si articolano su una dimensione online, il loro effetto è pienamente e compiutamente reale.**

Lo sport a sua volta presenta una duplice veste: da un lato, è spesso uno strumento di integrazione e trasmissione di valori, soprattutto quando praticato; dall'altro lato, specialmente per quanto riguarda la dimensione del tifo, lo sport è invece un elemento divisivo che inasprisce la competizione fino a trasformarla in conflitto tra individui che sostengono fazioni opposte.

Questa ricerca – che si colloca nel quadro più ampio delle attività di documentazione e contrasto di fenomeni di odio e discriminazione online realizzate dal progetto Odiare non è uno sport, finanziato dall'Agenzia Italiana per la Cooperazione allo Sviluppo (AID 011797) e promosso da CVCS insieme a 13 partner nazionali – intende fornire un quadro ampio e dettagliato sul fenomeno dell'*hate speech* analizzando le conversazioni che si sviluppano a partire dalle notizie sportive pubblicate dalle principali testate giornalistiche sportive italiane (*La Gazzetta dello Sport*, *Tuttosport*, *Il Corriere dello Sport*, *Sky Sport* e *Sport Mediaset*) su Facebook e Twitter.



Dopo una prima parte introduttiva che definisce i confini del problema e chiarisce fonti e periodo analizzati, il report si concentra sulla diversa diffusione che l'*hate speech* ha sui due social media, tra i diversi sport, tra le principali squadre italiane di calcio, tra i principali giocatori, nonché la sua distribuzione nel tempo e nei diversi tipi di interazione tra utenti.

Il primo risultato che salta agli occhi è che esiste un livello costante di *hate speech* al di sotto del quale non si scende, pari al 10,9% dei commenti su Facebook e 18,6% su Twitter.

I messaggi d'odio possono quindi a tutti gli effetti considerarsi come una componente strutturale delle conversazioni sportive su questi social media.

Tuttavia Facebook e Twitter si caratterizzano per essere due ambienti diversi in termini di numero di commenti: a parità di messaggi pubblicati **Facebook genera un volume di commenti 26 volte superiore a quello di Twitter**. I due social media hanno anche **livelli diversi di *hate speech***: su Twitter la presenza è percentualmente più elevata, ma la quantità di commenti contenenti odio bassa; su Facebook la presenza percentuale è più bassa, ma il numero di commenti con riferimenti all'*hate speech* elevato.

Il 13,4% dei commenti su Facebook e il 31% di quelli su Twitter contiene *hate speech*. Di questi, la maggior parte non si riferisce al cuore della definizione di *hate speech* (vedi pag 17), ma riguarda invece l'uso di **linguaggio volgare** (14% e 31%) o di **aggressività verbale** (73% e 60%). Comparativamente, i riferimenti a **discriminazione** (7% e 5%) e **aggressività fisica** sono minori (5% e 4%).

Questo non vuol dire tuttavia che siano una componente irrilevante poiché si tratta di circa **5.000 commenti di questo tipo pubblicati dagli utenti in un arco di tre mesi**.

Un'altra evidenza che si ricava dai dati esaminati è la **sovrarappresentazione del calcio nel flusso comunicativo** e anche nel volume totale di commenti. Basket, Formula 1 e tennis sono gli unici altri sport con un volume di commenti sufficiente per essere analizzato. Anche per questo motivo **il calcio è evidentemente lo sport a cui è associato il numero più alto di commenti di *hate speech***: 12,3% su Facebook e 31,3% su Twitter.

Le sei le squadre di calcio che riportano il maggior numero di commenti di *hate speech* risultano essere **Milan, Lazio, Roma, Juventus, Inter, Napoli** e tra queste non emergono grandi differenze nella tipologia di linguaggi d'odio; sono prevalentemente associate all'aggressività verbale e solo in percentuali minime ad aggressività fisica o discriminazione.



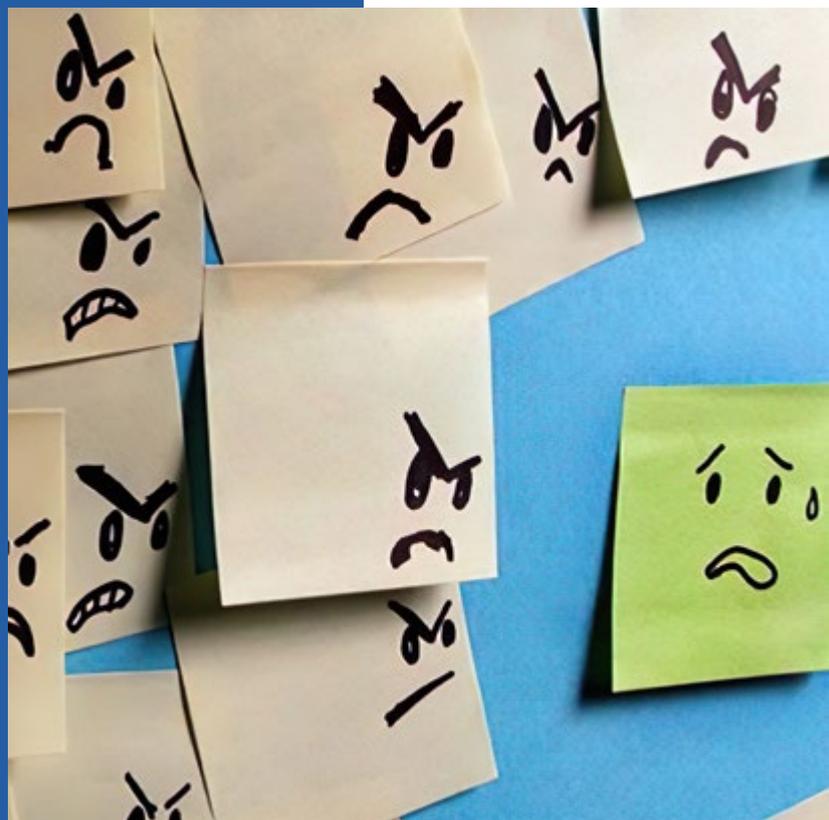
Mentre tra i personaggi sportivi **Mario Balotelli** e **Romelu Lukaku** spiccano come quelli su cui si rilevano livelli più alti di *hate speech* (rispettivamente 16,7% su Facebook e 38,3% su Twitter; 15,5% su Facebook e 40,6% su Twitter) e in particolare insulti contenenti discriminazione razziale per il loro colore della pelle (rispettivamente 2,1% su Facebook e 5,6% su Twitter; 1,9% su Facebook e 2,4% su Twitter).

Infine è interessante notare come su Facebook, dove è possibile commentare i commenti degli altri, ***l'hate speech* risulta più elevato nelle discussioni tra utenti rispetto ai commenti ai post**. Ovvero il maggior numero di riferimenti al linguaggio d'odio non risulta correlato al contenuto dei post ma piuttosto alle prese di posizione di altri utenti.



CAPITOLO 1

L'hate speech e i social media



Di *hate speech* si parla sempre più frequentemente, e con una certa apprensione. Tuttavia **dare una definizione precisa di cosa s'intenda per *hate speech* è questione tutt'altro che semplice.**

Se con questa espressione ci si riferisce (letteralmente) a "discorsi di incitamento all'odio", chiarire invece cosa rientri in questa categoria risulta più problematico. Entra in gioco il dibattito sulla libertà di espressione (Yong 2011) e sancire cosa contribuisca a definire un contenuto testuale o verbale come *hate speech* diventa materia di discussione.

Non è un caso se anche da un punto di vista legale si sia faticato non poco a giungere a una definizione condivisa in grado di circoscrivere precisamente il perimetro di questo fenomeno.

A livello internazionale la fonte di riferimento che per prima bandisce ogni forma di odio è il Patto internazionale relativo ai diritti civili e politici, adottato a New York nel 1966 (esecutivo in Italia nel 1977, legge n. 881). In quel testo si vietano appelli all'odio nazionale, razziale o religioso, che incitino la discriminazione, l'ostilità, la violenza e la guerra.

È solo in anni più recenti che si pone il problema di aggiornare questa definizione e soprattutto di offrire un riferimento normativo comune. In questo senso, il Comitato dei ministri del Consiglio d'Europa ha definito come *hate speech* quelle espressioni che diffondono, incitano, promuovono o giustificano l'odio razziale, la xenofobia, l'antisemitismo o più in generale l'intolleranza, ma anche i nazionalismi e gli etnocentrismi, gli abusi e le molestie, gli epiteti, i pregiudizi, gli stereotipi e le ingiurie che stigmatizzano e insultano (Raccomandazione n. 20 del Comitato dei ministri del Consiglio d'Europa del 30 ottobre 1997).

Sulla stessa linea, l'Unione Europea nel 2008 ha indicato come *hate speech* quelle espressioni verbali che pubblicamente incitano alla violenza o minacciano direttamente singoli o gruppi di persone in ragione di elementi legati alla razza, alla religione e alle origini etniche (*Council Framework Decision 2008/913/JHA of 2008*). È solo più recentemente che la definizione di *hate speech* da un punto di vista legale allarga la prospettiva.

Qual è la definizione di discorso d'odio?

È definibile come tale la promozione, l'incitamento alla denigrazione, l'attacco o la mortificazione di un singolo o di un gruppo di persone come pure la molestia, l'insulto, la stigmatizzazione, l'utilizzo di stereotipi negativi o la minaccia di persone sulla base di caratteristiche riferite alla razza, al colore della pelle, alla religione, alle origini nazionali o etniche come pure all'età, la disabilità, il genere, l'identità di genere o l'orientamento sessuale.

2016, European Commission against Racism and Intolerance (ECRI) - N. 15 - 8.12.2016

Si apre dunque il ventaglio delle categorie a rischio di *hate speech* riconosciute all'interno di documenti ufficiali ed è importante ricordare che l'ECRI sottolinea che la lista non è esaustiva. L'*hate speech*, dunque, viene riconosciuto come un **fenomeno ampio e composito**. E sempre nel 2016, la Commissione europea riconosce internet e i social media come un ambiente in cui questo tipo di narrative ed espressioni che incitano all'odio trovano spazio ed enfasi. Il 31 Maggio 2016 viene dunque diffuso un **Codice di condotta per il contrasto dell'*hate speech*** (Code of Conduct on Countering Illegal *Hate speech*): Facebook, YouTube, Twitter, Microsoft e Instagram sono chiamati in causa nella lotta all'*hate speech*.



In questa sede, non è importante entrare nel dettaglio del framework legale che norma il contrasto alle forme di *hate speech*, ma semmai ci preme sottolineare che, nonostante a livello internazionale esistano delle linee guida ufficiali e il tentativo di trovare una definizione comune del fenomeno trovi spazio in risoluzioni e policy regulation, quando si entra nel dettaglio delle normative nazionali il tema si fa ancora più complesso e si registra una grande eterogeneità. Sono molte le ricerche che sottolineano come il fenomeno dell'*hate speech* e la sua definizione – dunque la sua individuazione e contrasto – risentano di specificità culturali e normative nazionali.

In Italia la legislazione più rilevante in merito è la L. 205/1993 che definisce crimine la propagazione di idee basate su concetti di superiorità razziale o odio etnico, o che istighino a commettere atti di discriminazione razziale, etnica o religiosa. La legge, inoltre, punisce anche chi istighi o commetta atti di violenza o provocazione di violenza sulla base di motivazioni razziali, etniche o religiose. Il caso italiano è interessante da questo punto di vista, i media infatti per legge non possono trasmettere contenuti che presentino incitamento all'odio o comportamenti intolleranti sulla base di razza, genere, religione o nazionalità.

Ma – come è facile intuire – **la normativa italiana non tiene conto dell'evoluzione tecnologica** e, quando si tratta di gestire fenomeni di *hate speech* su ambienti online, la questione diventa più complessa ed elaborare strategie di contrasto più difficile.

Soprattutto, in Italia non esiste ancora una definizione normativa specificamente focalizzata sull'*hate speech* al punto che **in questa legislatura è stata istituita un'apposita Commissione parlamentare – la cosiddetta Commissione Segre – al fine di fornire strumenti legislativi di indirizzo e controllo** sui fenomeni di "intolleranza, razzismo, antisemitismo e istigazione all'odio e alla violenza nei confronti di persone o gruppi sociali sulla base di alcune caratteristiche quali l'etnia, la religione, la provenienza, l'orientamento sessuale, l'identità di genere o di altre particolari condizioni fisiche o psichiche".

Gli effetti di messaggi violenti e offensivi non sono circoscritti all'effimero e impalpabile ambiente online. Si parte da lì, ma **l'impatto è reale**, riguarda gli individui nella loro vita quotidiana, quella offline.

In altre parole, il legislatore punisce azioni di intolleranza e incitamento alla violenza ma fallisce nel circoscrivere il fenomeno e affrontarlo negli ambienti online dove oggi trova espressione e manifestazione più rilevante. Ed è proprio la dimensione online dell'*hate speech* a meritare particolare attenzione.

Siamo inclini a circoscrivere la diffusione dei discorsi d'odio a una dimensione online, laddove internet e più precisamente i social media diventano l'ambiente in cui il fenomeno emerge con maggiore evidenza. Tuttavia le dimensioni online e offline sono strettamente legate.

La ricerca scientifica ha raccontato con grande efficacia come i social media si configurino come un ambiente particolarmente confortevole per i messaggi d'odio e, data la loro diffusione e pervasività nel vissuto quotidiano delle persone, diviene estremamente urgente capirne di più, misurando il volume di questi contenuti e comprendendo quali siano le modalità di espressione (attraverso l'individuazione delle categorie più a rischio, vittime e target prediletti di parole d'odio che rimandano a violenza, aggressività e offese).

Nell'ambiente online le relazioni fra gli individui sono caratterizzate da una diluizione dello spazio-tempo.

Chiunque in possesso di un accesso internet può intervenire, offrire la propria opinione, contribuire alla conoscenza e alla creazione di una cultura condivisa, una "cyberculture" (Sacristán 2013: 126). Tuttavia l'interazione online prevede un **isolamento dalle convenzioni sociali**, e questo indurrebbe un **meccanismo di disinibizione** rendendo gli individui più inclini a esprimere disaccordo rispetto alle opinioni altrui (Santana 2014).

L'ambiente online è lo spazio in cui l'espressione e la diffusione di idee o credenze intolleranti trovano più **facile e massiccia diffusione.**

(Assmakopoulos et al. 2017).

Surel (2004) sottolinea che questo meccanismo può avere esiti positivi nel momento in cui contribuisce all'espressione di opinioni, ma può produrre effetti perversi quando l'ambiente online è percepito come separato da una dimensione reale offline, un'arena dove i codici comportamentali classici possono essere sovvertiti, o semplicemente non valgono e le relazioni sociali seguono dinamiche differenti. E questa percezione di dissociazione fra online e offline può essere favorita dalla possibilità di utilizzare l'anonimato nelle interazioni sul web o sui social media.

Gli utenti possono utilizzare questo spazio per esprimere i loro punti di vista e le loro opinioni e l'ambiente stesso offre la possibilità di farlo giocando sull'anonimato, proiettando sulle loro interazioni online una erronea e falsata idea di impunità.

La natura globale, immediata e partecipativa di internet è dunque contemporaneamente il punto di forza ma anche la sua debolezza.

Internet è certamente una delle più rilevanti innovazioni tecnologiche degli ultimi decenni. E più recentemente, i social media hanno rimodellato le modalità di interazione fra gli individui offrendo un ambiente dove le distanze spaziali possono essere facilmente superate e annullate.

La connessione e l'interazione fra gli individui sono istantanee e si muovono seguendo logiche del tutto nuove. I linguaggi e i codici comportamentali assumono connotazioni differenti e si ridefiniscono secondo dinamiche diverse rispetto a quelle offline.



Le regole e i codici sociali appaiono meno rigidi e dunque comportamenti, espressioni, parole e modalità che in ambiente offline sarebbero biasimati (o comunque comporterebbero una sanzione sociale) divengono legittimi. Il significato delle parole e il loro impatto sembrano più blandi perché confinati in uno spazio etereo, l'online. Al contrario, e qui si situa la crescente preoccupazione delle istituzioni e l'interesse della ricerca scientifica sul tema, l'*hate speech* ha un peso e una rilevanza molto reale e un impatto preciso sul vissuto degli individui che ne sono vittime. I discorsi d'odio si articolano su una dimensione online, ma il loro effetto è invece pienamente e compiutamente reale.



APPROFONDIMENTO

La ricerca scientifica e l'*hate speech*



La ricerca scientifica ha cercato di comprendere il fenomeno sotto punti di vista differenti, ma mettendo al centro lo spazio online e i social media in particolare come ambienti dove i discorsi d'odio trovano non solo spazio, ma anche risonanza e circolazione più capillare (Andrisani, 2014; Ziccardi, 2016; Hopp e Vargo, 2017). Il tema viene affrontato riconducendo l'*hate speech* all'interno della più ampia categoria di *incivility*.

Ad esempio, alcuni studi si sono interessati a comprendere la relazione fra individui e l'utilizzo di forme di *hate speech* o linguaggio aggressivo nell'ambiente online (Hmielowski et al. 2014), altri lavori hanno invece esplorato l'effetto dei commenti d'odio sulla propensione e capacità dei soggetti di comprendere efficacemente le informazioni offerte dai media (Anderson et al. 2014).

Questi lavori di ricerca aiutano a definire in maniera più precisa cosa si intende per *hate speech*, calando la definizione sull'ambiente online e aiutando a circoscrivere più precisamente il fenomeno laddove il legislatore e le istituzioni hanno lasciato zone di ambiguità.

Coe, Kenski and Rains (2014) ad esempio nel loro studio sull'utilizzo di discorsi d'odio all'interno dei forum dei media online hanno definito questo fenomeno come «una discussione che presenta toni irrispettosi sui temi affrontati e nei confronti dei soggetti coinvolti». Introducono un ulteriore elemento, però, definendo l'*hate speech* un contenuto irrispettoso che irrompe in maniera superflua, gratuita, non necessaria. Antoci et al. (2016) considerano invece espressione di *incivility* online quelle dinamiche di interazione che comprendono commenti aggressivi all'interno di discussioni rabbiose o critiche sgarbate, ma anche affermazioni offensive, parole d'odio o moleste.

Santana (2014) offre una definizione più precisa e focalizzata sul contenuto, suggerendo di considerare *hate speech* i contenuti che possiedono almeno uno dei seguenti elementi:

- | | | | |
|---|--|---|--|
| ① | insulto | ⑦ | sentimenti di razzismo o intolleranza |
| ② | minaccia | ⑧ | commenti denigratori sulla base di razza o etnia |
| ③ | volgarità | ⑨ | uso di stereotipi negati |
| ④ | linguaggio offensivo o turpiloquio | | |
| ⑤ | xenofobia | | |
| ⑥ | linguaggio aggressivo, epiteti o insulti | | |

Nel contesto italiano, Tullio De Mauro (2016: 2) ha definito invece le *hate words* come «parole per ferire, quelle che provocano dolore perché sono diventate dispregiativi per natura. Sono le parole peggiori che si possano usare, soprattutto se si appartiene a un gruppo che esercita il potere su un altro perché costituisce una minoranza o perché ha alle spalle una lunga storia di discriminazione, gli eterosessuali lo esercitano sugli omosessuali, i bianchi sulle minoranze razziali, gli uomini sulle donne, i cristiani sui fedeli di altre religioni, le persone cosiddette normali sulle persone con disabilità, e così via».



CAPITOLO 2

L'hate speech e lo sport



L'hate speech in ambito sportivo ha oggi assunto particolare rilevanza. In particolare lo sport più diffuso in Italia, il calcio, rappresenta non solo un'occasione di integrazione, nella sua forma giocata, ma anche un terreno di scontro che contrappone persone appartenenti a diverse tifoserie (Valeri, 2010).

Queste manifestazioni di parte che spesso sfociano in violenza verbale o fisica godono di ampia visibilità quando avvengono in occasione di partite di cartello come derby o competizioni internazionali. Se nella memoria collettiva resta scolpito il ricordo degli Hooligans inglesi degli anni '80 o, in anni più recenti, dell'ultras serbo Ivan Bogdanov, anche nel calcio nostrano non mancano esempi di violenza. Basti pensare agli scontri, in cui morì un tifoso interista, prima della partita Inter-Napoli nel novembre 2018. Anche i campionati minori non sembrano immuni da questo tipo di dinamiche che a volte degenerano in maniera ancora più violenta anche per la mancanza di prevenzione.

Le cronache più recenti riportano scene di guerriglia urbana con numerosi feriti in occasione della partita Albanova-Afragolese valida per il campionato di Eccellenza (15 febbraio 2020).

Questa frequente situazione di scontro porta a un processo di **normalizzazione della violenza** per cui è ormai opinione piuttosto diffusa che gli stadi e il tifo siano intrinsecamente connessi a forme di violenza e sopraffazioni.

30%

degli Italiani ritiene che sia **pericoloso** andare a vedere una partita perché **c'è troppa violenza**

51° Osservatorio sul Capitale Sociale degli Italiani: il tifo calcistico degli Italiani, 2016

L'equazione tra calcio e violenza sembra avere due conseguenze.

Da un lato una parte di tifosi – il 15% per cento – afferma di essere stufo della corruzione, delle scommesse e del mondo ultrà al punto da **allontanarsi dal calcio** e dichiararsi ex-tifoso.

Sebbene il calcio rimanga la principale passione degli italiani è infatti insidiato sempre più da vicino da altri sport, meno facili da praticare, ma capaci di coinvolgere sempre più persone: la Formula 1, il motociclismo, la pallavolo o il nuoto (TABELLA 1).

Dall'altro lato, la conseguenza di un livello diffuso di violenza nel calcio genera un processo che porta alcune persone interessate a questo sport a modellare il proprio comportamento per **imitazione o identificazione** rispetto a un ambiente che più di altri tollera le discriminazioni, accetta forme di prevaricazione, banalizza e deride il riconoscimento dell'altro. La conseguenza è quella di considerare come accettabili anche condotte riprovevoli o discriminatorie poiché applicate al contesto calcistico.

In questo contesto, **l'avvento del web e dei social media ha offerto un nuovo terreno di contrapposizione ai tifosi delle diverse fazioni**, non più obbligati a recarsi in uno stadio o nei pressi di un campo da gioco, per ostentare la propria fede sportiva.

Il problema è che il tifo ha una duplice componente: una positiva, di sostegno alla propria squadra o al proprio sportivo favorito, e una negativa, di denigrazione e dilleggio della parte avversa.

La componente positiva del tifo – benché combinata a quella negativa – trova il suo naturale spazio in prossimità spaziale e concomitanza temporale con l'evento sportivo perché è principalmente rivolta a qualcuno che può sentirla (gli atleti), nel momento in cui il tifoso pensa possa avere una qualche rilevanza (la gara).

Al contrario, la componente negativa trova la sua maggiore diffusione al di fuori del contesto sportivo perché è principalmente rivolta verso altri tifosi: mossi da un senso di appartenenza, la derisione, l'insulto, la minaccia non hanno bisogno di un luogo e di un momento specifico strettamente definito (la gara). Al contrario sono validi sempre, in un perpetuarsi anche al di fuori della competizione sportiva di una logica noi vs. loro.

L'ambiente dei social media - che consente di avere, in un unico flusso informazione, commenti e interazioni tra utenti - ha finito col creare le condizioni perché la componente negativa del tifo trovasse terreno fertile per prosperare e propagarsi con una visibilità senza precedenti.

TABELLA 1

Per quale sport tifano gli italiani?

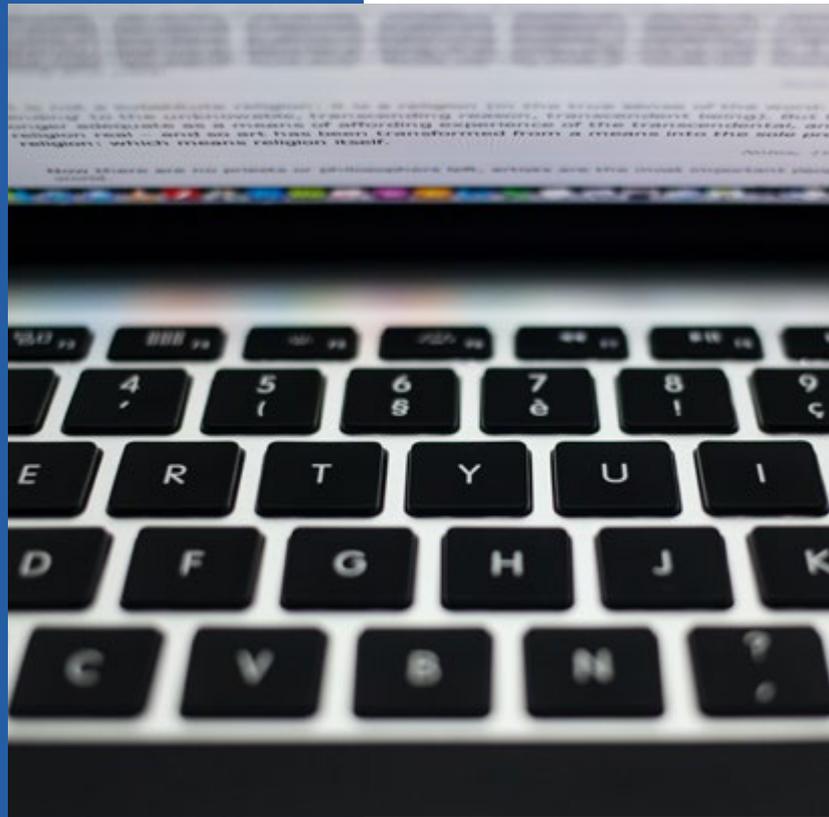
Sport	%
Calcio	38
Motociclismo	35
Formula 1	34
Pallavolo	32
Nuoto	32
Atletica	30
Ciclismo	25
Basket	21
Tennis	21
Rugby	11

(Fonte: Demos & Pi, 2016)



CAPITOLO 3

Come e dove si è svolta la ricerca



PARAGRAFO 3.1

Una definizione operativa di *hate speech*

Come accennato in precedenza, il termine *hate speech* ha avuto grande diffusione negli ultimi anni. Nel linguaggio comune, lo si utilizza per indicare il generico ricorso a un linguaggio volgare, offensivo, violento o discriminante.

Concentrando l'analisi sui social media e sugli ambienti digitali, abbiamo adottato una definizione ampia e articolata identificando quattro principali dimensioni che muovono da un livello più basso e non direzionato a un livello più alto e direzionato di *hate speech* (TABELLA 2).

① Linguaggio volgare

La prima dimensione riguarda l'utilizzo di linguaggio volgare. Pur non essendo una forma di istigazione d'odio in senso stretto, il ricorso ad un linguaggio volgare è stato considerato come un primo indicatore di una interazione che avviene al di fuori o ai confini dei requisiti base di una conversazione civile. Non a caso nella letteratura accademica, accanto all'espressione *hate speech*, ha fatto la sua comparsa anche il termine *incivility*, che include un ricorso generico a parolacce ed espressioni volgari non rivolte specificatamente contro un altro soggetto.

② Aggressività verbale

La seconda dimensione si riferisce all'aggressività verbale. In questo caso parolacce e insulti sono direzionati, vale a dire utilizzati con il fine di offendere un altro soggetto verbalmente. Questo tipo di linguaggio può rivolgersi ad almeno due diversi target. Da un lato, l'offesa può essere indirizzata verso la squadra, lo sportivo, il giornalista, o un altro soggetto passivo, con cui non può esserci interazione. Dall'altro lato, invece, questo tipo di linguaggio può essere rivolto verso uno o più utenti attivi che reagiscono e interagiscono tra loro, generando una spirale di *hate speech*.

③ Aggressività fisica

La terza dimensione implica (l'evocazione di) aggressività fisica. Si tratta, in buona sostanza, di una versione più violenta della dimensione precedente che implica minacce direzionate. Anche in questo caso, possono essere genericamente rivolte verso un soggetto passivo oppure verso un utente che può prendere parte alla conversazione. Questo tipo di aggressività si distingue dalla precedente per l'intenzione di minacciare un altro soggetto facendo riferimento alla sua incolumità, non soltanto a quella mentale e psichica.

④ Discriminazione

La quarta dimensione, infine, coglie in senso stretto l'essenza dell'*hate speech* poiché identifica quelle espressioni di discriminazione che sono volte ad offendere un altro soggetto sulla base delle sue caratteristiche quali etnia, religione, genere, orientamento sessuale, condizioni fisiche o psichiche o semplicemente per la sua diversità rispetto a modelli dominanti, considerati normali.

TABELLA 2

La composizione dell'*hate speech*: definizioni ed esempi

Definizione delle dimensioni di <i>hate speech</i>	Esempi
<p>① Linguaggio Volgare</p> <p>Uso generico di parolacce ed espressioni volgari non rivolte a un altro soggetto</p>	<p>«Anche questa è una stagione di merda» «Uno che diceva che quest'anno avrebbe pisciato in testa a tutti» «Non puoi farci niente se guidi una moto del cazzo»</p>
<p>② Aggressività verbale</p> <p>Uso di parolacce e insulti volti ad offendere un altro soggetto verbalmente</p>	<p>«Siete peggio del Beer Sheva! pagliacci sempre» «Stronzo Rubentino» «Da oggi non tifo più... fate schifo, andate affanculo»</p>
<p>③ Aggressività fisica</p> <p>Uso di parolacce e insulti volti ad offendere e/o minacciare un altro soggetto fisicamente</p>	<p>«Gli darei fuoco a tutti! Non ha senso cacciarne uno scarso per comprarne uno ancora più scarso» «So dove abiti, ti aspetto sotto casa» «Per quelli come te una sola cosa: calci in culo!»</p>
<p>④ Discriminazione</p> <p>Uso di parolacce e insulti volti ad offendere un altro soggetto a causa della diversità rispetto a modelli considerati normali</p>	<p>«Questa è una squadra di ritardati» «Buffone zingaro» «Lo sanno tutti che Cristiano è frocio...»</p>



PARAGRAFO 3.2

Una definizione operativa di *hate speech*

I social media sono indubbiamente il mezzo di comunicazione che contraddistingue l'epoca attuale. Su piattaforme come Facebook, Twitter o Instagram, le interazioni avvengono attraverso flussi comunicativi individuali o collettivi, mescolando la logica one-to-one con la logica *one-to-many* e *many-to-many*. Gli utenti, in altre parole, possono usare i social media per inviare messaggi privati ad altri utenti, oppure conversare all'interno di gruppi, pagine pubbliche o sulle proprie bacheche non solo con i loro contatti, ma anche con persone che non conoscono. Quest'ultimo elemento è particolarmente rilevante per quanto riguarda l'*hate speech*, perché questa **assenza di conoscenza reciproca** – mi rivolgo a qualcuno che non conosco e che non mi conosce di fronte a un pubblico che non mi conosce e che non conosco – favorisce la **mancanza di sanzioni sociali**, presenti invece nella vita reale.

Dal momento che questa ricerca si concentra sulla parola scritta – verificando la presenza di *hate speech* nei commenti – la scelta è caduta su due social media che prediligono questo tipo di interazione tra gli utenti, vale a dire Facebook e Twitter.



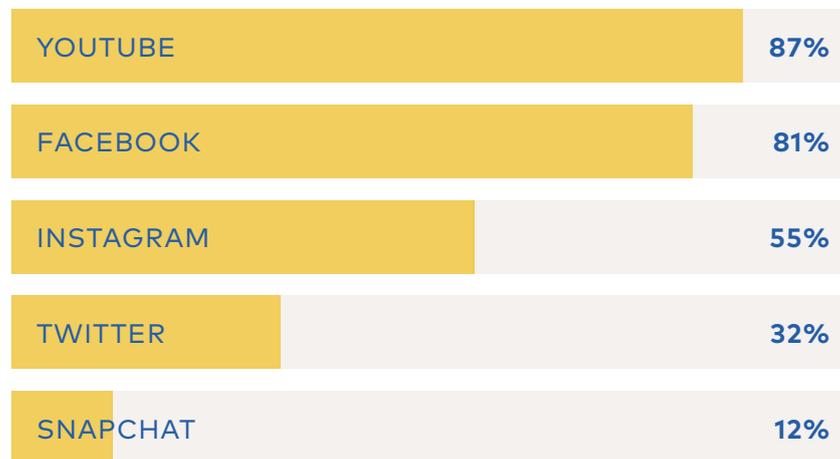
Mentre altri social sono incentrati sulla condivisione di immagini (Instagram, Snapchat, etc.) o video (Youtube, Tik Tok, etc.), Facebook e Twitter, pur avendo la possibilità di condividere contenuti multimediali, prevedono che l'interazione tra utenti avvenga principalmente attraverso le conversazioni scritte che si sviluppano attraverso i commenti ai post.

Facebook, come si sa, ha un'attitudine più generalista che lo ha portato a essere il social media più diffuso al mondo con una media di 2.271 milioni di utenti attivi al mese (Digital 2019).

Anche in Italia si conferma tra i social media più utilizzati, secondo solo a Youtube (FIGURA 1). Tuttavia, mentre gli utenti di Youtube sono prevalentemente passivi – accedono per vedere contenuti e non per caricare propri video – Facebook è di gran lunga la prima piattaforma per quanto riguarda gli utenti che attivamente contribuiscono al flusso comunicativo attraverso la pubblicazione di post, reazioni e commenti.

FIGURA 1

Percentuale di utenti che dichiarano di utilizzare i social media in Italia



(Fonte: Digital 2019)

Twitter segue piuttosto distaccata, dal momento che solo il 32% dichiara di usarlo abitualmente. Tuttavia rappresenta un ambiente peculiare – che merita di essere studiato – perché è luogo di incontro e scambio tra le élites tradizionali (giornalisti, personaggi famosi, politici, etc.) e le non-élites rappresentate sia da minoranze impegnate nella promozione della propria causa (sociale, politica, culturale, sportiva, etc.) sia da singoli cittadini particolarmente attivi online. Come si sa, la struttura di Twitter è meno libera di quella di Facebook, tra le altre cose perché limita il numero di caratteri utilizzabili in ogni tweet (inizialmente 140, attualmente 280 caratteri) e offre una gamma semplificata di commenti rispetto a Facebook: si può soltanto commentare il tweet originale, mentre su Facebook si possono commentare i commenti di altri utenti, funzionalità che favorisce la conversazione tra gli individui.



PARAGRAFO 3.3

Periodo e fonti

Dovendo coprire un'ampia gamma di eventi sportivi, l'analisi ha riguardato un periodo di tre mesi dal **7 ottobre 2019 al 6 gennaio 2020**. Durante questo periodo si sono disputati i campionati di calcio, pallavolo, basket (con cadenza almeno settimanale); ci sono stati cinque gran premi di Formula 1 e cinque di motociclismo, le finali ATP di tennis, gli europei di nuoto, i mondiali di Atletica e le vicende giudiziarie che riguardano il marciatore Alex Schwazer.

In prima battuta, la ricerca ha preso in esame tutti i post e i tweet pubblicati sugli account ufficiali dai principali mezzi di informazione sportivi italiani, vale a dire ***La Gazzetta dello Sport, Tuttosport, Il Corriere dello Sport, Sky Sport e Sport Mediaset***. In totale i post analizzati su Facebook sono stati 4.857, mentre su Twitter 4.554. Per ognuno di questi post e tweet sono stati quindi analizzati, tramite un dizionario sviluppato ad hoc, tutti i commenti ricevuti al fine di individuare l'eventuale presenza di *hate speech*. **Il totale dei commenti analizzati è stato pari a 443.567 su Facebook e 16.991 su Twitter**. La marcata differenza nel volume di commenti è principalmente dovuta alla disparità tra utenti attivi quotidianamente sulle due piattaforme.

Le TABELLE 3 E 4 mostrano la distribuzione di post e commenti tra le varie fonti informative analizzate. Su Facebook, *La Gazzetta* e *Sky Sport* spiccano per essere tra le più prolifiche in termini di post pubblicati e anche di commenti ricevuti. Una situazione diversa caratterizza invece Twitter dove *La Gazzetta* pubblica un volume di tweet più che doppio rispetto alle altre testate, risultando anche prima – seppure con livelli più in linea con le concorrenti – per volume medio di commenti.

TABELLA 3

Numero di post analizzati e commenti medi per post su Facebook

Testata	Post	Commenti medi per post
Corriere dello Sport	1.082	46,0
Gazzetta dello Sport	1.069	116,9
Sky Sport	1.006	155,6
Sport Mediaset	750	59,6
Tuttosport	950	71,1

TABELLA 4

Numero di tweet analizzati e commenti medi per tweet su Twitter

Testata	Post	Commenti medi per tweet
Corriere dello Sport	713	31,8
Gazzetta dello Sport	1.615	33,9
Sky Sport	753	14,1
Sport Mediaset	763	20,6
Tuttosport	710	12,6



CAPITOLO 4

Hate speech nello sport: un fenomeno diffuso



La presenza di *hate speech* legato allo sport sui social media può essere verificata, in prima battuta, guardando al numero di post che contengono espressioni riconducibili a una o più delle quattro dimensioni descritte in precedenza (FIGURA 2). Il quadro che emerge mostra alcune importanti differenze tra i due mezzi analizzati. Su Twitter, più della metà dei tweet non contiene nei commenti *hate speech*, mentre poco più di un tweet su quattro contiene un numero basso di commenti d'odio, compresi tra 1 e 10. Al contrario, **su Facebook** la distribuzione mostra un livello generale di *hate speech* più elevato: **solo un post su quattro non ha commenti d'odio**, più del 40% ne ha un numero basso (tra 1 e 10), il 15,1% dei post mostra livelli moderati (da 11 a 25) e il 13,6% registra livelli alti di *hate speech* (più di 25 commenti contenenti *hate speech*).

Un secondo indicatore che permette di valutare il livello di *hate speech* è la percentuale di insulti, aggressività e discriminazione presenti sul totale dei commenti pubblicati dagli utenti in relazione ai post analizzati (FIGURA 3). In questo caso vediamo che la situazione si ribalta.

Su Facebook nel 13,4% dei commenti ai 4.857 post analizzati è possibile individuare almeno un riferimento di *hate speech*, pari a 59.375 commenti su un totale di 443.567 analizzati. **Su Twitter**, invece, la presenza di *hate speech* è molto più evidente: **ogni dieci commenti, almeno tre risultano contenere riferimenti d'odio**, una percentuale più che doppia rispetto a Facebook.

L'immagine si completa guardando la rilevanza delle singole dimensioni che compongono l'*hate speech* sui due mezzi (FIGURE 4 E 5). Facebook è di gran lunga l'ambiente in cui si contano il maggior numero di insulti: tre messaggi d'odio su quattro contengono aggressività verbale.

Oltre a ciò, anche la componente più hard dell'*hate speech*, vale a dire l'aggressività fisica e la discriminazione, raggiunge percentuali significative pari rispettivamente a 6 e 7 percento. Anche Twitter presenta una distribuzione simile, ma il linguaggio volgare è comparativamente più elevato, al punto da superare il 90 percento se sommato all'aggressività verbale.

ANALISI DI FACEBOOK

4.857

post analizzati

443.567

commenti
analizzati

59.375

commenti con
hate speech

3 su 4

post ha ricevuto
commenti contenenti
linguaggi d'odio

ANALISI DI TWITTER

1 su 2

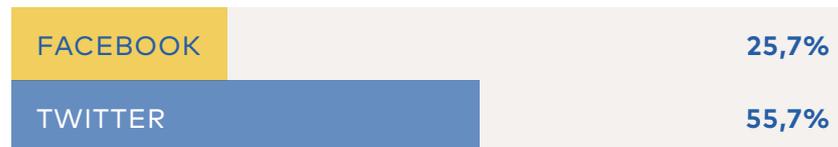
tweet che
contengono
riferimento d'odio

In sintesi, su Facebook è più facile imbattersi in un post che abbia commenti con *hate speech* rispetto a Twitter; i commenti di Facebook sono però molto più numerosi per cui la percentuale di commenti d'odio è comparativamente minore di quella di Twitter: se leggo dei commenti di un post sportivo scelto casualmente è più facile che io trovi *hate speech* su Twitter rispetto a Facebook; in entrambi gli ambienti, la percentuale di commenti che fa riferimento al cuore della definizione di *hate speech*, vale a dire la discriminazione, è una porzione bassa ma rilevante delle conversazioni tra utenti.

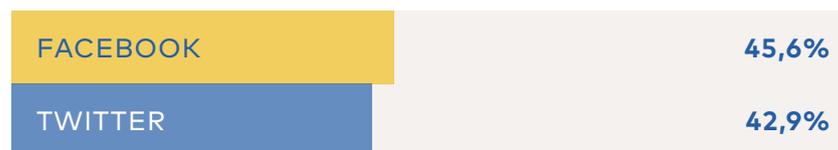
FIGURA 2

Percentuale di post con *hate speech* nei commenti

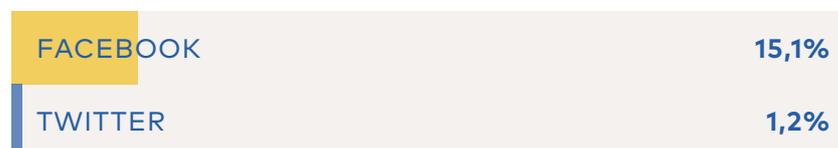
Nessun commento con hate speech



1-10 commenti con hate speech



11-25 commenti con hate speech



>25 commenti con hate speech

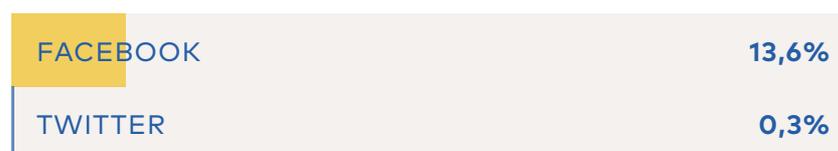
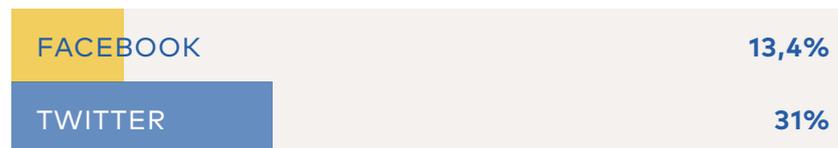


FIGURA 3

Percentuale di *hate speech* nei commenti

Hate speech



No hate speech



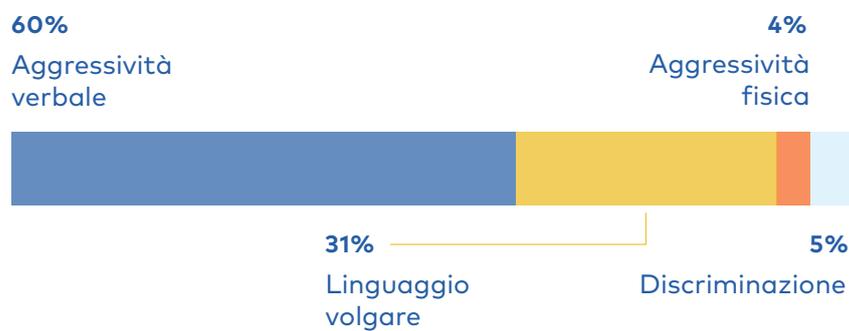
FIGURA 4

La rilevanza delle quattro dimensioni di *hate speech* su Facebook



FIGURA 5

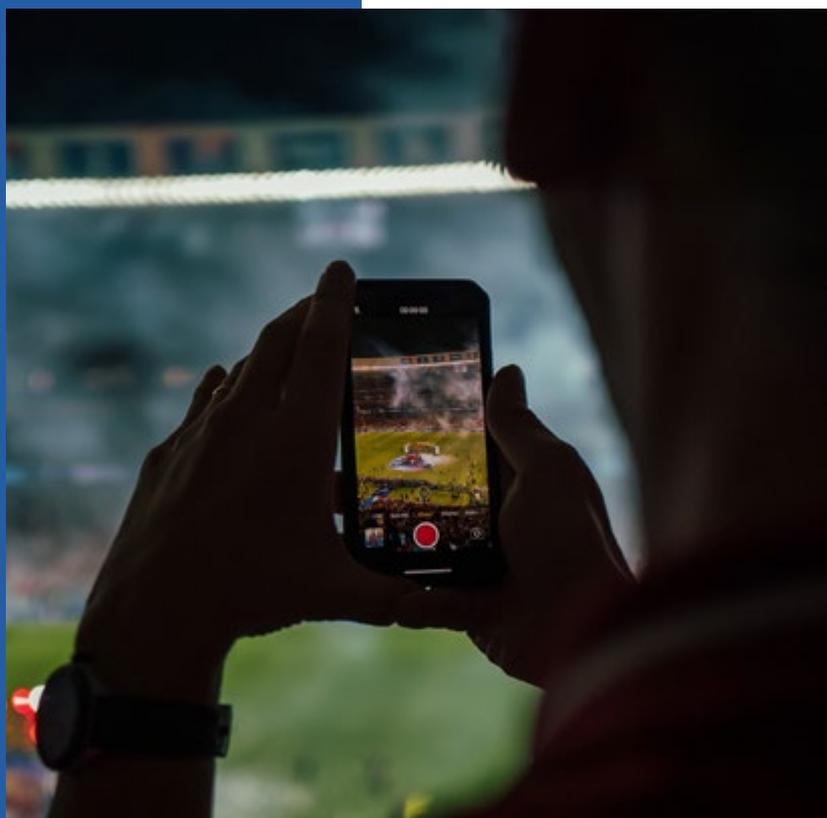
La rilevanza delle quattro dimensioni di *hate speech* su Twitter





CAPITOLO 5

Gli sport non sono tutti uguali sui social media e in relazione all'*hate speech*



PARAGRAFO 5.1

Gli sport più presenti su Facebook e Twitter

Per comprendere il fenomeno dell'*hate speech* all'interno dei due principali social media – Twitter e Facebook – è necessario innanzitutto identificare quali sono gli sport a cui le testate sportive hanno dedicato maggiore attenzione nei loro contenuti social. Le FIGURE 6 E 8 mostrano con grande evidenza che **il calcio, con 4161 post su Facebook e 4061 tweet, è lo sport più presente nei contenuti pubblicati sui social media** dai quattro quotidiani analizzati. Le proporzioni non lasciano dubbi. Infatti, i contenuti social dedicati ad altri sport rappresentano un numero decisamente residuale di post e tweet. Fra questi, come illustrato dalle FIGURE 7 E 9, sono in particolare il tennis, il basket e la Formula 1 a registrare il numero di contenuti più alto.

Il calcio

è lo sport **più trattato sulle pagine social** delle testate sportive e **più commentato** dagli utenti.

Su Facebook, infatti, risultano 55 post riferiti al tennis, 47 al basket e 30 alla Formula 1.

Su Twitter osserviamo dati analoghi, ma se il tennis si conferma il secondo sport più presente con 48 tweet, in questo caso segue la Formula 1 con 40 tweet e infine il basket con soli 12 contenuti.

FIGURA 6

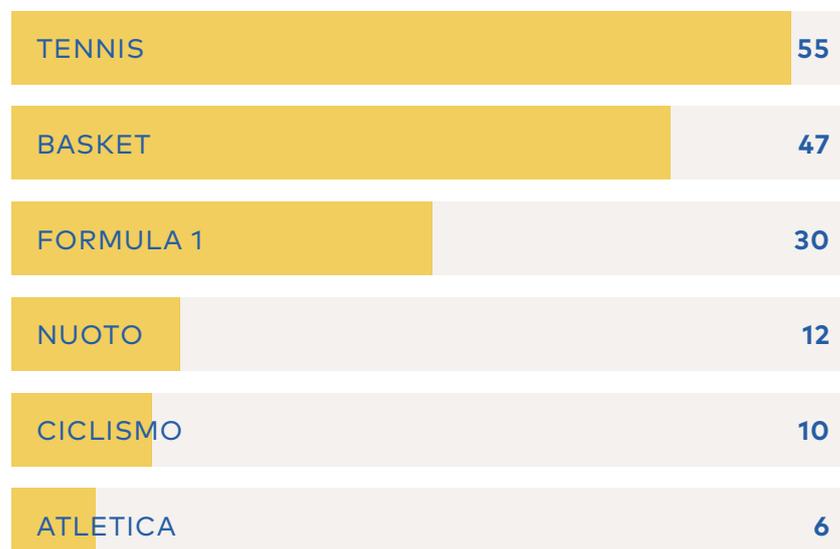
Sport principale trattato nei post analizzati su Facebook



4500 post analizzati in totale

FIGURA 7

Altri sport trattati nei post analizzati su Facebook, escluso il calcio

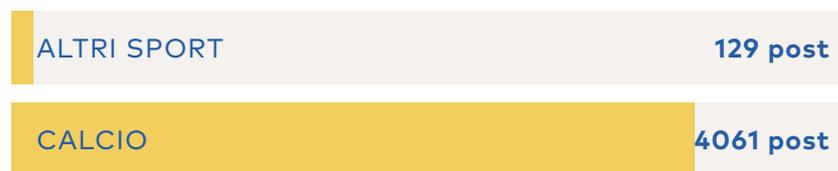


I nostri dati sul volume di contenuti dedicati ai diversi sport ci raccontano dunque di una centralità del calcio nelle pagine social delle testate sportive analizzate. L'attenzione infatti è concentrata – senza troppe sorprese – su quello che è considerato lo sport nazionale dell'Italia.

L'attenzione sugli altri sport è invece più marginale e focalizzata principalmente su altri tre sport: tennis, basket e Formula 1.

FIGURA 8

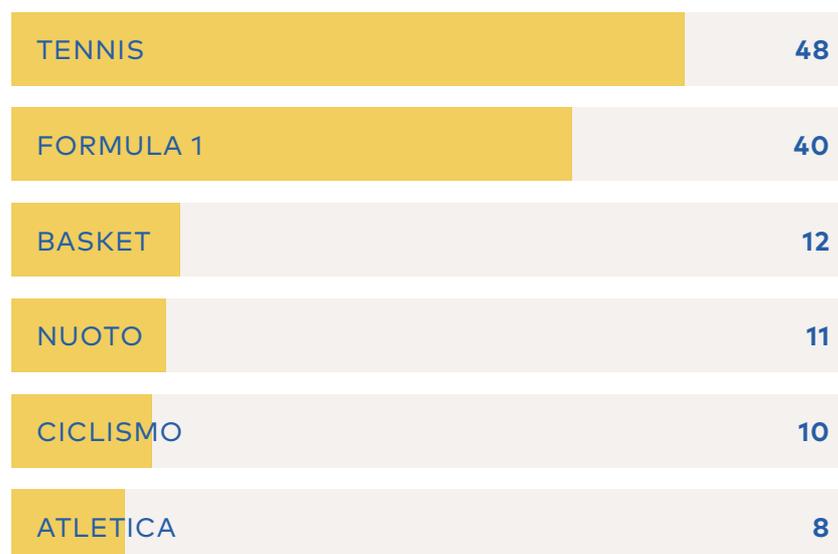
Sport principale trattato nei tweet analizzati su Twitter



5000 tweet analizzati in totale

FIGURA 9

Altri sport trattati nei tweet analizzati su Twitter, escluso il calcio





PARAGRAFO 5.2

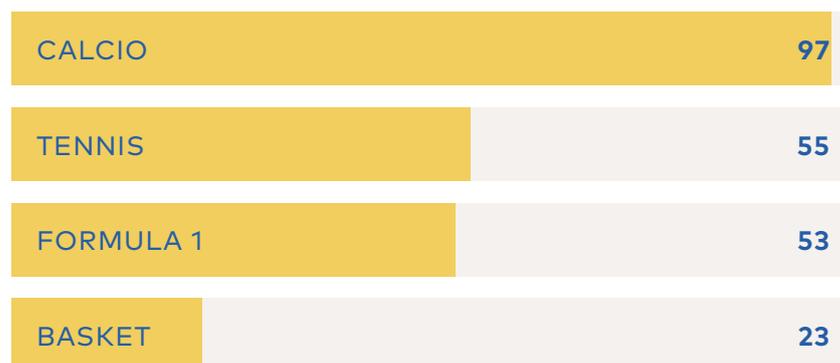
I post e i tweet con più commenti su Facebook e Twitter

Se spostiamo la nostra attenzione sul numero di commenti raccolti dai post o tweet dedicati ai diversi sport osserviamo alcune differenze fra le due piattaforme social.

Per quel che riguarda Facebook, osservando la FIGURA 10 notiamo che i post focalizzati sul calcio raccolgono in media un numero largamente superiore di commenti. Se i contenuti Facebook sono infatti dedicati a **temi calcistici i commenti in media sono 97 contro i 55 e i 53 commenti registrati in media per i post riferiti, rispettivamente, a tennis e Formula 1. Molto inferiore è il numero medio di commenti relativo ai post dedicati al basket (23).**

FIGURA 10

Numero medio di commenti agli sport trattati nei post



I dati riportati dalla FIGURA 11, invece, si riferiscono a Twitter.

Qui il primo elemento che può essere sottolineato è la differenza in termini di volume di commenti. Ancora una volta, il calcio si conferma come lo sport più commentato (4 commenti), seguito da Formula 1 e basket, entrambi con due commenti in media, e infine il tennis con un solo commento. Queste cifre sono decisamente più contenute rispetto a quelle registrate per i commenti ai contenuti Facebook, certificando la differenza fra le due piattaforme e facendo dunque registrare una differente propensione a commentare i contenuti offerti dalle testate sportive analizzate. In ogni caso, il calcio si conferma come lo sport non solo più presente ma anche come quello che stimola maggiori commenti.

FIGURA 11

Numero medio di commenti agli sport trattati nei tweet

CALCIO	4
TENNIS	2
FORMULA 1	2
BASKET	1



PARAGRAFO 5.3

La presenza di *hate speech* nei commenti su Facebook e Twitter

Andiamo ora a focalizzarci sul tema al cuore di questa ricerca, vale a dire la presenza di elementi di *hate speech* nei commenti ai post delle testate che abbiamo analizzato.

Le FIGURE 12 E 13, in particolare, riferiscono la percentuale di commenti che contengono *hate speech*, considerando lo sport trattato all'interno dei post Facebook o dei tweet analizzati.

Per quel che riguarda Facebook, notiamo che il 12,3% dei commenti a post riferiti a tematiche e questioni calcistiche contiene riferimenti a linguaggio d'odio.

COMMENTI SU FACEBOOK CONTENENTI *HATE SPEECH*

Calcio	Formula 1
12,3%	10%
Basket	Tennis
8,7%	7,6%

In seconda istanza, registriamo che quasi un commento su 10 riferito a post dedicati alla Formula 1 presenta elementi di *hate speech*. Infine, l'8,7% dei commenti a post riferiti al basket e il 7,6% di quelli riferiti al tennis includono espressioni o parole riferibili a linguaggio d'odio.

Se spostiamo la nostra attenzione su Twitter, il quadro presenta alcune interessanti differenze. I dati mostrano infatti che sono soprattutto il basket (31,8%) e il calcio (31,3%) a registrare commenti che presentano elementi di *hate speech*.

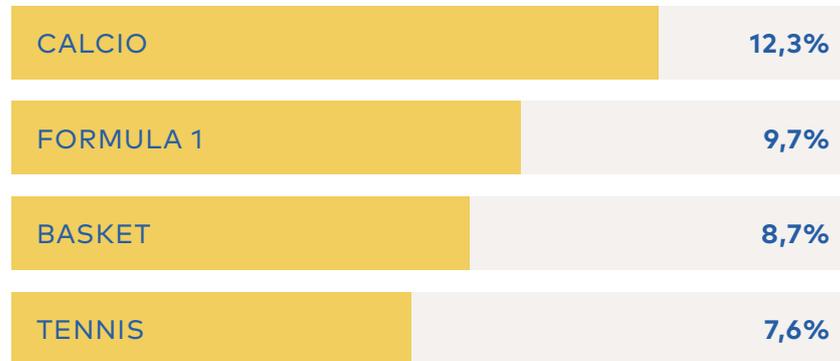
Vale a dire che su Twitter circa un commento su 3 a notizie che si occupano di basket o calcio si connota per espressioni di volgarità, aggressione verbale o fisica o discriminazione. La presenza di *hate speech* è più contenuta, ma pur sempre rilevante fra i commenti a contenuti sulla Formula 1 (21,4%), ossia un commento su 5.

COMMENTI SU TWITTER CONTENENTI *HATE SPEECH*

Basket	Calcio
31,8%	31,3%
Formula 1	Tennis
21,4%	15,2%

FIGURA 12

Percentuale di commenti contenenti *hate speech* nei post che parlano di sport

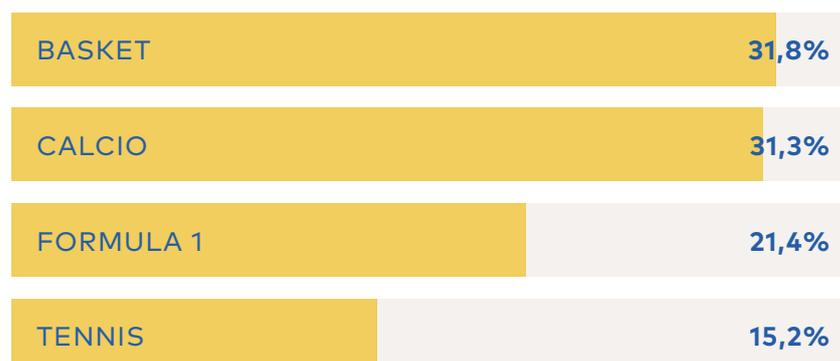


Infine, i commenti ai tweet che si interessano di tennis che includono elementi di *hate speech* sono circa il 15,2%.

In generale, rispetto a quanto osservato per i commenti ai post Facebook, su Twitter i commenti ai contenuti pubblicati dalle testate giornalistiche sportive analizzate sembrano essere caratterizzati da una quota decisamente più importante di *hate speech*. Tuttavia, occorre ricordare che - come segnalato sopra - su Twitter i commenti sono in generale più limitati. Infine, su entrambe le piattaforme social sono le notizie relative al calcio a registrare la porzione più alta di *hate speech*.

FIGURA 13

Percentuale di commenti contenenti *hate speech* nei tweet che parlano di sport





PARAGRAFO 5.4

Quale tipo di *hate speech* nei commenti

Entrando più nel dettaglio della nostra analisi, le FIGURE 14 E 15 chiariscono quali sono le modalità di *hate speech* più frequenti nei commenti ai contenuti riferiti ai quattro sport più visibili sulle piattaforme di social media.

Per quel che riguarda Facebook osserviamo che l'aggressività verbale è la modalità di *hate speech* più ricorrente all'interno dei commenti. Più precisamente, il dato più rilevante è registrato fra i commenti a post Facebook che trattano di tematiche calcistiche (12,3%), mentre Formula 1 (9,8%) e basket (8,4%) presentano quote più simili di aggressività verbale.

È il tennis, invece, a registrare la quota più bassa di commenti caratterizzati dall'utilizzo di aggressività verbale (6,4%).

Le altre forme di *hate speech* sembrano invece meno ricorrenti. Si osserva in particolare la maggiore presenza di linguaggio volgare nei commenti a post Facebook inerenti al calcio (2,8%), mentre registriamo che nei commenti ai post dedicati al basket sia più presente la dimensione della discriminazione.

I dati riferiti a Twitter riportati nella FIGURA 13 mostrano alcune analogie, anche se come prevedibile i volumi sono differenti. In primis, anche su Twitter prevale nettamente l'*hate speech* espresso in forma di aggressività verbale, il dato emerge con particolare evidenza fra i commenti dedicati ai post focalizzati su calcio (23,7%) e basket (22,7%), in entrambi i casi elementi di aggressività verbale possono essere rintracciati in più di un commento su cinque. L'aggressività verbale resta la modalità di *hate speech* prevalente anche per i commenti riferiti a contenuti che riguardano la Formula 1 (16,7%) e il tennis (13,6%). In seconda battuta, fra i commenti su Twitter si segnala una presenza importante di linguaggio volgare. In questo caso, la quota è osservata fra i commenti dedicati al basket (13,6%), ma è opportuno sottolineare che anche per quel che riguarda i commenti a contenuti riferiti al calcio si registra un dato simile (12,5%), mentre per la Formula 1 il linguaggio volgare si attesta al 9,5%.

Non si rintracciano invece elementi di linguaggio volgare per quel che concerne il tennis. Più articolata è invece la distribuzione di commenti che presentano discriminazione. È soprattutto il basket che registra la porzione più larga di commenti contenenti espressioni di natura discriminatoria, ben il 9,1% dei commenti. La discriminazione invece pare decisamente meno presente e più marginale fra i commenti ai tweet che si interessano degli altri sport.

FIGURA 14

Percentuale di commenti contenenti i diversi tipi di *hate speech* nei post che parlano degli sport

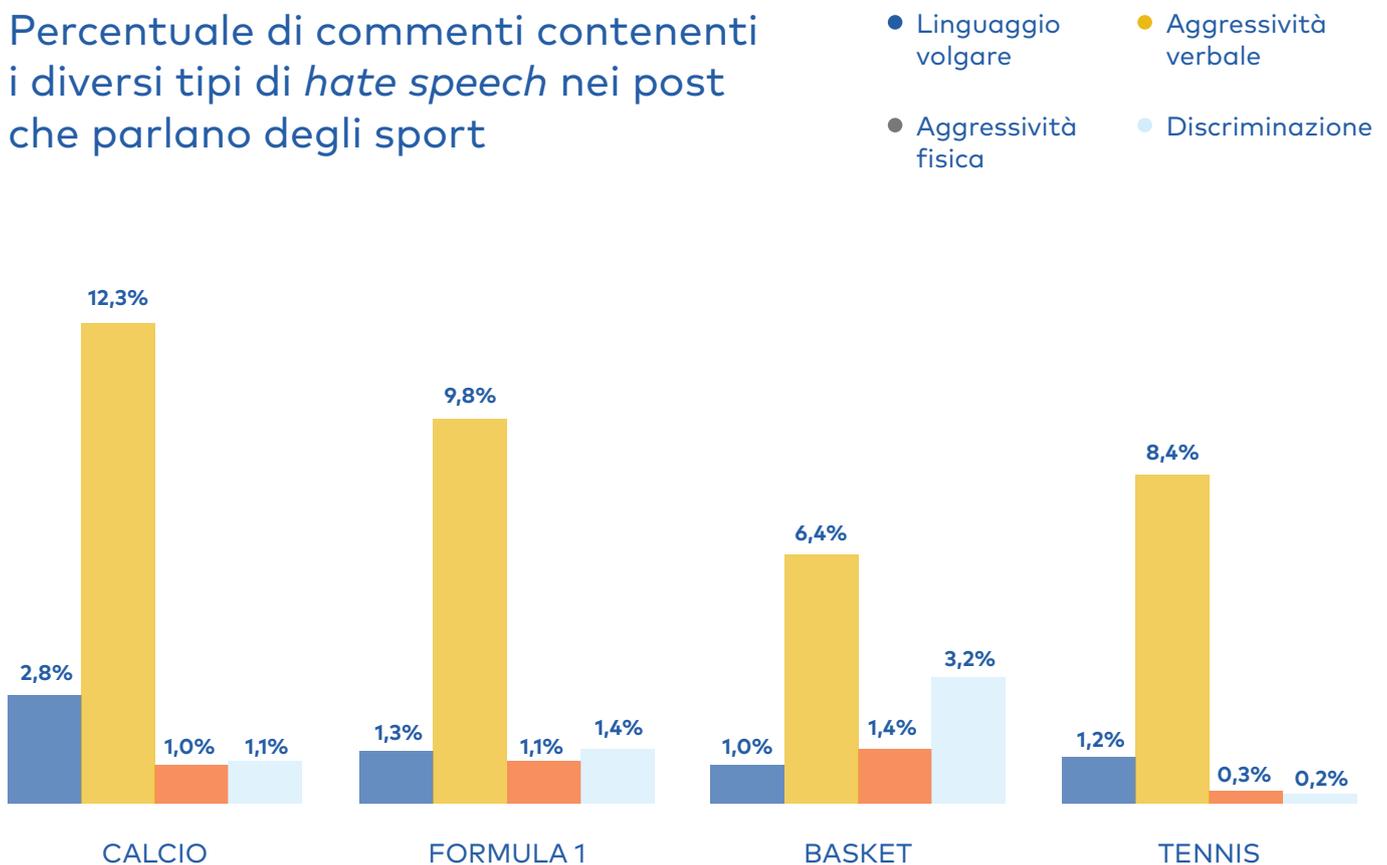
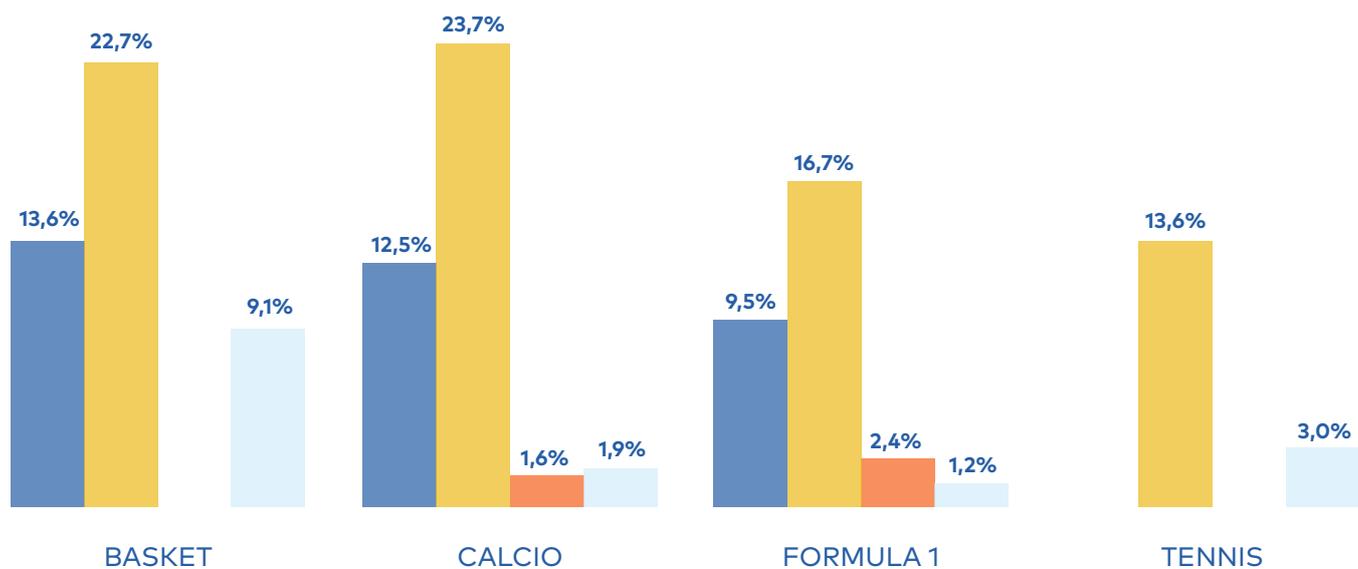


FIGURA 15

Percentuale di commenti contenenti i diversi tipi di *hate speech* nei tweet che parlano degli sport

- Linguaggio volgare
- Aggressività verbale
- Aggressività fisica
- Discriminazione





CAPITOLO 6

Focus sul calcio: le squadre e *l'hate speech*



Come si è visto, il calcio è lo sport più trattato dalle testate giornalistiche e anche quello più commentato. Con grande facilità, quindi, navigando su un social media si possono incrociare notizie e commenti su campionato, coppe o nazionale. E con altrettanta facilità all'interno di questi commenti sarà probabile leggere *hate speech*. Per approfondire meglio le ragioni di questo quadro, in questo capitolo ci concentriamo sulle principali squadre di calcio, mentre nel prossimo sui principali calciatori.



PARAGRAFO 6.1

Le squadre più presenti su Facebook e Twitter

Non sorprendentemente, visti i risultati sportivi e la disponibilità economica della proprietà, **il club che ottiene maggiore copertura è la Juventus**: sia su Facebook, sia su Twitter le notizie sulla Juventus sono quasi doppie rispetto a quelle ottenute dall'Inter, secondo club per visibilità (FIGURA 16 E 17). Con qualche minima differenza tra i due mezzi, seguono poi quelle squadre che per blasone, storia e risultati recenti risultano tra le più conosciute e tifate: Napoli, Roma, Milan e Lazio. Con netto distacco seguono le restanti squadre che ottengono una copertura sporadica, contando in media meno di 2 articoli o servizi al giorno su tutte le testate analizzate. Vista questa grande disparità nei volumi di post prodotti, soltanto le prime sei squadre vengono approfondite nei paragrafi seguenti.

FIGURA 16

Le squadre più presenti nei post analizzati su Facebook

JUVENTUS	1224
INTER	595
NAPOLI	439
MILAN	422
ROMA	397
LAZIO	236
ATALANTA	146
BRESCIA	126
TORINO	97
BOLOGNA	90
VERONA	67
FIorentina	28

FIGURA 17

Le squadre più presenti nei post analizzati su Twitter

JUVENTUS	1362
INTER	783
MILAN	545
NAPOLI	454
ROMA	289
LAZIO	223
ATALANTA	179
BRESCIA	144
BOLOGNA	69
VERONA	58
TORINO	28
FIorentina	8



PARAGRAFO 6.2

Le squadre con più commenti

Ma la disparità che si rileva nella copertura giornalistica delle diverse squadre trova una corrispondenza anche nel volume di commenti che i post sulle diverse squadre generano?

Le FIGURE 18 E 19 mostrano chiaramente che le conversazioni tra utenti sono molto meno differenziate tra loro e non seguono le stesse logiche adottate dai giornalisti. Su Facebook, Milan, Juventus e Inter hanno un numero medio di commenti per post molto simile (compreso tra 128 e 125), mentre Lazio (95), Napoli (90) e ancor di più la Roma (48) seguono distanziati. Un quadro diverso emerge invece da Twitter dove, come abbiamo visto, il numero di commenti è molto minore e tutti i club ottengono un numero medio di commenti simile (compreso tra 3 e 4).

SITUAZIONE SQUADRE SPORTIVE

+ visibilità mediatica

↳ Juventus

+ commenti

↳ Milan

+ hate speech

↳ Inter su Facebook

↳ Napoli su Twitter

FIGURA 18

Numero medio di commenti alle squadre presenti nei post

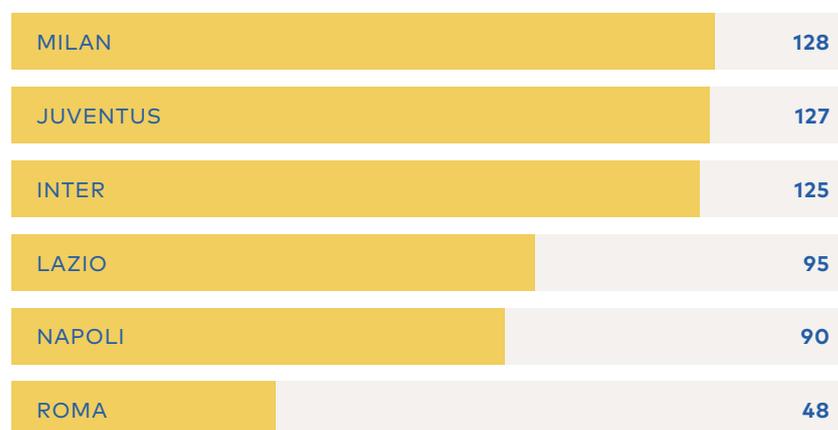
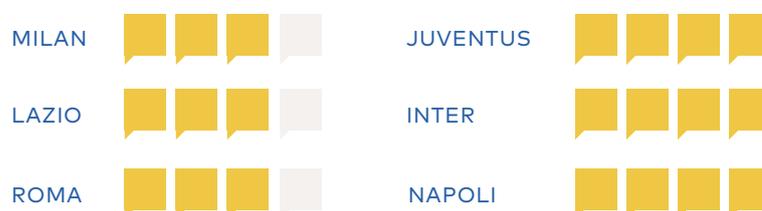


FIGURA 19

Numero medio di commenti alle squadre presenti nei tweet





PARAGRAFO 6.3

La presenza di *hate speech*

Una volta appurato che le sei squadre più visibili sono comparabili in termini di volume di commenti, è finalmente possibile entrare nel merito della ricerca focalizzandosi sulla presenza di *hate speech* nelle conversazioni sulle diverse squadre (FIGURE 20 E 21). Su Facebook, la squadra che ottiene la percentuale più alta di commenti con *hate speech* è l'Inter (14,1%). Seguono da vicino Lazio (13,5%), Juventus (13,2%), Napoli (12,8%) e Roma (12,3%). A livelli un po' più bassi troviamo invece il Milan con una percentuale di poco superiore al 10%.

FIGURA 20

Percentuale di commenti contenenti *hate speech* nei post che parlano delle squadre

INTER	14,1
LAZIO	13,5
JUVENTUS	13,2
NAPOLI	12,8
ROMA	12,3
MILAN	10,6

Su Twitter, come abbiamo visto nel Capitolo 4, i livelli di *hate speech* sono sistematicamente più alti di quelli registrati su Facebook. In questo caso, però, non si modifica solo la porzione di commenti d'odio, ma anche l'ordine delle squadre con più *hate speech*. È il Napoli, in questo caso, il club ad avere maggiori riferimenti d'odio nei commenti che lo riguardano (34%), seguito dall'Inter (32,8%), dalla Juventus (30,9%) e dalla Roma (30,6%). A livelli leggermente più bassi troviamo invece Milan (28,9%) e Lazio (27,1%).

FIGURA 21

Percentuale di commenti contenenti *hate speech* nei tweet che parlano delle squadre





PARAGRAFO 6.4

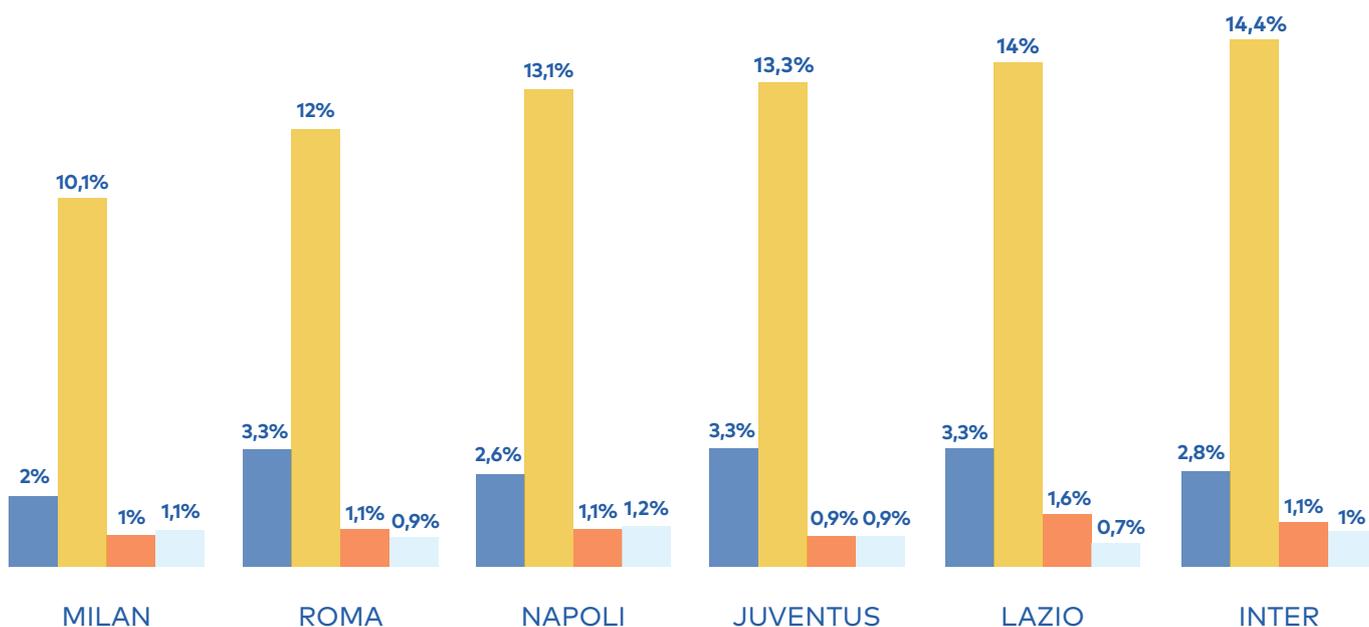
Il tipo di *hate speech* nei commenti

Se guardiamo più nel dettaglio alla composizione dell'*hate speech* notiamo che, per tutte le squadre e per entrambi i social media utilizzati, è l'aggressività verbale la componente trainante. Su Facebook varia tra un minimo di 10,1% (Milan) a un massimo di 14,4% (Inter), mentre su Twitter il livello è quasi doppio, compreso tra 20,4% (Milan) e 26,5% (Napoli). **Il linguaggio volgare si conferma più rilevante per Twitter** (compreso tra 10,6% e 13,3%) rispetto a Facebook (tra 2,0% e 3,3%), ma non si evidenziano significative differenze tra i diversi club.

FIGURA 22

Percentuale di commenti contenenti i diversi tipi di *hate speech* riferiti ai post che parlano delle squadre

- Linguaggio volgare
- Aggressività verbale
- Aggressività fisica
- Discriminazione fisica



Aggressività verbale e linguaggio volgare sono le componenti trainanti dei commenti di *hate speech*, ma questo non significa che le altre siano irrilevanti, sono stati individuati circa

5000 commenti

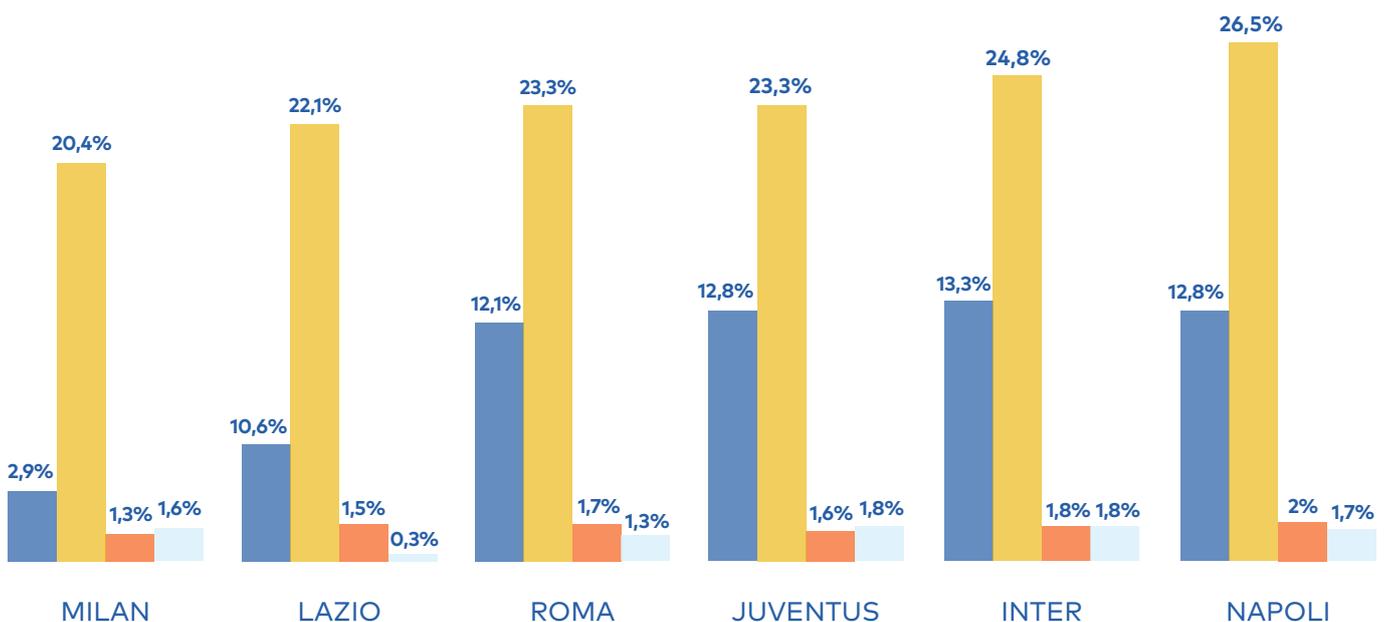
contenenti **discriminazione e aggressività fisica.**

La componente hard dell'*hate speech* – vale a dire l'uso di offese e minacce fisiche così come le espressioni di discriminazione – resta limitata e uniformemente distribuita tra le squadre. Questo non vuol dire tuttavia che sia una componente irrilevante dal momento che si tratta di circa 5000 commenti contenenti discriminazione e altrettanti contenenti aggressività fisica.

FIGURA 23

Percentuale di commenti contenenti i diversi tipi di *hate speech* riferiti ai tweet che parlano delle squadre

- Linguaggio volgare
- Aggressività verbale
- Aggressività fisica
- Discriminazione fisica





CAPITOLO 7

Focus sul calcio: i personaggi sportivi e l'*hate speech*



PARAGRAFO 7.1

I personaggi sportivi più visibili sui social

La nostra analisi si sposta ora sui personaggi sportivi.

Tenendo come focus il calcio che abbiamo visto essere lo sport più presente sui contenuti social delle testate sportive monitorate in questa ricerca, le TABELLE 24 E 25 ci mostrano chi sono i calciatori più visibili, rispettivamente, su Facebook e Twitter. Seppure con qualche differenza, i dati sono in verità piuttosto simili. Su tutti, **Cristiano Ronaldo è il più presente nei post Facebook (354) e nei tweet (442)**. La sua fama internazionale è chiaramente un elemento che guida le testate sportive a coprire assiduamente nei contenuti social il giocatore della Juventus.

FIGURA 24

I personaggi sportivi più presenti nei post analizzati su Facebook

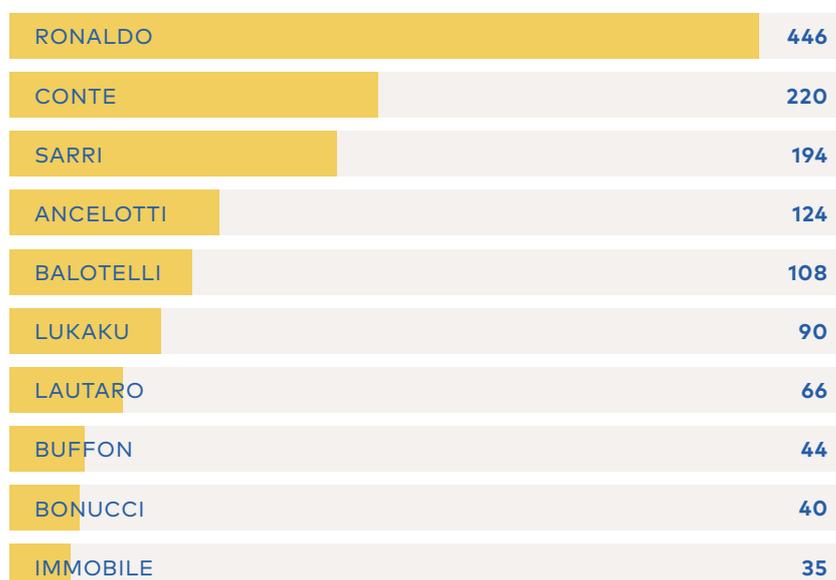
RONALDO	354
SARRI	218
CONTE	203
ANCELOTTI	136
BALOTELLI	77
LUKAKU	56
LAUTARO	53
BONUCCI	40
BUFFON	39
IMMOBILE	38

Sul podio dei personaggi sportivi più visibili troviamo poi **due allenatori: Antonio Conte e Maurizio Sarri**. Su Facebook, l'allenatore della Juventus conta 218 post, seguito da Conte (203 post).

Su Twitter, troviamo invece un'inversione e una distanza maggiore: il tecnico interista è citato in 220 tweet, mentre sono 194 i tweet riferiti a Sarri. Un altro allenatore, questa volta del Napoli, **Carlo Ancelotti** è invece presente in 136 post Facebook e 124 tweet. I tre allenatori sono seguiti in termini di visibilità da due attaccanti: **Mario Balotelli e Romelu Lukaku**. Il calciatore del Brescia è citato in 77 post Facebook e 108 tweet. Il centravanti interista, invece, è presente su Facebook in 56 post e su Twitter in 90 contenuti pubblicati dalle testate sportive analizzate. Gli altri giocatori inclusi nell'analisi sembrano invece meno visibili, con l'altro attaccante interista, Lautaro con 53 post e 66 tweet. Bonucci registra la stessa quota di post Facebook e tweet (40), e un volume di post Facebook e tweet simile per Buffon e Immobile.

FIGURA 25

I personaggi sportivi più presenti nei tweet analizzati





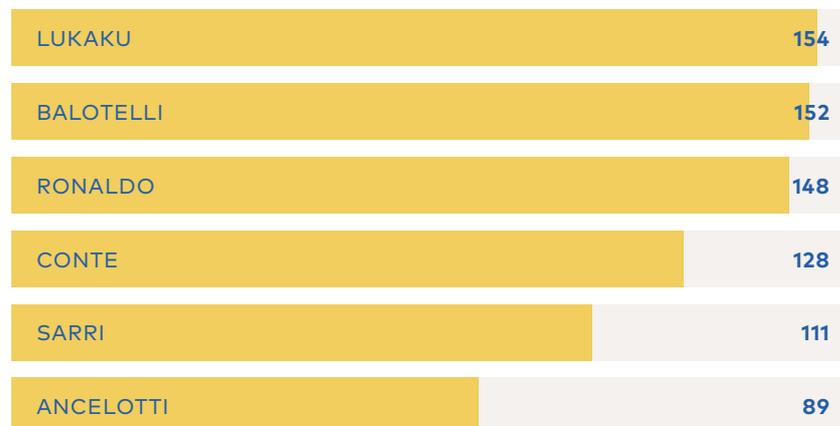
PARAGRAFO 7.2

Il tipo di *hate speech* nei commenti

La nostra analisi si focalizza ora sui sei personaggi sportivi più visibili: Cristiano Ronaldo, Antonio Conte, Maurizio Sarri, Carlo Ancelotti, Romelu Lukaku e Mario Balotelli. Ci interessano in particolare i commenti ai post Facebook e ai tweet che includono riferimenti a questi sei sportivi. La FIGURA 26 riporta il numero di commenti ai post dedicati ai 6 personaggi sportivi più visibili su Facebook. Osserviamo come **a maggiore visibilità, in realtà, non corrisponda necessariamente un analogo volume di commenti da parte degli utenti**. Su tutti, sono infatti Lukaku e Balotelli, con un numero medio di commenti rispettivamente di 154 e 152, a registrare il volume più alto di utenti attivi a intervenire a commento dei post che li riguardano. L'attaccante juventino, Cristiano Ronaldo, invece conta 148 commenti medi ai contenuti che parlano di lui.

FIGURA 26

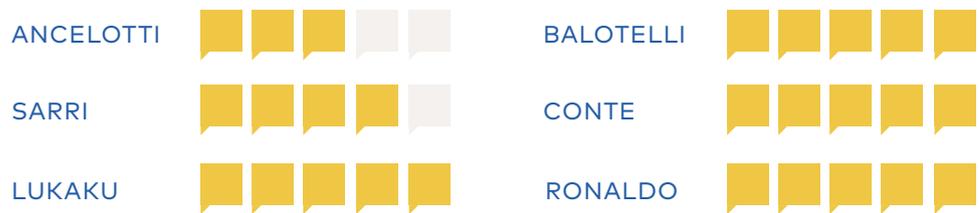
Numero medio di commenti ai personaggi sportivi presenti nei post su Facebook



Il trio di allenatori che invece spiccava per visibilità sembra registrare livelli più contenuti di commenti. In media i commenti ai post che parlano di Antonio Conte sono 128, mentre per quel che riguarda il tecnico juventino Sarri il numero di commenti medio è pari a 111. Decisamente più contenuto il volume di commenti rivolti ai post Facebook che si occupano di Carlo Ancelotti (89).

FIGURA 27

Numero medio di commenti alle squadre presenti nei tweet



Se osserviamo invece i dati presentati nella FIGURA 27, troviamo ulteriore conferma del numero decisamente contenuto dei commenti che accompagnano i contenuti pubblicati su Twitter dalle testate sportive. Più nel dettaglio i tweet che riguardano Cristiano Ronaldo, Mario Balotelli, Romelu Lukaku e Antonio Conte ottengono in media 5 commenti. Mentre nel caso di contenuti riferiti a Maurizio Sarri a Carlo Ancelotti i commenti sono in media, rispettivamente, 4 e 3.



PARAGRAFO 7.3

La presenza di *hate speech*

Le FIGURE 28 E 29 illustrano la presenza di contenuti di *hate speech* all'interno dei commenti ai post che parlano dei personaggi sportivi considerati nel nostro focus.

Per quanto riguarda Facebook, in prima battuta, notiamo che la percentuale più alta di commenti che fanno ricorso a linguaggio d'odio si registra fra i post dedicati a Mario Balotelli (16,7%) e – a pochissima distanza – Antonio Conte (16,1%). Viceversa, i post che riguardano Cristiano Ronaldo registrano una quota inferiore di *hate speech* (12,9%). Il dato è interessante, perché rispetto a quanto presentato sopra, riguardo alla visibilità dei personaggi sportivi, osserviamo un ribaltamento di prospettiva.

FIGURA 28

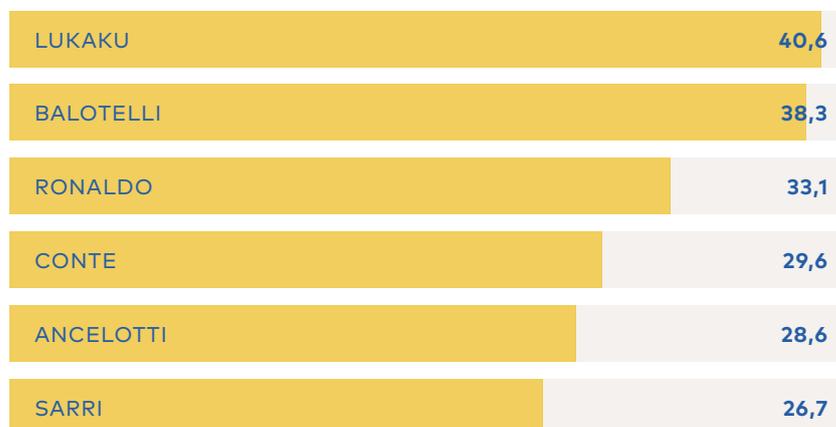
Percentuale di commenti contenenti *hate speech* nei post che parlano di personaggi sportivi su Facebook



Un quadro analogo è rintracciabile anche a partire dai dati relativi a Twitter presentati nella FIGURA 29. Anche in questo caso infatti si nota come i giocatori meno visibili in realtà siano quelli che invece sembrano raccogliere una quota più rilevante di *hate speech*.

FIGURA 29

Percentuale di commenti contenenti hate speech nei tweet che parlano di personaggi sportivi



I contenuti che si riferiscono a Lukaku, infatti, vedono una porzione di commenti caratterizzati da linguaggio d'odio pari al 40,6%; una configurazione simile a quella relativa ai post che riguardano invece Balotelli, dove i commenti d'odio sono il 38,3%.

Mario Balotelli

è il più colpito da commenti di *hate speech*.

I giocatori e gli allenatori con una minore presenza all'interno dei post delle testate giornalistiche sono quelli che raccolgono una porzione più alta di messaggi di *hate speech*.

In altre parole, circa un commento su quattro presenta almeno un richiamo ad elementi di *hate speech*. Se consideriamo i tweet riferiti a Cristiano Ronaldo, un terzo dei commenti presenta *hate speech*. Livelli simili si riscontrano per i commenti ai tweet relativi a Conte e Ancelotti, rispettivamente il 29,6% e il 28,3% contengono parole d'odio. E, infine, i commenti di *hate speech* ai tweet che trattano il tecnico juventino Sarri sono il 26,7%.

In sintesi, il quadro che emerge su Twitter è particolarmente interessante, perché a fronte di un numero ridotto di commenti come abbiamo visto dai dati presentati sopra osserviamo invece una maggiore densità di contenuti d'odio, laddove sui tweet riferibili ai sei personaggi sportivi analizzati registriamo una quota non inferiore a un quarto di linguaggio d'odio.



PARAGRAFO 7.4

Il tipo di *hate speech*

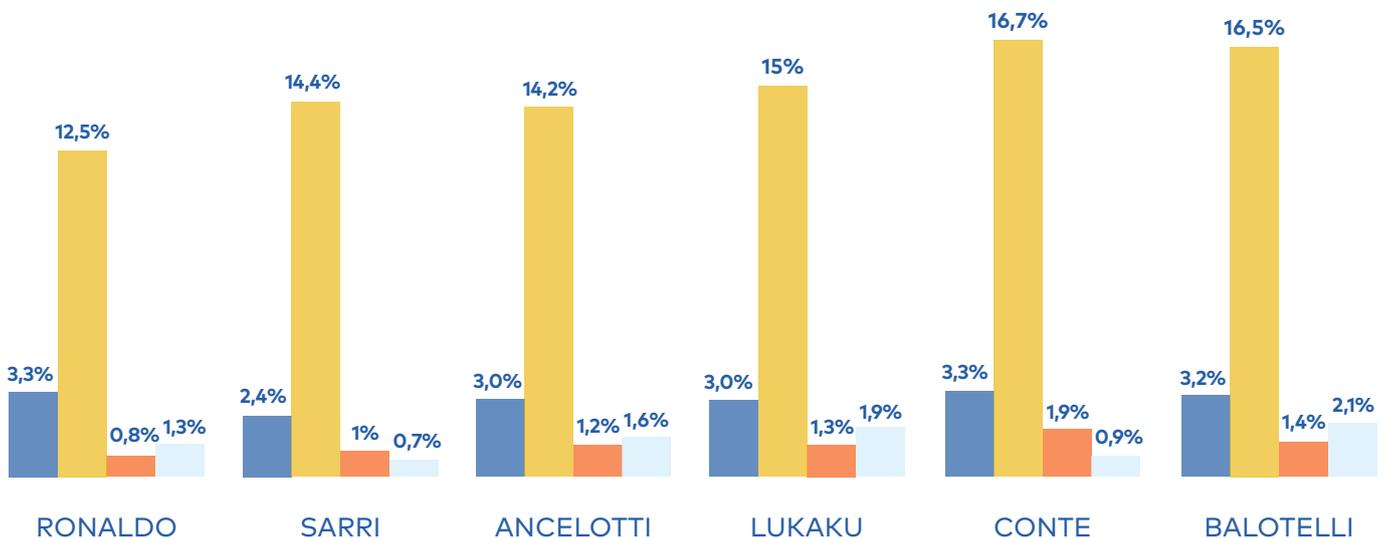
Nelle FIGURE 30 E 31 è possibile vedere la diversa distribuzione di elementi di *hate speech* all'interno dei commenti ai contenuti social sui sei personaggi sportivi considerati .

Se osserviamo i dati relativi a Facebook è facile identificare la prevalenza di elementi legati ad aggressività verbale, con una variazione piuttosto stretta. La percentuale più alta di commenti caratterizzati da **aggressività verbale** si osserva fra i commenti riferiti ai post in cui è presente Antonio Conte (16,7%) – ma occorre sottolineare che il dato è molto simile a quello riferito ai contenuti in cui si parla di Mario Balotelli. Al contrario, la porzione più contenuta di commenti connotati da aggressività verbale si registra per i post riferiti a Cristiano Ronaldo (12,5%).

FIGURA 30

Percentuale di commenti contenenti i diversi tipi di *hate speech* nei post che parlano di personaggi sportivi

- Linguaggio volgare
- Aggressività verbale
- Aggressività fisica
- Discriminazione fisica



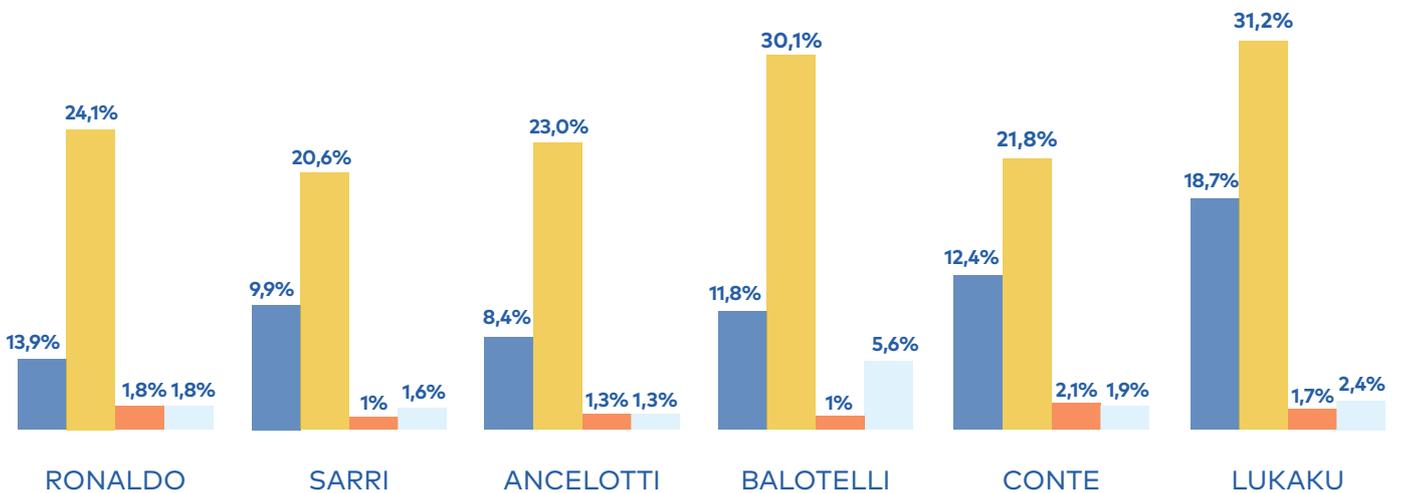
In sostanza, **non meno di un commento su 10 fra quelli dedicati a calciatori e tecnici sportivi inclusi nell'analisi presenta elementi di *aggressività verbale***. Le altre modalità di *hate speech* sembrano meno frequenti, ma presentano comunque alcuni spunti di riflessione interessanti. In prima istanza, infatti, la proporzione di commenti che presentano linguaggio volgare è piuttosto simile per tutti i personaggi sportivi considerati, ma ancora una volta sono i post dedicati a Conte e Balotelli – accompagnati in questo caso da quelli riferiti a Cristiano Ronaldo – a registrare le quote più alte di linguaggio volgare. E lo stesso può essere detto per quanto riguarda i commenti che presentano elementi di aggressività fisica. Anche in questo caso, per quanto si tratti di percentuali molto marginali, sono sempre il tecnico interista e il centravanti del Brescia a registrare livelli più alti di aggressività fisica.

Se veniamo poi al tema della discriminazione emergono i pattern razzisti che connotano l'*hate speech*. Osserviamo infatti che – per quanto limitati – **i commenti che presentano elementi di *discriminazione* sono soprattutto riconducibili a due giocatori in particolare: l'italiano Mario Balotelli e il belga Romelu Lukaku, entrambi di colore.**

FIGURA 31

Percentuale di commenti contenenti i diversi tipi di *hate speech* nei tweet che parlano di personaggi sportivi

- Linguaggio volgare
- Aggressività verbale
- Aggressività fisica
- Discriminazione



Se passiamo ad analizzare l'articolazione dell'*hate speech* sui commenti ai tweet dedicati a calciatori e allenatori osserviamo un quadro piuttosto simile. Infatti, i commenti che presentano elementi di *aggressività verbale* sono anche in questo caso quelli più frequenti per tutti i personaggi sportivi considerati. Sono soprattutto i tweet in cui sono presenti Balotelli e Lukaku a registrare le quote più alte di linguaggio riconducibile ad *aggressività verbale* (rispettivamente 30,1% e 31,2%). Mentre il dato più contenuto – se così si può dire – è osservabile fra i commenti ai tweet in cui si parla di Cristiano Ronaldo, dove comunque almeno un commento su cinque riporta aggressività verbale. Il secondo elemento più ricorrente di *hate speech* per tutti i personaggi sportivi è il *linguaggio volgare*, che caratterizza soprattutto i commenti dedicati ai tweet che si occupano di Lukaku (18,7%). L'*aggressività fisica* – come già osservato anche su Facebook – è un elemento più marginale e varia fra il 2,1% dei commenti ai tweet riferiti ad Antonio Conte e l'1% dei commenti ai contenuti Twitter riferiti a Sarri e Balotelli.

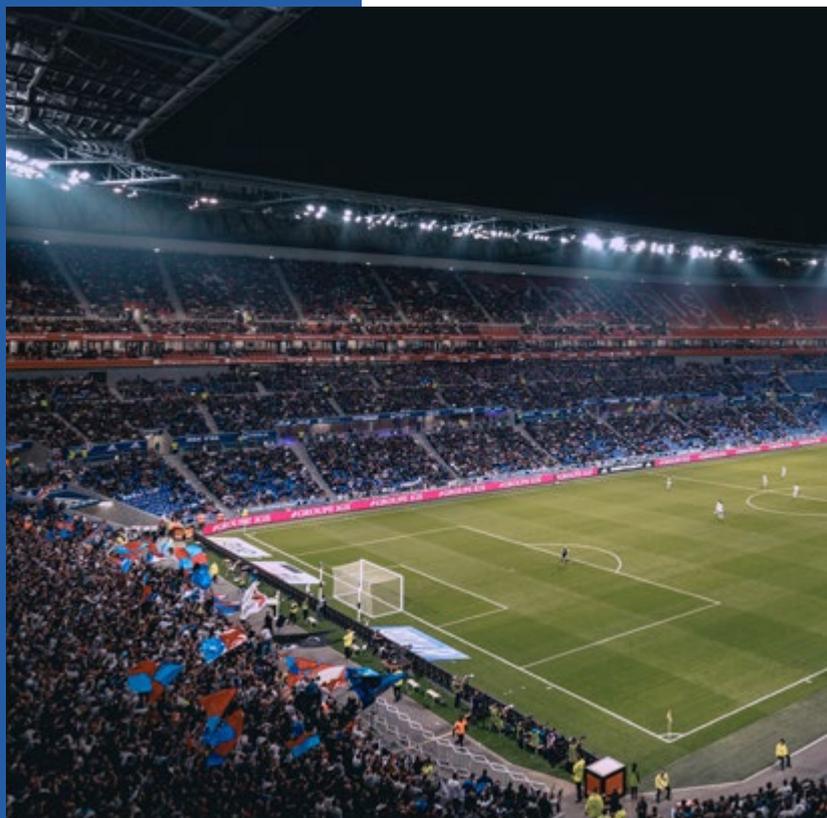


Più interessante è invece considerare il dato riferito agli elementi di *discriminazione*. Come già sottolineato per i commenti ai post Facebook, sono ancora una volta gli unici due giocatori di colore a essere associati a una porzione più larga di commenti che presentano parole discriminatorie. Addirittura, nel caso di Mario Balotelli il dato raggiunge il 5,6% dei commenti, mentre per Lukaku si attesta al 2,4%, suggerendo anche in questo caso che sia proprio il colore della pelle dei due attaccanti ad animare sentimenti razzisti che si traducono esattamente in *hate speech*.



CAPITOLO 8

Distribuzione nel tempo dell'*hate speech*



I dati medi presentati nei paragrafi precedenti non tengono ovviamente conto della distribuzione nel tempo dell'*hate speech* e, in particolar modo, non permettono di capire se sia un fenomeno quotidiano e costante oppure se viva di alcuni picchi, determinati da eventi significativi. Le FIGURE 32 E 33 ci aiutano a chiarire questo punto.

L'andamento delle due figure evidenzia dei picchi. Tra questi, i picchi più bassi si registrano in concomitanza di eventi sportivi – partite di calcio ovviamente, ma non solo – e testimoniano il fatto che in occasione di competizioni sportive il volume di *hate speech* aumenta sistematicamente.

I picchi più alti coincidono invece con partite di calcio a cui fanno seguito polemiche, spesso legate all'arbitraggio e all'uso del VAR. Questo è vero per i due rigori concessi alla Lazio contro l'Atalanta (19 ottobre), un rigore non concesso al Napoli contro l'Atalanta e la successiva espulsione di Carlo Ancelotti (30 ottobre), il rigore non dato per fallo di De Ligt in Torino-Juventus (2 ottobre), la gestione arbitrale di Fiorentina-Inter (15 dicembre).

Unica eccezione a questo tipo di eventi è la polemica tutta giornalistica (ma molto commentata dagli utenti dei social media) del titolo del *Corriere dello Sport*, "Black Friday", riferito a Romelo Lukaku e Chris Smalling due giocatori di colore che si incrociavano nel match di campionato Inter-Roma (5 dicembre).

LIVELLO MINIMO DI HATE SPEECH

18,6%
su Facebook

10,9%
su Twitter

I messaggi d'odio possono quindi a tutti gli effetti considerarsi come una componente strutturale delle conversazioni sportive su questi social media.

La notizia determina una conversazione accesa al punto che elementi d'odio sono riscontrabili in circa un quarto dei commenti del giorno.

Eventi minori, ma significativi in termini di generazione di *hate speech*, sono i cori contro Mario Balotelli in Brescia-Inter del 29 ottobre, in cui la percentuale di commenti contenenti discriminazione aumenta per gli insulti razziali rivolti contro il giocatore del Brescia, e la rapina subita da Claudio Marchisio che provoca invece un'impennata di aggressività verbale e fisica rivolta contro i rapinatori.

FIGURA 32

Distribuzione nel tempo di commenti contenenti hate speech su Facebook

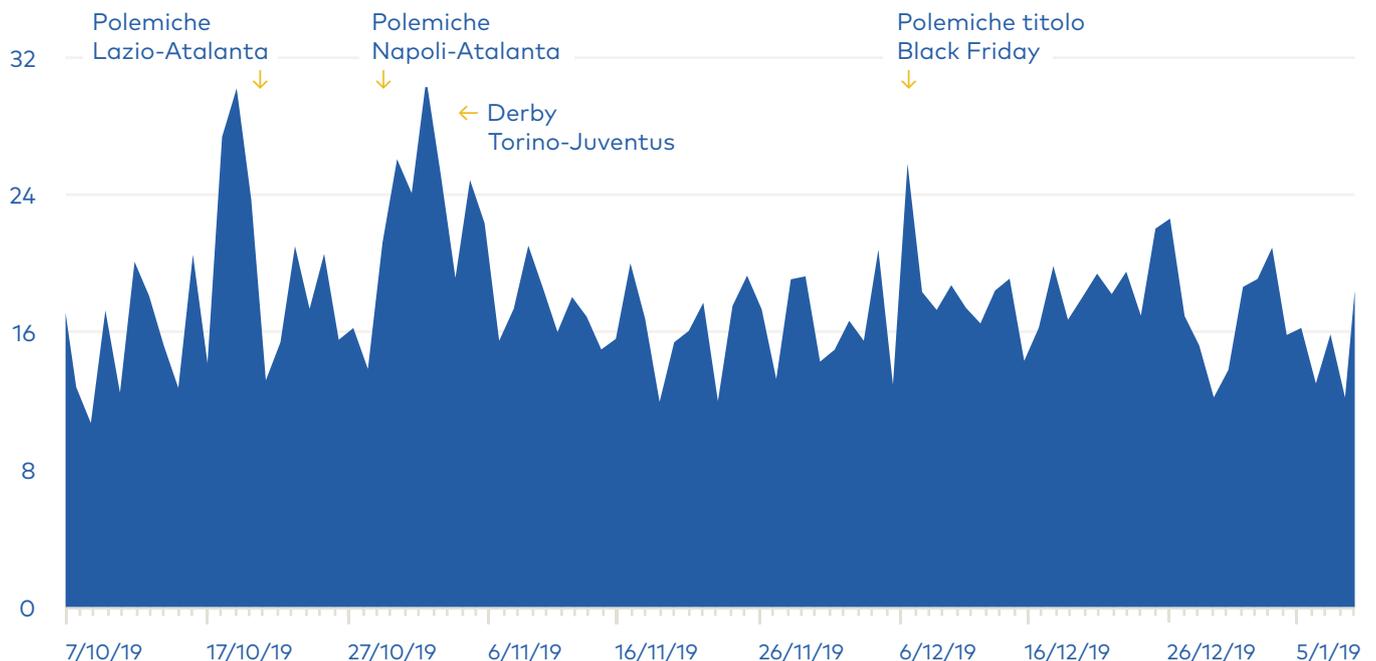
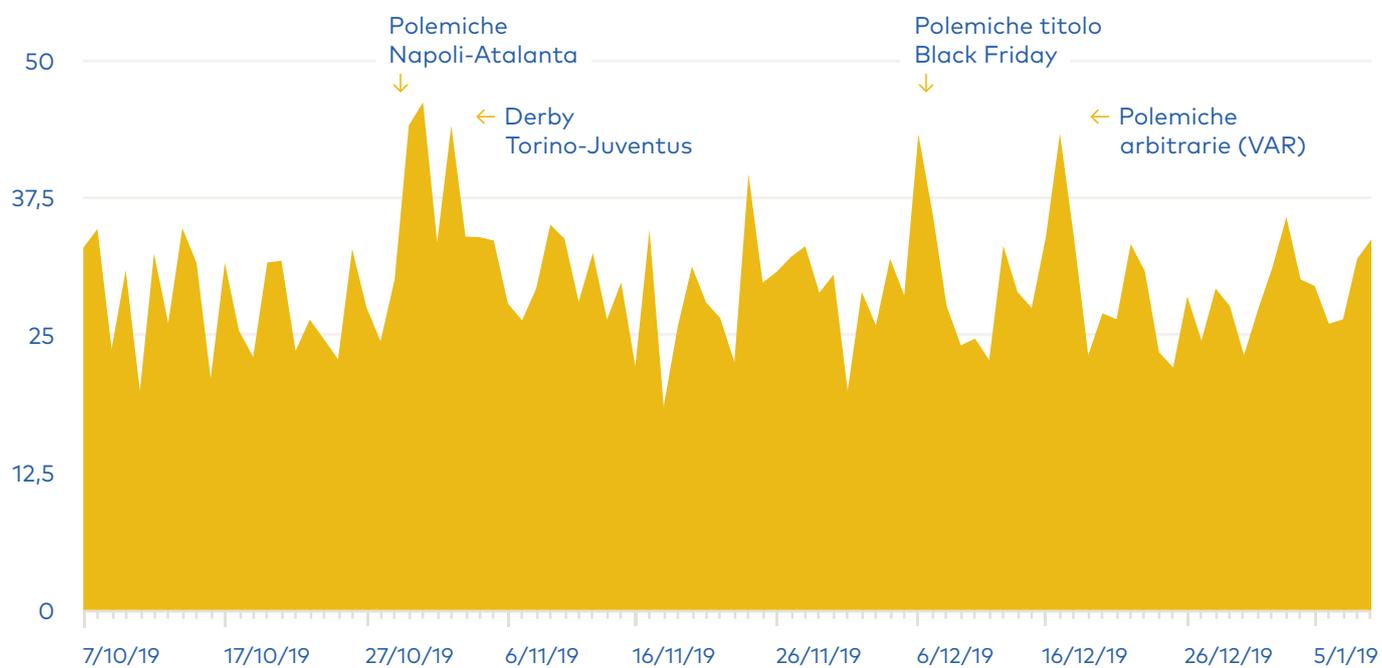


FIGURA 33

Distribuzione nel tempo di commenti contenenti hate speech su Twitter





CAPITOLO 9

Interazioni tra utenti e *hate speech*

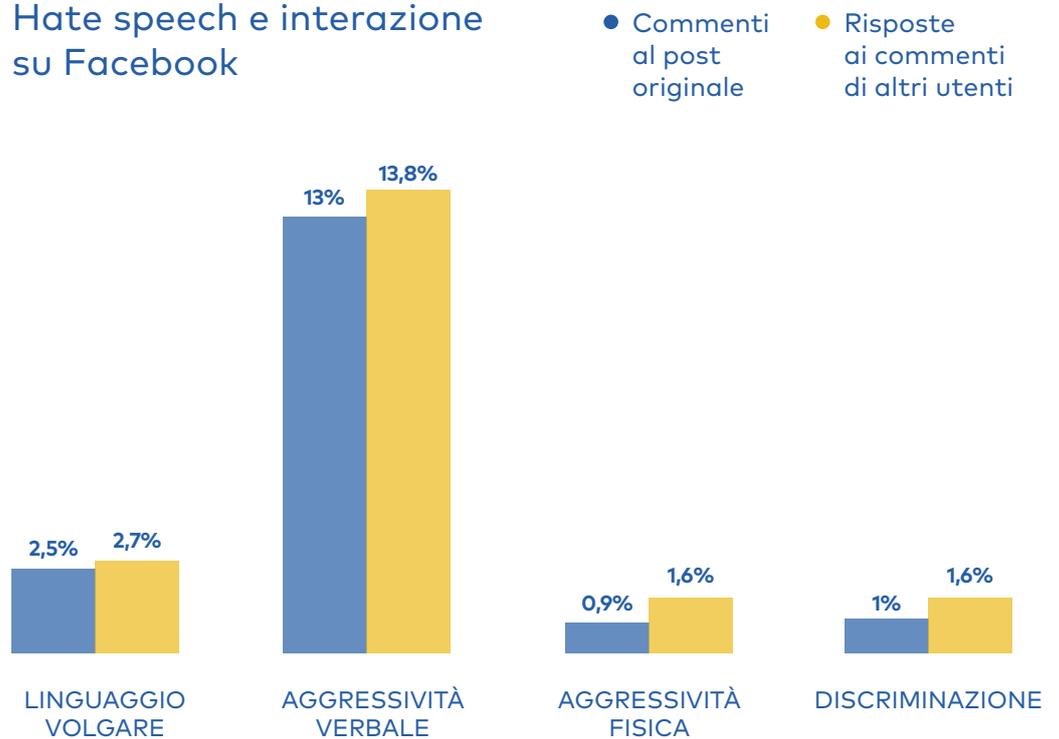


Un ultimo aspetto che merita di essere approfondito riguarda la struttura delle conversazioni per capire se l'*hate speech* si manifesti maggiormente in reazione o in relazione al contenuto di un post oppure se sia maggiormente collegabile all'interazione tra utenti. Facebook, fin dal 2012, ha introdotto la possibilità di commentare i commenti degli altri utenti, generando una struttura ad albero che parte dal commento al post ma si può ramificare in discussioni generate da un commento degli utenti. Questa opzione di interazione, non disponibile su Twitter dove tutti i commenti sono automaticamente riferiti al tweet iniziale, permette in maniera semplice di vedere se i commenti al post originale presentano un livello di *hate speech* più elevato rispetto alle risposte ai commenti di altri utenti.

La FIGURA 34 mostra che su Facebook l'*hate speech* è moderatamente più elevato nelle discussioni tra utenti rispetto ai commenti ai post, vale a dire che il maggior numero di riferimenti al linguaggio d'odio non risulta correlato al contenuto dei post ma sembra piuttosto svilupparsi in reazione alle prese di posizione di altri utenti.

FIGURA 34

Hate speech e interazione su Facebook



INTERAZIONI SU FACEBOOK

+ hate speech

- ↳ discussioni tra utenti

+ riferimenti d'odio

- ↳ prese di posizione di altri utenti

Questo è particolarmente vero per quei commenti che contengono insulti e offese, siano esse di natura verbale, fisica o discriminatoria. Seppur con differenze non particolarmente significative, le conversazioni tra utenti sembrano il luogo in cui l'*hate speech* si sviluppa maggiormente rispetto al commento diretto al contenuto del post: la litigiosità tra utenti produce più *hate speech* del commento alla notizia sportiva.



CAPITOLO 10

Conclusioni

- Facebook e Twitter si caratterizzano per essere due ambienti diversi in termini di numero di commenti: a parità di messaggi pubblicati, Facebook genera un volume di commenti 26 volte superiore a quello di Twitter.
- I due social media hanno anche livelli diversi di *hate speech*: su Twitter la presenza è percentualmente più elevata, ma la quantità di commenti contenenti odio bassa; su Facebook la presenza percentuale è più bassa, ma il numero di commenti con riferimenti all'*hate speech* elevato.
- Il calcio è sovrarappresentato come sport nel flusso comunicativo e anche nel volume totale di commenti. Basket, Formula 1 e tennis sono gli unici altri sport con un volume di commenti sufficiente per essere analizzato.
- Il calcio è lo sport a cui sono associati il numero più alto di commenti di *hate speech*: 12,3% su Facebook e 31,3% su Twitter.
- Tra le sei squadre di calcio con associato il maggior numero di commenti con *hate speech* non emergono grandi differenze in termini di dimensioni rilevanti. Milan, Lazio, Roma, Juventus, Inter, Napoli sono prevalentemente associate con aggressività verbale e solo in percentuali minime con aggressività fisica o discriminazione.

- Tra i personaggi sportivi Balotelli e Lukaku sono quelli associati a livelli più alti di *hate speech* (rispettivamente 16,7% su Facebook e 38,3% su Twitter; 15,5% su Facebook e 40,6% su Twitter) e in particolare a insulti contenenti discriminazione razziale per il loro colore della pelle (rispettivamente 2,1% su Facebook e 5,6% su Twitter; 1,9% su Facebook e 2,4% su Twitter).
- Le partite di calcio e in particolare quelle in cui ci sono decisioni arbitrali controverse sono gli eventi a cui sono associati livelli più alti di *hate speech*.
- Il 13,4% dei commenti su Facebook e il 31% di quelli su Twitter contiene *hate speech*. Di questi, la maggior parte dei commenti non si riferisce al cuore della definizione di *hate speech*, ma riguarda invece l'uso di linguaggio volgare (14% e 31%) o di aggressività verbale (73% e 60%). Comparativamente, i riferimenti a discriminazione (7% e 5%) e aggressività fisica sono minori (5% e 4%). Questo non vuol dire tuttavia che siano una componente irrilevante poiché si tratta di circa 5.000 commenti di questo tipo pubblicati dagli utenti in un arco di tre mesi.
- L'*hate speech* ormai si caratterizza per essere non solo una parte rilevante ma una vera e propria dimensione strutturale delle conversazioni sportive tra utenti online.



Riferimenti bibliografici

- Andrisani, P. (2014). Il perverso intreccio tra odio virtuale e odio «virale», in Lunaria (a cura di) *Cronache di ordinario razzismo*. Terzo libro bianco sul razzismo in Italia. Roma.
- Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The "nasty effect:" Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19(3), 373-387.
- Antoci, A., Delfino, A., Paglieri, F., Panebianco, F., & Sabatini, F. (2016). Civility vs. incivility in online social interactions: An evolutionary approach. *PLoS One*, 11(11).
- Assimakopoulos, S., Baider, F. H., & Millar, S. (2017). *Online hate speech in the European Union: a discourse-analytic perspective*. London: Springer.
- Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication*, 64(4), 658-679.
- De Mauro, T. (2016). Le parole per ferire. *Internazionale*, www.internazionale.it/opinione/tullio-de-mauro/2016/09/27/razzismo-parole-ferire.
- Hmielowski, J. D., Hutchens, M. J., & Cicchirillo, V. J. (2014). Living in an age of online incivility: Examining the conditional indirect effects of online discussion on political flaming. *Information, Communication & Society*, 17(10), 1196-1211.
- Hopp, T., & Vargo, C. J. (2017). Does negative campaign advertising stimulate uncivil communication on social media? Measuring audience response using big data. *Computers in human behavior*, 68, 368-377.
- Sacristán, Ana (ed.). 2013. *Sociedad del conocimiento, tecnología y educación*. Madrid: Morata.
- Santana, A. D. (2014). Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism practice*, 8(1), 18-33.
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & behavior*, 7(3), 321-326.
- Yong, C. (2011). Does freedom of speech include hate speech?. *Res Publica*, 17(4), 385.
- Ziccardi, G. (2016). *L'odio online. Violenza verbale e ossessioni in rete*. Milano: Cortina.



Appendice: metodo e descrizione del dizionario utilizzato

Per analizzare il contenuto dei commenti ai post su Facebook e delle risposte ai tweet, abbiamo utilizzato uno degli strumenti più semplici e al contempo più efficaci messi a nostra disposizione dall'analisi quantitativa dei testi: la classificazione delle parole presenti in diverse categorie predefinite di *hate speech*. Questa procedura consiste in due stadi: (1) la creazione di un vocabolario contenente una lista di termini che assumiamo essere propri dell'*hate speech* e (2) la ricerca all'interno del corpus dei commenti e dei tweet di tali termini.

La creazione del vocabolario è ovviamente lo stadio più laborioso, dalla cui qualità dipende però l'esito dell'intera operazione. Con il termine "vocabolario" si intende una collezione di parole o brevi locuzioni associate a diverse categorie qualitative. L'esempio più classico di utilizzo di un vocabolario nell'analisi quantitativa di testi è la *sentiment analysis*, la tecnica attraverso la quale il livello di positività o negatività espresso da un testo viene ricavato dalla quantità di parole classificate come "positive" o "negative" presenti in esso. Nel nostro caso, per produrre un dizionario di *hate speech* abbiamo creato liste di parole partendo dalle diverse categorie di linguaggio d'odio definite nel Capitolo 3. La redazione della lista di parole è partita prendendo il vocabolario *Hurtlex* (Bassignana et al., 2018), una delle rare risorse per analisi di *hate speech* in lingua italiana. Da questo vocabolario abbiamo eliminato alcune parole a nostro avviso cadute in disuso (ad es. "gagliofo", "marrano"), e aggiunto alcuni sinonimi delle parole rimaste. In una seconda fase, abbiamo utilizzato i dati messi a disposizione per *HaSpeeDe* (*Hate Speech Detection*), una competizione promossa nel contesto della campagna *EVALITA 2018* organizzata dall'Associazione Italiana di Linguistica Computazionale ([qui](#) e [qui](#) per ulteriori informazioni). I dati *HaSpeeDe* consistono in una serie di post Facebook e tweet che sono stati classificati manualmente come contenenti *hate speech* o meno da due team di linguisti computazionali italiani (vedi Del Vigna et al., 2017 per quanto riguarda i dati Facebook, e Poletto et al., 2017 per quanto riguarda i dati Twitter). Usando delle tecniche di *machine learning*, abbiamo identificato le parole con il più alto valore discriminante (le parole la cui presenza predice con maggiore successo la probabilità che un post o un tweet sia classificato come contenente *hate speech*), e utilizzato queste per aggiornare il dizionario. In una terza fase, abbiamo espanso il dizionario includendo termini da noi valutati come propri di linguaggio d'odio nelle diverse categorie.

La divisione in categorie e sottocategorie risultante, con alcuni esempi per ognuna di esse, è riportata nella TABELLA A1. La tabella mostra come le quattro categorie generiche di *hate speech* si compongono da una o più sottocategorie più specifiche (per esempio la categoria Discriminazione è l'aggregato di diversi tipi specifici di discriminazione verbale, come il *body shaming*, la discriminazione etnica o religiosa, l'omofobia, ecc.). Per ognuna di queste sottocategorie, la terza colonna della tabella mostra alcuni esempi. Notare come il vocabolario include parole vere e proprie oppure "temi" (tutti i termini che finiscono con l'asterisco), ovvero le parole private delle desinenze. Il vantaggio di utilizzare i temi, invece delle parole complete, è che in questo modo è possibile identificare tutte le possibili coniugazioni delle parole cercate. Inoltre, il vocabolario contiene anche alcune brevi locuzioni, come ad esempio "*tua_madre*". Questo ci permette di identificare un insulto direzionato, con maggiore precisione che includendo semplicemente il termine "madre". In totale, il vocabolario contiene 886 termini, di cui 53 si ripetono in più di una categoria (ad esempio "*affanculo*" conta sia come Linguaggio Volgare che come Insulto Generico). La quarta colonna della tabella mostra il numero di termini in ogni categoria (e la percentuale rispetto al totale dei termini nel vocabolario).

Per determinare la capacità del vocabolario di riconoscere effettivamente il contenuto di *hate speech* dei commenti Facebook e dei tweet, abbiamo analizzato i commenti e i tweet presenti nei dati HaSpeeDe, confrontando la classificazione dei testi come "contenenti *hate speech*" o meno sulla base del vocabolario con la classificazione manuale operata dai team che hanno prodotto i dati HaSpeeDe. I risultati sono alquanto incoraggianti: nel caso dei commenti Facebook, la classificazione dei testi come *hate speech* utilizzando il nostro metodo si sovrappone al 76,3% alla classificazione manuale. Il che significa che, in più di 3 commenti su 4, il nostro vocabolario identifica la presenza di *hate speech* tanto accuratamente quanto lo farebbe un essere umano. Nel caso dei tweet, questo valore è leggermente più basso, ma sempre incoraggiante: 72,4%. Riassumendo, questo esercizio conferma che, sia per quanto riguarda Facebook che per quanto riguarda Twitter, il nostro vocabolario è un metodo valido di rilevazione di *hate speech* nei commenti degli utenti.

TABELLA A1

Categorie del vocabolario di *hate speech* con esempi

Categorie	Sottocategorie	Esempi	Numero di termini nel vocabolario
<i>Linguaggio Volgare</i>		<i>cazzo, figa, stronzat*</i>	127 (14.3%)
<i>Aggressività Verbale</i>	Insulto generico	<i>scemo, stronzo</i>	383 (43.2%)
	Insulto alla famiglia	<i>tua_madre, tua_nonna</i>	16 (1.8%)
	Insulto politico	<i>zecca, fascio, sinistroid*</i>	16 (1.8%)
	Insulto sportivo	<i>gobb*, rubentus</i>	35 (3.95%)
	Riferimento al disgusto	<i>lurid*, puzza, puzzolente</i>	32 (3.6%)
	Espressione di odio	<i>detesto, disprezzo, ti_odio</i>	9 (1%)
			<i>ammazz*, stupr*, bruci*</i>
<i>Aggressività Fisica</i>	Body shaming	<i>ciccion*, nano, tappo</i>	15 (1.7%)
<i>Discriminazione</i>	Abilismo	<i>spastic*, storpio, storpia</i>	46 (5.2%)
	Discriminazione etnica o religiosa	<i>negr*, terron*, rabbin*, rom</i>	35 (3.95%)
	Omofobia	<i>finocc*, froci*, lesbic*</i>	21 (2.4%)
	Sessismo	<i>puttana, cagn*, donnacc*</i>	73 (8.2%)

RIFERIMENTI BIBLIOGRAFICI DELLA TABELLA A1

Bassignana, E., Basile, V., & Patti, V. (2018). Hurtlex: A Multilingual Lexicon of Words to Hurt. Proceedings of the Fifth Italian Conference on Computational Linguistics, CLiC-It 2018. 5th Italian Conference on Computational Linguistics, CLiC-it 2018, Turin, Italy. <http://ceur-ws.org/Vol-2253/paper49.pdf>

Del Vigna, F., Cimino, A., Dell'Orletta, F., Petrocchi, M., & Tesconi, M. (2017). Hate me, hate me not: Hate speech detection on Facebook. Proceedings of the First Italian Conference on Cybersecurity (ITASEC17). First Italian Conference on Cybersecurity (ITASEC17), Venice, Italy. <http://ceur-ws.org/Vol-1816/paper-09.pdf>

Poletto, F., Stranisci, M., Sanguinetti, M., Patti, V., & Bosco, C. (2017). Hate Speech Annotation: Analysis of an Italian Twitter Corpus. Proceedings of the Fourth Italian Conference on Computational Linguistics, CLiC-It 2017. 4th Italian Conference on Computational Linguistics, CLiC-it 2017, Rome, Italy. <http://ceur-ws.org/Vol-2006/paper024.pdf>



Odiare non è uno Sport

è progetto educativo per
prevenire e contrastare
l'hate speech nello sport

SCOPRI IL PROGETTO →



UNIVERSITÀ
DEGLI STUDI
DI TORINO

