

**Statu quo e misure contro i discorsi d'odio in Internet:
panoramica e raccomandazioni**

Dr. Lea Stahel
Istituto di sociologia, Università di Zurigo

Agosto 2020



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Dipartimento federale dell'interno DFI

Segreteria generale DFI

Servizio per la lotta al razzismo

INDICE

1	SINTESI	1
2	INTRODUZIONE	4
3	DEFINIZIONI	5
3.1	Discorsi d'odio	5
3.2	Razzismo	8
3.3	Social media	8
3.4	Distinzione dei discorsi d'odio razzisti online.....	9
3.4.1	Dalla discriminazione digitale strutturale	9
3.4.2	Da altri attacchi online diretti	10
4	FREQUENZA DEI DISCORSI D'ODIO RAZZISTI IN RETE.....	11
4.1	Dati disponibili e situazione internazionale	11
4.2	Situazione in Svizzera	13
4.3	Bambini e adolescenti: un gruppo particolare	14
5	AMBIENTE DIGITALE DEI DISCORSI D'ODIO RAZZISTI.....	16
5.1	Grado di organizzazione degli autori	16
5.2	Fattori d'influenza individuali e sociali	17
5.3	Principali piattaforme	19
5.4	Infrastruttura online che favorisce il razzismo	22
5.4.1	Condizioni comunicative.....	22
5.4.2	Architettura delle piattaforme online	24
6	CONSEGUENZE NEGATIVE DELL'ODIO RAZZISTA ONLINE	29
6.1	Osservazione preliminare sull'amplificazione digitale	29
6.2	Vittime	29
6.3	Pubblico e società nel suo complesso.....	30
7	CONTROMISURE, ATTUAZIONE ED EFFICACIA	32
7.1	Legislazione e giurisprudenza in materia di discorsi d'odio	32
7.2	Gestori di social media	33
7.3	Media tradizionali in Svizzera	35
7.4	Attività di ricerca in Svizzera.....	36
7.5	Società civile in Svizzera e all'estero.....	37
7.5.1	Prevenzione e sensibilizzazione.....	38
7.5.2	Segnalazione e consulenza.....	41
7.5.3	Monitoraggio.....	43
7.5.4	Contronarrazione	46
8	SERVIZI SVIZZERI: SFIDE ED ESIGENZE	50
8.1	Servizi amministrativi	50
8.2	Consultori pubblici (o finanziati con fondi pubblici).....	51
8.3	Organizzazioni private	53
9	MISURE DI PREVENZIONE PER LA SVIZZERA.....	54
9.1	Competenze e funzioni	54
9.2	Strumento nazionale di segnalazione.....	55
9.3	Gruppi target e loro raggiungibilità	55
9.4	Criteri per la valutazione dei progetti presentati	57
10	CONSULENZA E INTERVENTI IN SVIZZERA	59
10.1	Conoscenze e competenze di cui i consultori devono disporre	59
10.1.1	Consulenza nei casi più semplici.....	59
10.1.2	Consulenza in casi più complessi e intervento.....	59
10.2	Misure per far conoscere e rendere fruibili i servizi di consulenza	60
11	RACCOMANDAZIONI PRATICHE	61
12	ELENCO DELLE ABBREVIAZIONI.....	64
13	BIBLIOGRAFIA.....	66

1 SINTESI

Come affrontare in futuro i discorsi d'odio razzisti online? Il presente rapporto, commissionato dal Servizio per la lotta al razzismo (SLR), offre solidi spunti: un riassunto critico dei dati disponibili, un ampio squarcio del mondo digitale e delle condizioni di comunicazione che favoriscono il fenomeno e una panoramica delle contromisure finora adottate in Svizzera e all'estero. Nella parte finale sono esposte le sfide e le esigenze dei servizi amministrativi, dei consultori e delle organizzazioni private del nostro Paese e formulate raccomandazioni.

Il fenomeno dei discorsi d'odio in Internet è complesso e dinamico. Il termine «discorsi d'odio» definisce affermazioni che discreditano o denigrano determinati gruppi o loro appartenenti. Nel caso della discriminazione razziale, gli esseri umani sono suddivisi in gruppi supposti naturali secondo le loro peculiarità etniche, nazionali, culturali o religiose in base alle quali sono poi trattati ingiustamente, screditati o umiliati. L'espressione «discorsi d'odio razzisti online» è dunque riferita ai discorsi d'odio caricati in rete che discriminano persone in forma diretta denigrandole in base alla loro appartenenza. Il rapporto non approfondisce per contro la discriminazione indiretta o strutturale nello spazio digitale né gli attacchi diretti online come il cybermobbing o il cyberstalking.

I discorsi d'odio razzisti, sempre più diffusi in rete, suscitano nella popolazione interesse e indignazione, ma anche perplessità. L'attribuzione dell'etichetta di «discorso d'odio» dipende da diversi fattori storici, culturali e giuridici. Secondo il contesto, la stessa affermazione può essere vista come non problematica in una situazione ed essere considerata un discorso d'odio in un'altra. I discorsi d'odio dipendono dal contesto, ragion per cui non ne esiste una definizione universalmente riconosciuta. Per affrontare efficacemente il fenomeno non si può non tener conto dell'eterogeneità delle sue interpretazioni.

Con quale frequenza sono postati discorsi d'odio? La dipendenza dal contesto rende difficile dare una risposta a questa domanda. Spesso sono utilizzate definizioni e fonti diverse, il che non giova alla comparabilità dei dati rilevati. Inoltre, i discorsi d'odio razzisti online possono essere tanto sottovalutati a causa di una loro «normalizzazione» a lungo termine quanto sopravvalutati per via della loro forte visibilità mediatica. Un'importante fonte di dati sono attualmente le analisi dei contenuti online (p. es. commenti, video). La quota dei discorsi d'odio sull'universo statistico di base varia secondo i Paesi e le piattaforme dall'1 al 20 per cento. D'altro canto, sondaggi sovranazionali mostrano che fino a due terzi della popolazione constatano la presenza in Internet di discorsi d'odio, in gran parte discorsi d'odio razzisti. Da una a quattro persone su dieci sono già state oggetto di attacchi personali. Bambini e giovani sono più rappresentati della media sia tra gli spettatori, che tra gli autori e tra le vittime. Per la Svizzera non sono disponibili dati paragonabili rilevati sistematicamente. Un confronto tra dati di diverse fonti mostra tuttavia che anche in Svizzera succede molto online, ma viene segnalato poco. Ufficialmente sono rilevate alcune decine di casi di consulenza e di casi giuridici. Eppure, da sondaggi sugli attacchi online in generale (non soltanto di stampo razzista) risulta che le vittime sono molte di più.

Ma chi sono gli autori? Nel complesso sistema delle piattaforme digitali i confini tra gruppi di odiatori organizzati, reti «fluide» come il movimento statunitense alt-right (destra alternativa) e singoli individui sono sfumati. Nonostante la grande varietà di «tipi» di autore, dagli studi scientifici svolti emerge tuttavia un identikit ricorrente: si tratta perlopiù di maschi pieni di pregiudizi, diffidenti, privi di scrupoli o impulsivi che si informano su una vasta gamma di media populistici di destra, frequentano piattaforme di odiatori o sono stati essi stessi vittima di discorsi d'odio, cercano attenzione e intrattenimento o vogliono difendere il proprio gruppo. Delle loro caratteristiche sociostrutturali (p. es. famiglia, reddito, livello d'istruzione) si sa poco. È tuttavia chiaro che gli sviluppi sociali (p. es. rivolgimenti economici, politici o culturali) influenzano a lungo termine tanto i discorsi d'odio offline che quelli online e che singoli avvenimenti, per esempio attentati, possono provocarne un picco temporaneo.

I discorsi d'odio razzisti possono annidarsi in tutti gli spazi virtuali di Internet, ma alcuni sono più adatti di altri ad esprimerli e diffonderli, in particolare siti, blog, forum (p. es. 4chan), piattaforme social (p. es. Facebook, Twitter), videogiochi, mail e chat (p. es. WhatsApp). Le condizioni della comunicazione online abbattano barriere che nella vita reale arginano i discorsi d'odio. Gli autori sono più disinibiti perché la distanza dagli interlocutori li fa sentire meno visibili. Poiché s'incontrano persone e opinioni di diversa provenienza sociale e geografica, i conflitti sono inoltre più probabili. Gli autori

possono dare immediatamente pubblico sfogo ai propri risentimenti: l'attenzione che otterranno sarà tanto maggiore quanto più semplicistici e drammatizzanti saranno i loro contenuti.

In questo contesto comunicativo le architetture delle piattaforme mettono a disposizione degli utenti molteplici strumenti di diffusione (p. es. video, hashtag), che possono essere usati anche per produrre rapidamente e diffondere su larga scala contenuti razzisti. Gli autori di discorsi d'odio online sanno sfruttare al meglio l'estremo dinamismo delle strutture interattive e partecipative delle piattaforme social, dei servizi di messaggistica e dei forum per linkare i loro contenuti, creare reti e ispirarsi a vicenda. Spesso attivi su più di una piattaforma, sviluppano fitte reti nelle quali possono diffondere discorsi d'odio e propaganda rapidamente, a largo raggio e praticamente gratis. Inoltre, possono utilizzare profili fasulli e bot (programmi informatici) per far credere di rappresentare una vasta maggioranza (manipolazione metrica) o postare fake news e teorie del complotto semplicistiche ma di forte impatto emotivo, che generano molti click e quindi una larghissima visibilità (manipolazione contenutistica). Sovente dissimulano contenuti razzisti affinché non vengano riconosciuti come tali dagli internauti e dagli appositi algoritmi (p. es. con memi umoristici). Infine, introducono sotteraneamente l'ideologia razzista tra il vasto pubblico digitale integrandola nel panorama considerato legittimo (*mainstream*).

I discorsi d'odio razzisti online non hanno conseguenze negative soltanto su chi ne è vittima, ma anche sugli spettatori e sulla società. Le conseguenze per le vittime sono paragonabili a quelle del razzismo offline, ma alcuni aspetti del razzismo digitale sono potenzialmente peggiori. Le vittime sono esposte a un pubblico ben più numeroso, non hanno alcuna possibilità di sfuggire agli attacchi né nello spazio né nel tempo e devono convivere con messaggi d'odio permanentemente visualizzabili. Una percentuale significativa di queste persone dichiara di dover combattere con disturbi non soltanto emotivi, ma anche fisici. Inoltre, molti si ritirano regolarmente dalla vita digitale. Secondo i risultati di alcuni esperimenti, i discorsi d'odio intaccano anche la coesione sociale, in quanto inducono gli spettatori a pensare e ad agire in maniera più ostile. Questo danneggia la cultura della discussione democratica online. Infine, vi sono indizi che i discorsi d'odio online costituiscano un terreno fertile per i delitti d'odio offline.

Per combattere i discorsi d'odio online, tanto in Svizzera quanto all'estero, attori come le imprese mediatiche, i gestori di social media, il mondo scientifico, le istituzioni giuridiche e la società civile adottano diverse contromisure. I media classici possono arginarli con la qualità dei loro servizi giornalistici, soggetti a linee guida istituzionalizzate, e con la gestione dei commenti delle loro community. I gestori di social media sono meno regolamentati, ma gli utenti possono segnalare i discorsi d'odio direttamente sulle piattaforme. La trasparenza dei criteri di verifica e la qualità dell'esame automatizzato devono tuttavia ancora essere valutate criticamente. La ricerca contribuisce costantemente alla comprensione del fenomeno, ma i dati sul contesto svizzero sono (ancora) scarsi. Nella giurisprudenza svizzera l'applicazione della norma penale contro la discriminazione razziale (art. 261^{bis} CP) ai discorsi d'odio digitali è in continua evoluzione. Tuttavia, poiché i discorsi d'odio online sono spesso opera di reti internazionali, perseguirli con le nostre leggi può essere molto difficile.

In Svizzera come all'estero, la società civile attua contromisure in quattro settori centrali: prevenzione e sensibilizzazione, consulenza e segnalazione, monitoraggio e contronarrazione. All'estero la maggior parte degli approcci è consolidata, in Svizzera siamo ancora agli inizi. Gli approcci di prevenzione e sensibilizzazione intendono impedire i discorsi d'odio online o contenerne gli effetti promuovendo la competenza mediale. Lo scopo è una società civile digitalmente matura. La loro efficacia a lungo termine è difficile da stimare, ma studi svolti in materia hanno riscontrato primi effetti positivi a breve termine. In caso di attacchi concreti, i servizi di segnalazione e consulenza offrono consulenza psicologica, sociale e/o giuridica e strumenti di segnalazione. Questi strumenti dovrebbero essere quanto più conosciuti e quanto più facilmente accessibili. Il monitoraggio, che segue e descrive i discorsi d'odio nei contesti locali, fornisce un quadro quantitativo delle forme e delle tendenze del fenomeno e dati qualitativi sulle sue caratteristiche. Il monitoraggio quantitativo è informativo, ma oneroso nell'attuazione. Nella contronarrazione, gli utenti impiegano testi, immagini e video per confutare i discorsi d'odio con argomenti oggettivi. L'obiettivo è rafforzare norme come il rispetto e l'oggettività e segnalare solidarietà alle vittime. In linea di massima, la contronarrazione sembra una contromisura efficace, ma anche in questo caso non vanno sottovalutati gli oneri di una sua attuazione a lungo termine.

In Svizzera, i servizi del settore dispongono delle più disparate competenze e facoltà per la lotta al razzismo. I servizi amministrativi, in particolare, possono informare, sensibilizzare, creare reti e versare sussidi. I consultori possono fornire sostegno nei casi più semplici e trasmettere conoscenze ad attori e istituzioni. Grazie alla loro indipendenza finanziaria, le organizzazioni private possono svolgere una

gamma più ampia di compiti, tra cui la contronarrazione e l'impegno nei casi, più complessi, con implicazioni politiche. Per quanto riguarda i discorsi d'odio razzisti online, i servizi dispongono di sensibilità e competenza in materia di razzismo, ma hanno conoscenze molto scarse sulla dimensione digitale del fenomeno e mancano quindi delle competenze necessarie per una consulenza efficace.

Per dotare i consultori delle conoscenze e competenze necessarie, bisogna distinguere tra consulenze semplici e consulenze complesse (interventi). Per le prime si raccomanda lo sviluppo di conoscenze tematiche, tecniche e giuridiche di base. Per le seconde sono richieste anche altre competenze, tra cui una buona rete di contatti e conoscenze delle dinamiche mediatiche.

Il rapporto raccomanda diverse misure per impedire e arginare i discorsi d'odio online in Svizzera. Le misure attualmente in atto sono ritenute suscettibili di molti miglioramenti. Una vasta opera di sensibilizzazione e il rafforzamento delle competenze tematiche, tecniche e giuridiche sono un inizio. In futuro tutte le misure contro il razzismo dovrebbero sempre tener conto anche dello spazio digitale. Progetti contro il razzismo dovrebbero essere valutati anche secondo criteri riferiti ai discorsi d'odio online. Questi progetti, infatti, richiedono per esempio diverse competenze (p. es. tecniche), diversi canali per raggiungere i gruppi target, un lavoro di pubbliche relazioni e misure di protezione. Inoltre bisognerebbe elaborare misure specifiche per lo spazio digitale. Per elaborare misure adeguate, i gruppi target andrebbero definiti suddividendo accuratamente le persone vulnerabili secondo il loro ruolo (vittima, autore, spettatore), la loro età (in particolare bambini e giovani) e altre caratteristiche di gruppo (p. es. migranti, politici, ONG) e identificandone le abitudini medial, spesso soggette a un'evoluzione dinamica. A questo scopo è necessario investire nella ricerca e nel monitoraggio. Le organizzazioni statali e non statali dovrebbero collaborare più strettamente: per combattere efficacemente il complesso problema dei discorsi razzisti online è richiesto un approccio complessivo che tenga conto delle esigenze e delle competenze delle più disparate istituzioni e ne sfrutti al meglio le possibilità di azione. Si dovrebbero dunque adottare soluzioni olistiche e misure scientificamente fondate. Si raccomanda infine di sviluppare e pubblicizzare consulenze e interventi locali e nazionali, tra l'altro uno strumento di segnalazione nazionale. Grazie a tutte queste misure, lo squilibrio tra i raffinati discorsi d'odio digitali e le pertinenti contromisure, spesso ancora riferite al mondo analogico, a lungo termine potrebbe essere compensato.

2 INTRODUZIONE

«Qualsiasi cosa abbiamo creato con il web, quello che mi preoccupa di più è che abbiamo permesso al razzismo di muoversi liberamente sul pianeta, e le conseguenze per la civiltà e la democrazia sono molto, ma molto spiacevoli se non vengono affrontate.»

Tim Berners-Lee, fondatore del World Wide Web¹

Oramai per capire il razzismo bisogna capire anche Internet. Perché oggi il razzismo non è più (soltanto) il razzismo apertamente ostentato da teste rasate con stivali anfibi, ma può diffondersi in Internet dissimulato in una miriade di contenuti e simboli. Non che Internet sia di per sé distruttivo, anzi: grazie alle nuove possibilità di comunicazione che offre, possiamo fruire di una più ampia gamma d'informazioni e opinioni. Basti pensare ai movimenti democratici o di emancipazione supportati dai social media come la Primavera araba, Occupy Wall Street o #metoo. Anche gli svizzeri comunicano sempre più tramite Internet: nel 2019 il 92 per cento ha usato Internet e il 66 per cento le reti social². Se l'ideologia razzista continuerà a diffondersi nello spazio digitale, c'è il rischio che prima o poi diventi «presentabile». Eppure, le conseguenze negative dei discorsi d'odio razzisti per le vittime, gli spettatori e la società sono ormai ben documentate. I discorsi d'odio online devono tuttavia essere affrontati diversamente dai discorsi d'odio offline. Sul come stanno discutendo vivacemente Governi, gestori di social media, organizzazioni della società civile e ricercatori.

Al giorno d'oggi, gli incarichi d'informazione, prevenzione e consulenza contro il razzismo dovrebbero tener conto anche della dimensione digitale del fenomeno. Le disposizioni repressive da sole non bastano: i discorsi d'odio razzisti online, infatti, spesso si muovono al di sotto della soglia del penalmente perseguibile, ma rappresentano ugualmente una minaccia per la convivenza sociale. Alla luce di queste constatazioni, il SLR intende fornire ai consultori specializzati in razzismo e discriminazione razziale le competenze necessarie per affrontare i discorsi d'odio razzisti online. Inoltre, nel quadro di una prioritizzazione degli aiuti finanziari sosterrà maggiormente apposite misure dei più svariati tipi, tra cui progetti per la sensibilizzazione del pubblico, misure di prevenzione, iniziative di organizzazioni statali e private e offerte d'intervento (p. es. nei Comuni). Per disporre delle basi necessarie per il raggiungimento di questi obiettivi, il SLR ha commissionato il presente rapporto, con l'incarico di rispondere alle seguenti domande:

1. In che forme si manifestano i discorsi d'odio online? Come vengono favoriti da Internet? Che effetti hanno sulle vittime, gli spettatori e la società?
2. Quali contromisure sono state adottate in Svizzera e all'estero? Come sono attuate? Qual è la loro efficacia?
3. A quali sfide sono confrontati i servizi statali e non statali attivi in Svizzera che si occupano di discorsi d'odio razzisti online? Quali sono le loro esigenze?
4. Quali misure di prevenzione potrebbero essere attuate in Svizzera? Da chi? In base a quali criteri possono essere valutati i progetti presentati?
5. Di quali conoscenze e competenze hanno bisogno i consultori per poter affrontare i discorsi d'odio razzisti online? Come possono pubblicizzare e quindi rendere fruibile la loro offerta?

Il rapporto, la cui stesura è stata accompagnata da un gruppo di esperti³, è fondato sulle più recenti acquisizioni della ricerca, su rapporti d'attività, su informazioni online di organizzazioni rilevanti e su interviste con esperti svizzeri ed esteri⁴.

¹ Tim Berners-Lee, The Guardian, 12 marzo 2017: [I invented the web. Here are three things we need to change to save it.](#)

² Latzer, Büchi e Festic, 2019a.

³ Eva Wiesendanger (SLR), Michele Galizia (SLR), Alma Wiecken (SLR), Nora Refaeil (avvocata e mediatrice, vicepresidente della CFR), Nina Hobi (UFAS), Magdalena Küng (SLR).

⁴ Michael Bischof (Promozione dell'integrazione, Città di Zurigo), Eveline Lüönd (Segreteria Protezione dalla discriminazione Svizzera centrale), Estelle Rechsteiner (Cardis – Centro Ascolto Razzismo e Discriminazione,

3 DEFINIZIONI

3.1 Discorsi d'odio

Per i discorsi d'odio (ted. *Hassrede*; fr. *discours de haine*; ingl. *hate speech*) non c'è una definizione universalmente riconosciuta⁵. Il presente sottocapitolo è dunque un tentativo di esporre come la scienza, il diritto, le istituzioni internazionali e i gestori di social media lo intendono. Vista l'eterogeneità delle interpretazioni, il presente rapporto presuppone una definizione in senso lato del concetto, che considera un iperonimo di tutte le affermazioni intese a screditare o denigrare gruppi o singoli appartenenti a gruppi. Il concetto di «discorsi d'odio» rende più universale e più incisiva la discussione sul discredito e la denigrazione di gruppi o appartenenti a gruppi, anche se rischia di annacquare (in confronto a concetti specifici come antisemitismo, razzismo ecc.)⁶.

La *scienza* ha affrontato il tema del significato del concetto e delle difficoltà di definirlo. Il concetto è composto dalle parole «odio» e «discorso». L'«odio» è un «sentimento permanente di forte antipatia nel senso di un atteggiamento di fondo ostile» ed è la forma più intensa dell'avversione⁷. L'odio vuole l'annientamento dell'«odiato» e va al di là dell'ira, che è determinata da situazioni contingenti. La parola «discorso», invece, descrive un atto linguistico, cioè un'«affermazione linguistica» «pronunciata, scritta o comunicata mediante segni, immagini, gesti o mimica che in un contesto sociale specifico hanno un determinato significato»⁸. Possono essere vittima di discorsi d'odio diversi gruppi, sulla base di categorie come «razza», religione, orientamento sessuale ecc., in particolare chi è storicamente e/o attualmente oggetto di oppressione e stigmatizzazione. Questo non significa tuttavia che chi in questo momento occupa una posizione di potere (p. es. politico) ne sia al riparo⁹. Gli attacchi possono avvenire anche sulla base di più caratteristiche, per esempio colore della pelle e sesso (*intersezionalità*)¹⁰.

Il concetto di «discorsi d'odio» è difficile da definire anche perché il fenomeno dipende molto dal contesto in cui si sviluppa e può manifestarsi in diversi modi. A seconda del contesto, per molti contenuti possono essere adeguate le più disparate definizioni. Ma il contesto raramente è considerato nelle definizioni¹¹. L'esclusione del contesto è evidente nel tentativo di riconoscere discorsi d'odio in contenuti online¹² con l'ausilio di liste di parole automatizzate. Analisi di questo tipo sollevano un ulteriore problema: chi decide se un'affermazione è un discorso d'odio? Le vittime? Gli autori? O forse terzi non coinvolti (p. es. ricercatori)? Specialmente nello spazio digitale le persone si muovono in contesti diversi, il che rende più difficili interpretazioni comuni. Un'ulteriore difficoltà è che alcune componenti definitorie (p. es. l'«intenzione di far male») non possono essere determinate inequivocabilmente o non sono riconoscibili in quanto dissimulate (cfr. «Camuffamento» nel cap. 5.4.2). Questi ultimi casi possono essere smascherati come discorsi d'odio soltanto da chi è a conoscenza del camuffamento. In ultima analisi, il concetto non deve nemmeno rinviare a quanto sanno già tutti, cioè che i «diffusori di discorsi d'odio online», che lo facciano con commenti, condivisioni o «mi piace», per principio non devono provarlo essi stessi – anche se il tono delle loro affermazioni trasuda d'odio. Né sono penalmente perseguibili tutti i discorsi d'odio. Se si vogliono riconoscere i discorsi d'odio, la regola «I know it when

Ticino), Jolanda Spiess-Hegglin (#Netzcourage), Dominic Pugatsch (GRA), Stéphane Koch (intelligentzia, Ginevra), Claire Pershan (Renaissance Numérique, Parigi), Jonathan Birdwell (Institute of Strategic Dialogue, Londra), Catherine Blaya (Università Nice Côte; LASALÉ della HEP Vaud, Losanna), Hansi Voigt (dasnetz.ch; bajour), Johannes Baldauf (esperto di radicalizzazione e discorsi d'odio in Internet, precedentemente per la fondazione Amadeu Antonio, attualmente per Facebook).

⁵ Naguib 2014: pag. 13 e 17.

⁶ Rafael e Ritzmann 2018: pag. 13.

⁷ Naguib 2014: pag. 82.

⁸ Naguib 2014: pag. 83.

⁹ Naguib 2014: pag. 84.

¹⁰ L'intersezionalità è la sovrapposizione di forme di discriminazione che genera esperienze di discriminazione a sé stanti. Per esempio, un'azione con movente sessista può essere strumentalizzata con una motivazione razzista (Naguib 2014: pag. 26).

¹¹ Sellars 2016: pag. 14 e 32.

¹² I contenuti online sono dati elaborati e messi a disposizione in forma digitale, inclusi video, contenuti audiovisivi, e testi, commenti, immagini e fotografie digitali.

I see it» («lo so quando lo vedo») è insufficiente¹³. Giuristi, organi internazionali e gestori di social media¹⁴ hanno proposto diverse definizioni.

Dal *punto di vista giuridico* i discorsi d'odio non sono un concetto consolidato e legalmente definito – nemmeno nella prassi delle autorità svizzere¹⁵. La mancanza di una definizione è «sintomatica per il discorso oramai politicizzato e in pieno sviluppo»¹⁶. Tutti i discorsi d'odio si esprimono in termini spregiativi – e, secondo la prospettiva, anche moralmente riprovevoli – ma non tutti sono ritenuti sufficientemente pericolosi da essere vietati per legge¹⁷. Fondamentalmente infatti vale il diritto alla libertà di espressione (p. es. art. 19 della Dichiarazione universale dei diritti umani), che però non è illimitato. Nel diritto svizzero, i discorsi d'odio possono rientrare nel campo d'applicazione di determinate disposizioni penali¹⁸. Sono rilevanti i divieti penali in materia di discriminazione razziale: si può dunque procedere legalmente contro i discorsi d'odio che prendono di mira «razze» o gruppi etnici o religiosi. Il divieto di discriminazione sulla base dell'orientamento sessuale è in vigore dal 1° luglio 2020 (art. 261^{bis} del Codice penale svizzero [CP]). Le fattispecie concrete criminalizzano l'incitamento all'odio o alla discriminazione, la diffusione di ideologie razziste, la preparazione di propaganda razzista e la negazione, giustificazione o minimizzazione di genocidi o altri crimini contro l'umanità.

Nonostante la difficoltà di definirli, vi sono Paesi che cercano sempre più di disciplinare i discorsi d'odio. Le loro definizioni si differenziano nella ponderazione dei singoli aspetti del fenomeno. Alcuni ritengono per esempio che i discorsi d'odio debbano essere vietati soltanto se proferiti in pubblico, altri che lo debbano essere sempre. Alcuni pensano che la veridicità dei discorsi sia rilevante, altri no. Per alcuni il danno dev'essere intenzionale, per altri basta la negligenza¹⁹. Queste differenze sfociano in ultima analisi in legislazioni più o meno «tolleranti»: relativamente permissiva, per esempio, quella americana, relativamente restrittiva quella tedesca²⁰.

Senza entrare nel merito delle legislazioni di singoli Stati, possono essere menzionati i tentativi di adottare standard internazionali dei più disparati *organi internazionali*²¹. Le loro definizioni sono in parte precedenti all'avvento dei social media, ma sono applicabili anche alle forme digitali dei discorsi d'odio. Il Consiglio dei Ministri del Consiglio d'Europa ha proposto una definizione in senso lato alla quale ancora oggi fanno spesso riferimento pubblicazioni scientifiche e non scientifiche:

«Ogni forma di espressione destinata a diffondere, incitare, promuovere o giustificare l'odio razziale, la xenofobia, l'antisemitismo o altre forme di odio basate sull'intolleranza, tra cui quella tesa a promuovere forme aggressive di nazionalismo e di etnocentrismo, di discriminazione e di ostilità verso minoranze, immigrati e persone di origine straniera.»²²

In questo caso, le sobillazioni possibili, oltre a quelle nazionaliste, razziste o religiose, sono soltanto sottintese. La Commissione europea contro il razzismo e l'intolleranza (ECRI), invece, specifica le forme espressive dei discorsi d'odio ed elenca ancora più gruppi concreti di vittime:

«(...) il fatto di fomentare, promuovere o incoraggiare, sotto qualsiasi forma, la denigrazione, l'odio o la diffamazione nei confronti di una persona o di un gruppo e di sottoporre a soprusi, insulti, stereotipi negativi, stigmatizzazione o minacce una persona o un gruppo e la giustificazione di tutte queste forme o espressioni di odio sulla base della «razza», del colore della pelle, dell'ascendenza, dell'origine nazionale o etnica, dell'età, di una disabilità, della

¹³ Sellars 2016: pag. 14.

¹⁴ Per una definizione di social media, cfr. cap. 3.3.

¹⁵ Naguib 2014: pag. 90.

¹⁶ Naguib 2014: pag. 91; l'autore rinvia al Rabat Plan of action (2012). Cfr. United Nations Human Rights. [Rabat Plan of action on the prohibition of advocacy of national, racial or religious hatred that constitute incitement to discrimination, hostility or violence](#). Rabat (4 e 5 ottobre 2012).

¹⁷ George 2015: pag. 1.

¹⁸ Naguib 2014.

¹⁹ Per un quadro più dettagliato delle analogie e delle differenze tra le legislazioni di diversi Paesi, cfr. Sellars 2016: pag. 18.

²⁰ Hawdon, Oksanen e Räsänen 2017: pag. 258.

²¹ Anche i discorsi d'odio in Internet sono stati tematizzati, cfr. Gagliardone et al. 2015.

²² [Council of Europe's Committee of Ministers' Recommendation 97\(20\)](#), 30 ottobre 1997.

lingua, della religione o delle convinzioni, del sesso, del genere, dell'identità di genere, dell'orientamento sessuale e di altre caratteristiche o stato personale.»²³

Nella Convenzione internazionale sull'eliminazione di ogni forma di discriminazione razziale (CERD)²⁴ sono proposti aspetti contestuali suscettibili di essere rilevanti per la punibilità di una specifica affermazione. Sono considerati il contenuto e la forma dell'affermazione (crudeltà, tipo di diffusione, stile), il clima economico, sociale e politico (schema di discriminazione dominante), la posizione di chi si esprime (p. es. *opinion leader*), la portata del discorso (tipo di pubblico e di canale di comunicazione) e il suo obiettivo (i discorsi a tutela dei diritti umani non dovrebbero p. es. essere punibili).

Da qualche tempo, anche i *gestori di social media* hanno preso a sviluppare definizioni sulla cui base vengono moderati sulle loro piattaforme i contenuti generati dagli utenti. Nelle sue norme (2020), la piattaforma video YouTube definisce come segue i «contenuti che incitano all'odio»:

«(...) contenuti che promuovono o giustificano la violenza contro singoli individui o gruppi per motivi legati a etnia, religione, disabilità, sesso, età, nazionalità, condizione di reduce di guerra, casta, orientamento o identità sessuale e i contenuti che incitano all'odio sulla base di tali caratteristiche.»²⁵

Analogamente, negli standard della sua community (2020), Facebook definisce come segue i «contenuti che incitano all'odio»:

«(...) discorsi di incitamento all'odio come un attacco diretto alle persone sulla base di aspetti tutelati a norma di legge, quali razza, etnia, nazionalità di origine, religione, orientamento sessuale, casta, sesso, genere o identità di genere e disabilità o malattie gravi. Offriamo protezione contro gli attacchi in base all'età se questa è associata a un'altra caratteristica protetta. Inoltre, offriamo determinate protezioni per lo stato di immigrato. Definiamo l'attacco come un discorso violento o disumanizzante, dichiarazioni di inferiorità o incitazioni all'esclusione o alla segregazione.»²⁶

Infine, nelle sue «Norme sulla condotta che incita all'odio» (2020), Twitter si rivolge così ai suoi utenti:

«Non puoi promuovere la violenza contro altre persone, attaccarle o minacciarle sulla base di razza, etnia, origine nazionale, casta, orientamento sessuale, sesso, identità sessuale, religione, età, malattia seria o disabilità.»²⁷

Il ricercatore americano A. Sellars ha tentato una definizione cercando e riunendo i denominatori comuni tra le definizioni del diritto e quelle delle imprese²⁸. Secondo la sua definizione, con ogni probabilità sarebbero considerati discorsi d'odio dalla maggior parte dei Paesi e dei gestori di social media tutti i discorsi che:

- attaccano un gruppo o individui che ne fanno parte;
- esprimono in qualche modo odio;
- intendono provocare danni;
- provocano danni (violenza fisica, ma anche in senso strutturale, p. es. nelle relazioni sociali);
- sobillano (p. es. incitamento alla violenza);
- sono pubblici o destinati a un membro del gruppo che si intende colpire;
- sono formulati in un contesto sociopolitico in cui sono possibili risposte violente;
- non hanno alcuno scopo «legittimo» al di là dell'odio.

²³ [ECRI General Policy Recommendation No. 15 on Combating Hate Speech](#), 8 dicembre 2015.

²⁴ [CERD, Allgemeine Bemerkungen Nr. 35 betreffend «Combating racist hate speech»](#), 26 settembre 2013.

²⁵ Youtube. [Norme e sicurezza](#). Percorso: Youtube > Informazioni > Norme e sicurezza.

²⁶ Facebook. [Contenuti che incitano all'odio](#). Percorso: Facebook > Standard della community > Contenuti deplorabili > Contenuti che incitano all'odio.

²⁷ Twitter. [Norme sulla condotta che incita all'odio](#). Twitter > Centro assistenza > Regole e norme > Regole e norme di Twitter > Norme sulla condotta che incita all'odio.

²⁸ Sellars 2016: pag. 24-31.

Per concludere, si può osservare che ognuna delle definizioni in uso ha le sue priorità e nessuna presenta tutti gli elementi citati, ma che il numero di elementi comuni è significativo²⁹. Si potrebbe quindi parlare di una «sfocatura definitoria» che offusca le conoscenze acquisite sui discorsi d'odio e le loro contromisure³⁰. Va tuttavia constatato che la dipendenza dei discorsi d'odio dal loro contesto non può permettere una definizione e una regolamentazione universalmente valida. L'eterogeneità delle interpretazioni è dunque insita nella natura dei discorsi d'odio. Per combatterli efficacemente non si può dunque prescindere dall'accettazione di questo dato di fatto.

3.2 Razzismo

Il presente rapporto è focalizzato sui discorsi d'odio razzisti. Del concetto di razzismo, politicamente e giuridicamente complesso, viene data qui soltanto una breve spiegazione. Secondo il SLR, il concetto di razzismo descrive:

«(...) un'ideologia che, fondata su una suddivisione degli esseri umani in gruppi supposti naturali («razze») in base all'appartenenza etnica, nazionale o religiosa, giustifica la supremazia di un gruppo sugli altri. Le persone non sono giudicate e trattate come individui, ma come appartenenti a gruppi pseudo-naturali con caratteristiche collettive ritenute immutabili.»³¹

La «razza» è dunque un costrutto sociale fondato su caratteristiche esteriori e/o presunte peculiarità culturali, religiose o inerenti all'origine. L'appartenenza a questi gruppi è utilizzata per spiegare differenze di condizione socioeconomica e di livello d'istruzione, che vengono quindi motivate biologicamente. Il razzismo può manifestarsi a livello interpersonale, strutturale, istituzionale e culturale³².

I discorsi d'odio razzisti riflettono in primo luogo una forma interpersonale, diretta e immediata di «discriminazione razziale», cioè ogni atto finalizzato a «privare persone dei loro diritti, a trattarle in modo ingiusto o intollerante, a denigrarle, a offenderle o a metterne in pericolo la vita e l'integrità fisica, soltanto a causa delle loro caratteristiche fisionomiche, etniche, culturali (lingua, nome) e/o religiose (reali o presunte)»³³. In questi casi, persone diverse che si trovano in situazioni comparabili, «per ragioni inammissibili, sono svantaggiate»³⁴. Comportamenti di questo genere non hanno tuttavia necessariamente un fondamento ideologico. Ed è proprio questa la differenza tra la discriminazione e il razzismo: gli atteggiamenti razzisti forse a monte della discriminazione possono certo essere stereotipati, ma non devono per forza essere riconducibili a convinzioni ideologiche.

Il presente rapporto tralascerà in gran parte la discriminazione indiretta³⁵, cioè le situazioni in cui «nonostante la loro apparente neutralità, pratiche, politiche o basi legali hanno come risultato una disparità di trattamento illecita»³⁶. Per «razzismo online», s'intende nel presente rapporto soprattutto il razzismo diretto dei discorsi d'odio razzisti in Internet («discorsi d'odio razzisti online»).

3.3 Social media

Il presente rapporto si focalizza sui discorsi d'odio razzisti in Internet, specialmente sui social media. I social media sono «piattaforme più o meno aperte, interattive e partecipative che permettono agli utenti di comunicare, costruire e intrattenere relazioni. Inoltre, con uno sforzo minimo, gli utenti dei social media possono scambiare informazioni e contenuti di e su terzi oltre che generare contenuti, sia personalmente che in cooperazione con altri, per poi renderli accessibili ad altri utenti»³⁷. I contenuti possono essere testi, immagini, fotografie, simboli, video e musica, ma anche collegamenti ipertestuali

²⁹ Sellars 2016: pag. 24 e 31.

³⁰ Siegel 2020: pag. 5.

³¹ Servizio per la lotta al razzismo 2019: pag. 10.

³² Bliuc et al. 2018: pag. 75; Krieger 1990.

³³ Servizio per la lotta al razzismo 2019: pag. 11.

³⁴ Servizio per la lotta al razzismo 2019: pag. 11.

³⁵ Per le sue manifestazioni in Internet cfr. la breve panoramica al cap. 3.4.1.

³⁶ Servizio per la lotta al razzismo 2019: pag. 11.

³⁷ Il Consiglio federale 2013: pag. 7.

e download. In Svizzera, i social media sono diffusi: nel 2019, gli svizzeri hanno trascorso in media 3 ore e 33 minuti al giorno in Internet e oltre il 70 per cento di loro è stato almeno una volta al giorno sui social media. Tra i quattordici-diciannovenni, la quota era del 99 per cento³⁸.

L'offerta di piattaforme social è molto ricca. L'utilizzo dipende dall'architettura della piattaforma e dalle possibilità di interazione³⁹. In confronto alla comunicazione precedente alla loro comparsa, gli esperti constatano uno slittamento dei confini. I ruoli di produttore e consumatore di contenuti si confondono: oggi i profani non sono più «condannati» al consumo passivo di contenuti dei media ufficiali. Possono produrre autonomamente contenuti e diffonderli sulle piattaforme social. Sono cioè a un tempo produttori e consumatori. Inoltre, i canali per la comunicazione privata e per quella pubblica sono meno divisi; entrambi i tipi di comunicazione sono spesso praticati sulla stessa piattaforma. I dati non sono più memorizzati soltanto localmente, ma sono registrati in Internet e possono essere «riesumati digitalmente» in qualsiasi momento anche dopo il giorno in cui sono stati pubblicati. Passato, presente e futuro si confondono. E poiché la comunicazione globale non conosce barriere di gruppo né frontiere, scompaiono anche confini sociali e spaziali⁴⁰.

Poiché le piattaforme social si sviluppano dinamicamente, è difficile categorizzarle. Alcune vogliono trasmettere conoscenze e informazioni (p. es. Wiki, blog, forum specifici), altre intendono offrire intrattenimento o consentire di sperimentare mondi virtuali (p. es. i videogiochi o YouTube), di coltivare relazioni (p. es. Facebook, piattaforme di dating) o di gestire la propria identità e la propria reputazione (p. es. blogs, LinkedIn). Nel contesto del razzismo online sono significativi in particolare, anche se non soltanto, i seguenti tipi di piattaforma⁴¹:

- social media (p. es. Facebook, LinkedIn);
- microblog (p. es. Gab, Twitter, Weibo);
- forum (p. es. 4chan, 8chan, Reddit);
- pagine audiovisive (p. es. YouTube, Instagram, TikTok);
- piattaforme di social news su cui gli utenti possono caricare o offrire contenuti (p. es. BuzzFeed);
- piattaforme di messaggistica (p. es. Threema, WhatsApp, Telegram);
- piattaforme per la ricerca di un partner (p. es. Tinder, Parship).

Elementi tipici dei social media possono inoltre essere integrati in pagine preesistenti. Ne sono un esempio i giornali online, che non sono gestori di social media, ma offrono rubriche per i commenti e pulsanti di condivisione. Questo significa che il razzismo online non è soltanto un «problema dei social media», ma che si è piuttosto annidato in un sistema di piattaforme social e pagine preesistenti tra loro connesse. In teoria, tutte queste piattaforme possono agevolare il razzismo, nella misura in cui offrono condizioni di comunicazione e architetture utili allo scopo (cfr. cap. 5.4.1 e 5.4.2).

3.4 Distinzione dei discorsi d'odio razzisti online

3.4.1 Dalla discriminazione digitale strutturale

Il concetto di «razzismo online», che nel presente rapporto è utilizzato come sinonimo di «discorsi d'odio razzisti online», altrove può essere inteso in senso più lato. Il razzismo può infatti essere già insito nelle strutture più profonde di Internet. In passato questa discriminazione strutturale è stata discussa tra l'altro in relazione all'industria hightech (p. es. la composizione etnicamente distorta della dirigenza). Di conseguenza, anche certe piattaforme e certi algoritmi sono potenzialmente discriminatori. Tra questi si annoverano design di piattaforma (p. es. categorie razziste codificate nelle superfici utente e nei menu dropdown), intelligenza artificiale (p. es. riconoscimento del viso) e analisi predittive (p. es. predictive

³⁸ Latzer, Büchi e Festic 2019b: pag. 8.

³⁹ Il Consiglio federale 2013: pag. 7.

⁴⁰ Boyd 2010: pag. 10.

⁴¹ Per altri tipi di piattaforma in generale, cfr. Il Consiglio federale 2017.

policing⁴²). La disparità digitale strutturale può essere osservata anche tra gli utenti, in particolare nel caso dei videogiochi (p. es. sovrarappresentazione dei maschi bianchi tra i designer e i giocatori)⁴³. La discriminazione strutturale non è oggetto del presente rapporto. Tuttavia, pur essendo meno visibile dei discorsi d'odio espliciti e intenzionali, è almeno altrettanto incisiva, in quanto può promuovere azioni e abitudini razziste inconscie.

3.4.2 Da altri attacchi online diretti

Bisogna inoltre distinguere i discorsi d'odio razzisti online (e i discorsi d'odio online in generale) da attacchi diretti online analoghi. Anche questi ultimi infatti violano regole di comportamento, infrangono limiti e possono essere violenti, ma non in primo luogo discriminatori. Ne sono un esempio il cybermobbing (p. es. molestie attraverso i media digitali, frequenti in particolare nelle scuole), il vigilantismo digitale (giustizia fai da te in Internet, perlopiù denunce o gogna mediatica), il cyberstalking (persecuzione e molestie nei confronti di individui, gruppi o organizzazione mediante diffamazione o furto d'identità), e il trolling (provocazioni e azioni di disturbo di dibattiti online perlopiù subliminali)⁴⁴. Nella loro forma «incontaminata», tutti questi fenomeni si distinguono dal razzismo online soprattutto per il fatto che possono colpire tutte le persone e le istituzioni in egual misura. Eventuali denigrazioni vanno intese come dirette a individui specifici (in questi casi si tratta perlopiù di conflitti personali o di comportamenti ben precisi della persona attaccata). I contenuti e le intenzioni degli attacchi possono essere dei tipi più disparati. Il razzismo online, invece, è diretto contro determinate persone in quanto ritenute appartenenti a un determinato gruppo. Gli attacchi non sono dunque riferiti all'individuo, ma alla sua appartenenza. Nella prassi, tuttavia, i confini tra razzismo e altri attacchi diretti online sono più sfumati. Uno scolaro può, per esempio, essere oggetto di mobbing nella chat della classe anche con accenni denigratori al colore della sua pelle. E i troll possono diffondere contenuti razzisti spiegando però di essere apolitici e di voler «soltanto provocare»⁴⁵.

⁴² Termine che designa l'«attività predittiva della polizia», cioè «l'analisi di dati finalizzata a stimare la probabilità di futuri reati quale ausilio concreto alla pianificazione degli interventi della polizia». La base è l'elaborazione di grandi quantità di dati. Cfr. Öffentliche IT (2015). [Vorhersagende Polizeiarbeit](#) (marzo 2015). Percorso: Öffentliche IT > Trendschau > Vorhersagende Polizeiarbeit.

⁴³ P. es. Banaszczuk 2019: pag. 5; Daniels 2013: pag. 700; Shapiro 2019.

⁴⁴ Per una descrizione più dettagliata degli attacchi online, cfr. Fortuna e Nunes 2018: pag. 8.

⁴⁵ Marwick e Lewis 2017: pag. 4.

4 FREQUENZA DEI DISCORSI D'ODIO RAZZISTI IN RETE

Nel dibattito pubblico, non di rado le piattaforme dei social media sono descritte come roccaforti dell'odio⁴⁶. Spesso giornalisti, ricercatori e politici giustificano la loro attenzione per il fenomeno con il fatto che si è accentuato con il passare del tempo. Raramente però affermazioni del genere sono comprovate da dati empirici sistematici. Nel presente capitolo sarà quindi analizzato per quali motivi è così difficile determinare la frequenza del razzismo in rete e saranno presentati i dati disponibili.

4.1 Dati disponibili e situazione internazionale

Accertare l'effettiva frequenza degli episodi di razzismo online è difficile per i seguenti motivi:

- *Assenza di una definizione uniforme*: gli studi in materia denominano e definiscono in maniera eterogenea il fenomeno da quantificare, che spazia da fattispecie di reato come la sobillazione popolare a espressioni difficili da circoscrivere e spesso tutelate dalla libertà d'opinione come i discorsi d'odio. Questo incide sul numero di casi rilevati e rende difficile la comparazione delle cifre. La circostanza non è di per sé problematica, dato che la dipendenza dei discorsi d'odio dal loro contesto esclude una definizione universalmente valida. Tuttavia, quando si procede a una comparazione sistematica dei risultati degli studi scientifici, il contesto va tenuto in considerazione.
- *Interpretazione discordante dei discorsi d'odio in rete*: anche quando la terminologia utilizzata è la stessa («discorsi d'odio in rete»), i paragoni sono difficili, dato che il concetto dipende fino a un certo punto dalla sensibilità soggettiva. Pur avendo vissuto le stesse esperienze, alcune persone dichiarano per esempio di essere state vittima di «discorsi d'odio», mentre altre affermano di no, e anche gli stessi codificatori di contenuti online sono discordi su cosa debba essere considerato un discorso d'odio. Se e in che misura certi contenuti online vengano percepiti come discorsi d'odio dipende quindi dalle sensibilità, molto variabili, dei diversi Paesi, gruppi e individui⁴⁷. L'interpretazione diventa più univoca in caso di contenuti estremi, mentre diverge di più in quelli di media intensità, il che potrebbe dipendere da una maggiore frequenza di contenuti mascherati, sarcasmo e ironia⁴⁸. Inoltre sembra che alcuni gruppi tendano più di altri a considerare certi discorsi come intrinseci d'odio. Sondaggi effettuati negli Stati Uniti attestano per esempio una sensibilità maggiore tra le donne, le persone «politicamente moderate e liberali» e i membri di gruppi storicamente emarginati e svantaggiati come gli afroamericani⁴⁹.
- *Eterogeneità delle fonti di dati*: gli studi scientifici si basano sui risultati di sondaggi, contenuti online generici o specifici di determinate piattaforme oppure su contributi segnalati provenienti da tutta la rete. A causa di questo quadro frammentario, quando si operano delle generalizzazioni occorre tener conto dei limiti di ogni base di dati.
- *Sopravalutazione a causa dell'elevata visibilità*: vista la grande attenzione mediatica per i discorsi d'odio in rete, e in particolare per le campagne d'odio collettivo, è ipotizzabile che in generale la popolazione sopravvaluti la frequenza del fenomeno. Mancano prove per avvalorare questa tesi, ma processi simili sono stati riscontrati nelle ricerche sulla percezione del terrorismo. Se in passato i discorsi d'odio erano circoscritti al contesto privato o alle chiacchiere da bar, oggi hanno luogo nell'agorà digitale e risultano quindi più visibili – il che non vuol dire che siano più frequenti.
- *Cambiamento di percezione a lungo termine*: qualora il razzismo si dovesse diffondere ampiamente in Internet, potrebbe verificarsi una «normalizzazione» per cui non verrebbe più

⁴⁶ Lizza, New Yorker, 19 ottobre 2016: [Twitter's anti-semitism problem](#).

⁴⁷ Ciò emerge da 18 125 valutazioni relative a discorsi d'odio online espresse da internauti di 50 Paesi. Cfr. Salminen et al. 2018.

⁴⁸ Salminen et al. 2019: pag. 213 e 216.

⁴⁹ Kenski, Coe e Rains 2017: pag. 809; Costello et al. 2019; Tynes e Markoe 2010.

percepito come tale. A causa di questa «assuefazione», le risposte nei sondaggi potrebbero suggerire una diminuzione del fenomeno malgrado un suo costante aumento⁵⁰.

La frequenza dei discorsi d'odio razzisti in rete è determinata soprattutto sulla base di contenuti postati sulle piattaforme online e di sondaggi condotti in vari Paesi. Le indagini riguardano perlopiù i discorsi d'odio in generale e solo raramente si focalizzano in modo specifico su quelli razzisti. Poiché il razzismo è ripetutamente attestato come forma di discorso d'odio più diffusa (cfr. fig. 1), appare giustificato fare riferimento, nelle considerazioni seguenti, sia a statistiche sui discorsi d'odio in generale che a dati sui discorsi d'odio razzisti⁵¹.

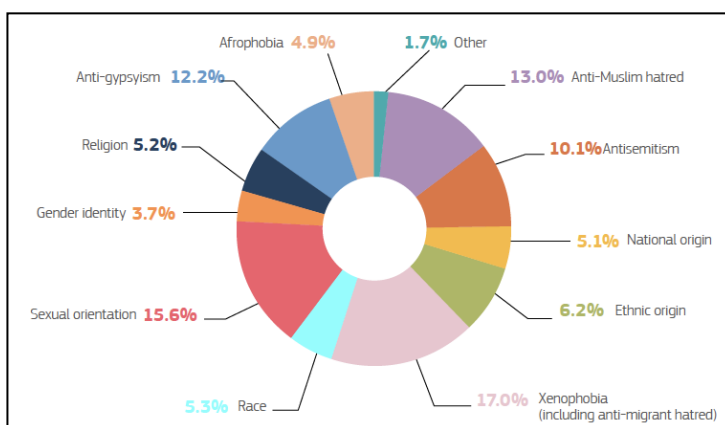


Figura 1: classificazione tipologica di 4392 contributi online segnalati ai gestori di social media nel quadro del *Code of Conduct* da utenti e segnalatori attendibili (*trusted flaggers*) tra il 5 novembre e il 14 dicembre 2018 (maggiori informazioni sul *Code of Conduct* nel cap. 7.2).

Contenuti postati sulle piattaforme online. I discorsi d'odio razzisti postati sulle piattaforme online sono riconosciuti e analizzati automaticamente. Siegel et al.⁵², per esempio, hanno trovato discorsi d'odio nell'1 per cento circa di oltre un miliardo di tweet di utenti statunitensi. Cifre simili (0,7 %) si riscontrano nelle pagine Facebook etiopi. Nella twittersfera italiana, la frequenza rilevata, pari al 15 per cento, è invece insolitamente elevata: il 10 per cento di questi messaggi era rivolto contro gli immigrati, il 6 per cento contro i musulmani e il 6 per cento contro gli ebrei⁵³. Questi dati suggeriscono che la quota dei discorsi d'odio sulla totalità dei contenuti online varia a seconda dei Paesi e delle piattaforme. Inoltre sembra che i discorsi d'odio riguardino soprattutto determinate comunità online: nel thread */pol/ (politically incorrect)* del popolare forum Internet 4chan, il 12 per cento dei contributi era per esempio costituito da discorsi d'odio⁵⁴. Queste cifre non ci dicono però con quale frequenza tali contenuti vengano visualizzati. Per giunta il confronto tra Paesi e piattaforme è difficile vista la diversità nelle pratiche di moderazione e cancellazione. La quantità di contributi pubblicamente accessibili e quindi analizzabili rappresenta dunque un indicatore soltanto parziale del numero di discorsi d'odio razzisti caricati in rete per essere diffusi. Oltre ai contenuti online, i ricercatori esaminano infine i profili sui social media: basandosi su questi ultimi, Guhl et al.⁵⁵ stimano che tra 15 000 e 50 000 persone di lingua

⁵⁰ Barnidge et al. 2019: pag. 8.

⁵¹ Jourová 2019a. Va ricordato ad ogni modo che la ripartizione nel grafico è il risultato di una selezione da parte dell'autrice delle segnalazioni degli utenti e non offre necessariamente un quadro rappresentativo della frequenza dei diversi tipi di discorso d'odio sui social media.

⁵² Siegel et al. 2019.

⁵³ Gagliardone et al. 2016; Lingiardi et al. 2019.

⁵⁴ Hine et al. 2017.

⁵⁵ Guhl, Ebner e Rau 2020: pag. 8.

tedesca con posizioni di estrema destra pubblicano un numero variabile di contenuti su piattaforme alternative come Gab o Reddit.

Dati tratti da sondaggi. In aggiunta a campioni rappresentativi della popolazione o a campioni non rappresentativi di utenti Internet è stato chiesto con che frequenza abbiano osservato, vissuto direttamente o diffuso in prima persona discorsi d'odio. Stando ai risultati disponibili, è probabile che la maggior parte delle persone si sia già imbattuta in discorsi d'odio in rete. Su 2592 giovani adulti di sei Paesi, il 70 per cento ha dichiarato di aver già notato la presenza di discorsi d'odio in Internet⁵⁶, con quote variabili a seconda dei Paesi (più elevate in Finlandia e Spagna, più basse in Francia e in Gran Bretagna). I discorsi d'odio osservati più di frequente sono quelli a sfondo razzista (fig. 2). Il fatto che quattro interpellati su cinque siano stati confrontati con questi contenuti per caso, senza averli cercati, dimostra la facilità con cui si può arrivare al consumo passivo di discorsi d'odio in rete.

	Total sample	Finland	France	Poland	Spain	U.K.	U.S.A.
Ethnicity or Race	964	90	143	147	102	201	281
	37.2%	36.4%	29.4%	37.8%	26.4%	39.0%	49.6%
Nationality or Immigrant Status	853	125	138	149	122	155	164
	32.9%	50.6%	28.3%	38.3%	31.5%	30.1%	28.9%
Sexual Orientation	889	93	146	153	163	133	201
	34.3%	37.7%	30.0%	39.3%	42.1%	25.8%	35.4%
Religious Conviction and Belief	640	81	123	118	52	124	142
	24.7%	32.8%	25.3%	30.3%	13.4%	24.1%	25.0%
Political Views	615	59	84	107	117	87	161
	23.7%	23.9%	17.3%	27.5%	30.2%	16.9%	28.4
Sex/Gender/Gender Identity	686	57	99	75	165	112	178
	26.5%	23.1%	20.4%	19.3%	42.6%	21.7%	31.4%
Disability Status	223	16	32	34	46	33	62
	8.6%	6.5%	6.6%	8.7%	11.9%	6.4%	10.9%
Appearance	470	49	91	83	85	64	98
	18.1%	19.8%	18.7%	21.3%	22.0%	12.4%	17.3%
Sample Size	2,592	247	487	389	387	515	567

Figura 2: domanda a tutte le persone che hanno osservato discorsi d'odio: quali erano le caratteristiche sociali prese di mira?

Nella vicina Germania sono già stati effettuati diversi sondaggi sui discorsi d'odio. Nella più vasta inchiesta rappresentativa finora promossa su scala nazionale, condotta da Geschke et al.⁵⁷, il 40 per cento dei 7349 internauti interpellati ha dichiarato di essersi già imbattuto in discorsi d'odio in rete. I contenuti erano rivolti in massima parte contro migranti, politici, musulmani e rifugiati e, in misura leggermente minore, contro ebrei, sinti e rom. In linea di massima quindi anche in Internet gli attacchi prendono di mira gruppi storicamente sfavoriti. Un sondaggio su scala un po' più ridotta⁵⁸ evidenzia una crescita significativa dei discorsi d'odio osservati negli ultimi anni (dal 65 % nel 2016 al 78 % nel 2018).

Una quota più ridotta di persone afferma di essere stata personalmente vittima di discorsi d'odio. Nel sondaggio realizzato in Germania da Geschke et al., si tratta dell'8 per cento del campione; per le persone con un passato migratorio la possibilità di subire simili attacchi è doppia rispetto a quelle senza passato migratorio. Le percentuali al riguardo variano a seconda dei Paesi: studi effettuati in Turchia e negli Stati Uniti attestano valori più elevati compresi tra il 25 e il 40 per cento⁵⁹.

4.2 Situazione in Svizzera

Siccome per la Svizzera non vi sono studi sistematici sulla prevalenza dei discorsi d'odio e del razzismo in Internet, è sconsigliabile generalizzare le poche, seppur interessanti, evidenze attualmente disponibili. Il quadro provvisorio della situazione fa pensare che siano numerosi i contenuti di questo tipo caricati in rete, ma che pochi vengano segnalati e denunciati. Secondo il rapporto «Episodi di razzismo trattati nell'attività di consulenza. Gennaio – Dicembre 2019»⁶⁰, su 352 episodi di

⁵⁶ Reichelmann et al. 2020: pag. 4.

⁵⁷ Geschke et al. 2019: pag. 19.

⁵⁸ Landesanstalt für Medien NRW 2018: pag. 2.

⁵⁹ Duggan 2017; Celik 2019: pag. 1456.

⁶⁰ Rete di consulenza per le vittime del razzismo 2020: pag. 11.

discriminazione razziale trattati dai consultori soltanto 23 (2 % in più rispetto all'anno precedente) hanno avuto luogo in Internet. Anche i dati provvisori per il 2019 relativi ai casi giuridici concernenti commenti razzisti sulle piattaforme dei social media (6 casi), nei blog e nelle rubriche dei commenti dei media tradizionali (6 casi) sono modesti⁶¹.

Se si considera che servizi rilevanti giudicano le istigazioni all'odio in rete come un «problema cronico», l'esiguità di queste cifre è sorprendente⁶². Anche alcune testimonianze suggeriscono un numero significativo di commenti intrisi d'odio nella cybersfera svizzera. Alcuni membri del gruppo Facebook autoorganizzato «Meldezentrale für Eidgenossen» («Centrale di segnalazione per Confederati»), che denunciano sistematicamente contenuti di questo tipo, affermano di essere riusciti a ottenere la cancellazione, nel solo anno 2017/2018, di ben 7000 commenti del genere⁶³. Inoltre, su 523 episodi di antisemitismo rilevati complessivamente nell'«Antisemitismusbericht für die Deutschschweiz 2019» («Rapporto sull'antisemitismo nella Svizzera tedesca»)⁶⁴, ben il 96 per cento del totale si è verificato online, per la maggior parte sui social media (cfr. fig. 3).

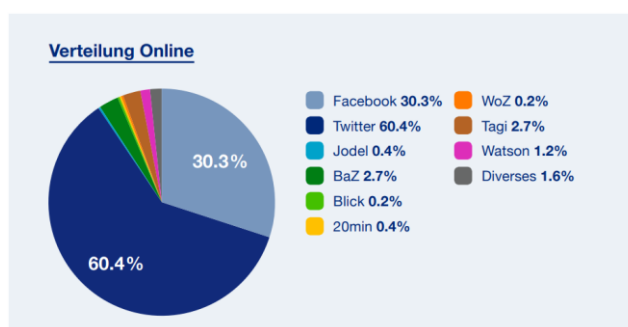


Figura 3: ripartizione dei casi online tra social media e media tradizionali (GRA e SIG 2019).

In questo contesto vanno considerati anche studi su altri tipi di attacco diretto in rete, in quanto anch'essi possono avere una connotazione razzista. In un sondaggio condotto tra la popolazione nel 2019⁶⁵, che ha coinvolto oltre 2000 persone, circa l'8 per cento ha dichiarato di essere stato dileggiato, offeso, insultato, minacciato oppure di essere stato vittima di maldicenze o della diffusione di informazioni private; il 7 per cento ha subito molestie sessuali. L'indagine non ha tuttavia rilevato differenze dovute alla presenza o meno di un retroterra migratorio. Sembra inoltre che siano soprattutto i personaggi pubblici o di una certa visibilità a essere esposti a simili attacchi: nel 2017, oltre la metà dei 637 giornalisti svizzeri intervistati⁶⁶ ha asserito di aver subito ingiurie, minacce od oltraggi di altro genere dal proprio pubblico (9 su 10 tramite i media digitali). In questo caso invece il 7 per cento degli interpellati ha dichiarato che gli attacchi prendevano di mira il loro retroterra migratorio o la loro provenienza.

4.3 Bambini e adolescenti: un gruppo particolare

Merita un'attenzione particolare la coorte d'età più giovane, che si muove quotidianamente e con dimestichezza in Internet. I social media sono parte integrante della vita dei ragazzi, il che li rende più visibili, ma anche più attaccabili. Ricerche condotte in diversi Paesi dimostrano che sono proprio loro a rilevare più spesso discorsi d'odio (razzisti) o a esserne autori o vittime. In un sondaggio condotto in Germania, il 96 per cento degli interpellati tra i 14 e i 24 anni ha dichiarato di «aver visto di persona

⁶¹ Commissione federale contro il razzismo. Dati provvisori: maggio 2020.

⁶² Humanrights.ch, 23 agosto 2017: [Hassaufrufe im Internet - Schweizer Fälle und Politik der Newsportale](#).

⁶³ Germann, WOZ, 20 settembre 2018: [7000 gesperrte Kommentare](#).

⁶⁴ Fondazione contro il razzismo e l'antisemitismo (GRA) e Federazione svizzera delle comunità israelite (FSCI) 2019. Questi dati includono casi segnalati da osservatori o dalle stesse vittime di attacchi, episodi menzionati nei media e ricerche proprie condotte nei social media dalle organizzazioni autrici del rapporto.

⁶⁵ Baier 2019: pag. 39.

⁶⁶ Stahel e Schoen 2019: pag. 11.

commenti d'odio in Internet» (rispetto all'85 % delle persone tra i 25 e i 44 anni, al 75 % di quelle tra i 45 e i 59 anni e al 60 % degli ultrasessantenni)⁶⁷. In un'altra indagine, focalizzata sui giovani tra i 15 e i 30 anni di quattro Paesi, un terzo degli intervistati tedeschi ha dichiarato di essersi imbattuto almeno una volta in discorsi d'odio online negli ultimi tre mesi, una quota che sale al 39 per cento per la Gran Bretagna, al 48 per cento per la Finlandia e al 53 per cento per gli Stati Uniti⁶⁸. Il numero di giovani vittima di discorsi d'odio è più basso, i giovani autori di discorsi d'odio sono il gruppo più piccolo: in un sondaggio effettuato in Francia tra quasi 2000 adolescenti, il 7 per cento del campione ha affermato di aver già pubblicato o condiviso notizie, commenti o immagini «umilianti o intrisi d'odio» rivolti contro una persona o un gruppo⁶⁹. L'11 per cento ha invece dichiarato di esserne stato vittima. Negli Stati Uniti, sono state rilevate cifre più elevate: gli autori sono il 20 per cento e le vittime quasi un terzo del totale⁷⁰.

In Svizzera sono disponibili soltanto dati riguardo all'osservazione di discorsi d'odio in rete. In confronto ad altri Paesi, le percentuali sono relativamente modeste: nel 2019, tra gli internauti svizzeri dai 12 ai 16 anni interpellati nell'ambito dello studio «EU Kids Online», uno su otto ha affermato di trovare discorsi d'odio in rete almeno una volta al mese. Per uno su cinque, la frequenza è inferiore, mentre due terzi non si sono mai imbattuti in simili contenuti⁷¹. Lo stesso studio dimostra che l'esposizione a tali contenuti cresce con l'aumentare dell'età: tra gli adolescenti di 15-16 anni, la percentuale di coloro che hanno osservato discorsi d'odio è doppia (21 %) rispetto al gruppo dei dodici-quattordicenni (11 per cento).

⁶⁷ Landesanstalt für Medien NRW 2018: pag. 2.

⁶⁸ Hawdon et al. 2017 : pag. 260.

⁶⁹ Blaya e Audrin 2019 : pag. 6.

⁷⁰ Costello e Hawdon 2018: pag. 58.

⁷¹ Smahel et al. 2020: pag. 66 e 67.

5 AMBIENTE DIGITALE DEI DISCORSI D'ODIO RAZZISTI

In Internet, le idee razziste si stanno trasformando in un sistema sempre più complesso e stratificato di piattaforme e applicazioni digitali. Si propagano all'interno di reti di utenti e tra di esse⁷² sotto molteplici forme, dalle esplicite minacce di distruzione a pregiudizi stereotipati ridiffusi. Il presente capitolo esplora l'organizzazione dei discorsi d'odio razzisti online e le piattaforme su cui sono presenti. Illustra inoltre i fattori individuali e sociali che ne favoriscono la propagazione e l'importante ruolo delle condizioni comunicative e dell'architettura delle piattaforme.

5.1 Grado di organizzazione degli autori

I diffusori di discorsi d'odio razzisti online possono essere suddivisi in base al grado di organizzazione: gruppi d'odio strutturati gerarchicamente, reti «fluide» o anche singoli utenti. Tutte entità dai confini permeabili che permettono ai singoli di passare agevolmente dall'una all'altra. Il grado di organizzazione ne influenza tuttavia sensibilmente il tipo di piattaforma, gli obiettivi e il modo di comunicare⁷³.

I più datati e noti sono i cosiddetti *gruppi d'odio (hate groups)*⁷⁴. Si tratta di gruppi generalmente organizzati in modo gerarchico, che esistevano già prima di Internet. Dagli anni 1990 sono anche online e si esprimono per lo più in modo esplicitamente razzista. Con l'ulteriore sviluppo digitale, l'avvento dei social media e la crescente pressione a sottrarsi alla stigmatizzazione sociale dei contenuti apertamente di estrema destra, negli ultimi anni si sono trasformati in entità meno organizzate e meno esplicite.

Queste *reti digitali*, più informali, ospitano utenti che si muovono su una o più piattaforme, sono collegati più o meno strettamente tra di loro⁷⁵ e interagiscono occasionalmente, per esempio in caso di attacchi coordinati. Si tratta di reti fluide e ambivalenti, prive di chiari centri organizzativi⁷⁶. Ne è un esempio il movimento digitale alt-right⁷⁷, nato negli Stati Uniti e noto al pubblico al più tardi da quando ha sostenuto attivamente la candidatura di Donald Trump alla presidenza. Alt-right non ha un capo, ma soltanto opinion leader che si alternano dinamicamente sui social media, sono chiamati anche «Intellectual Dark Web» e possono essere influencer su Twitter, blogger su YouTube o podcaster⁷⁸. Aniché impartire ordini, esercitano un'influenza emotiva. Contestano i «media politicamente corretti» e l'opinione pubblica con tutti quelli che l'accettano (i cosiddetti *normy*⁷⁹). Si considerano vittime, una comunità in pericolo che può difendersi soltanto con un'alternativa e soluzioni radicali. Tra i loro seguaci, i confini tra attivisti organizzati e semplici sostenitori si confondono. I loro troll sono tuttavia molto attivi: mirano ad attirare la massima attenzione disturbando dibattiti e non esitano a diffamare persone e organizzazioni non gradite⁸⁰.

In alcune forme di organizzazione, tuttavia, gerarchia e struttura a rete convivono. Un esempio di forma ibrida era il «movimento identitario» di estrema destra tedesco Reconquista Germanica⁸¹: gli utenti di questo forum online non pubblico avevano «inondato» pagine dei commenti e profili social con i loro contenuti, seguendo rigorose catene di comando⁸².

La transizione verso *singoli utenti* poco organizzati o non organizzati è fluida. Questi ultimi diffondono contenuti razzisti in modo relativamente isolato, volutamente o inconsapevolmente, per scarsa

⁷² Guhl et al. 2020: pag. 5.

⁷³ Bliuc et al. 2018: pag. 76.

⁷⁴ Winter 2019: pag. 54.

⁷⁵ Fielitz e Marcks 2019: pag. 16.

⁷⁶ Fielitz e Marcks 2019: pag. 1.

⁷⁷ Descritto come «a set of far-right ideologies, groups and individuals whose core belief is that «white identity» is under attack by multicultural forces using «political correctness» and «social justice» to undermine white people and «their» civilization». Southern Poverty Law Center 2020. www.splcenter.org > What We Do > Fighting Hate > Extremist Files > Ideologies > See All Ideologies > [Alt-Right](#).

⁷⁸ Winter 2019.

⁷⁹ [Qui](#) è disponibile un'autodescrizione dettagliata del movimento: Anglin, Daily Stormer, 31 agosto 2016.

⁸⁰ Marwick e Lewis 2017: pag. 4.

⁸¹ Nel novembre del 2019 il gruppo ha dichiarato il proprio scioglimento su [Youtube](#): Youtube > Reconquista Germanica meldet sich ab.

⁸² Fielitz e Marcks 2019: pag. 13.

competenza mediale o intenzione strategica, impulso dettato dalla situazione o radicata convinzione ideologica – in ogni caso non a nome di un gruppo. Sono spesso presenti sulle piattaforme social più note come Facebook o Twitter.

5.2 Fattori d'influenza individuali e sociali

Chi diffonde razzismo in Internet e perché? Indagini sui diffusori di discorsi d'odio in generale possono fornire indicazioni. Da un lato sono rilevanti *caratteristiche individuali*. Gli internauti che disturbano regolarmente dibattiti online e provocano («trollano») hanno tendenze sadiche e psicopatiche (subpatologiche) più marcate rispetto agli utenti che non trollano. Sono più impulsivi e hanno meno scrupoli, cercano maggiormente il brivido e provano meno empatia⁸³. Hanno una loro influenza anche le esperienze fatte in Internet. Non è chiaro se gli autori di discorsi d'odio trascorrono più tempo online. In ogni caso si muovono più spesso su piattaforme e in comunità in cui sono propagati discorsi d'odio⁸⁴. È dimostrata la significativa compresenza di autore e vittima nella stessa persona: la vittima diventa autore e/o viceversa⁸⁵. L'influsso dei videogiochi sui discorsi d'odio non è ancora stato appurato dalla scienza. Gli effetti rilevati sono contraddittori, anche se metanalisi (ossia studi che sintetizzano studi precedenti) riscontrano tendenzialmente da effetti positivi minimi a nessun effetto⁸⁶. Gli effetti positivi si spiegano in particolare con due meccanismi fondamentali⁸⁷. Le persone che presentano già una tendenza all'aggressione sono maggiormente attratte da determinati giochi (se non ci fossero i giochi, avrebbero sfogato la loro aggressività altrove). Al tempo stesso, chi gioca agli sparatutto (*shooter*) potrebbe assimilare comportamenti violenti osservando le figure violente nel gioco, inserire questi comportamenti nel repertorio comportamentale analogico e adottarli nella vita reale. La diffusione di discorsi d'odio è infine influenzata anche dagli obiettivi e dagli atteggiamenti personali degli utenti. I moventi sono a largo spettro: in interviste e indagini, gli autori rivelano di cercare, attraverso i discorsi d'odio, attenzione, riconoscimento, vendetta, influenza, potere e intrattenimento. Vogliono fare proseliti, eliminare le ingiustizie percepite, difendere il proprio gruppo sociale e veder confermate le proprie convinzioni⁸⁸. Questa eterogeneità rispecchia l'ampia «tipologia» degli autori: dal «cittadino arrabbiato» frustrato, che si permette una «deriva d'odio» isolata in un acceso dibattito online, all'influencer d'odio, che utilizza i discorsi d'odio strategicamente e a lungo termine quale strumento politico davanti a un ampio pubblico online. Anche i pregiudizi e la diffidenza sono considerati terreno fertile per i discorsi d'odio. Da un'indagine condotta in Francia è emerso che i giovani tendono maggiormente a diffondere discorsi d'odio se vedono di buon occhio violenza e razzismo e diffidano della scuola e della politica⁸⁹.

Le conoscenze sulle *caratteristiche sociali* degli autori e la loro integrazione nella società sono scarse. In ogni caso, la quota di uomini è nettamente superiore⁹⁰. Anche i giovani sembrano relativamente sovrarappresentati (cfr. cap. 4.3). Mancano tuttavia confronti sistematici con coorti di età superiore. È presumibile che generazioni distinte «odino» in modo specifico a seconda dei temi e delle vittime: per esempio, la maggior parte dei commentatori denunciati grazie all'associazione svizzera #Netzcourage sono persone anziane, per lo più personaggi pubblici di sesso femminile, come donne attive in politica⁹¹. Il contesto sociale sembra avere un chiaro influsso: i giovani autori appartengono spesso a un gruppo giovanile deviante (p. es. criminale), per il resto hanno relativamente poche relazioni offline e molti contatti online, per esempio in comunità online⁹². Per ora non sono classificabili chiaramente in «classi sociali»: i pochi dati disponibili non sembrano evidenziare effetti legati alla formazione e di conseguenza all'attività professionale e al reddito⁹³, il che è sorprendente perché in generale queste caratteristiche sociostrutturali sono uno dei fattori che spiegano i reati. Gli autori degli studi ipotizzano tuttavia che in

⁸³ Buckels et al. 2014.

⁸⁴ Costello e Hawdon 2018: pag. 58.

⁸⁵ Blaya e Audrin 2019: pag. 11.

⁸⁶ Elson e Ferguson 2014.

⁸⁷ Elson e Ferguson 2014: pag. 3.

⁸⁸ P. es. Craker e March 2016: pag. 83; Erjavec e Kovačič 2012: pag. 912; Guhl et al. 2020: pag. 47.

⁸⁹ Blaya e Audrin 2019.

⁹⁰ P. es. Blaya e Audrin 2019: pag. 3; Costello e Hawdon 2018: pag. 58.

⁹¹ Intervista con Jolanda Spiess-Hegglin, marzo 2020.

⁹² P. es. Blaya e Audrin 2019: pag. 11; Ribeiro et al. 2018.

⁹³ P. es. Costello e Hawdon 2018: pag. 58; Lowry et al. 2016: pag. 978.

futuro, sulla scia della crescente educazione ai media, si potrebbero osservare più spesso effetti sociostrutturali⁹⁴.

Considerando che il contesto reale e lo spazio virtuale online si fondono sempre più, non stupisce che anche gli *sviluppi sociali* si riflettano in discorsi d'odio razzisti online⁹⁵. Un sovvertimento delle strutture di potere sociali, economiche, politiche o culturali, come avviene nel caso della migrazione, dei valori postmaterialisti, della parità di diritti o dell'uguaglianza (o disuguaglianza), può alimentare conflitti sociali. Gruppi privilegiati possono perdere il loro statuto o temere di perderlo. Nell'intento di scongiurare la minaccia, possono ricorrere ai discorsi d'odio, analogici o digitali, per ripristinare il confine chiaro tra coloro che considerano la norma e quindi superiori e coloro che considerano diversi e quindi inferiori e per «ricordare» a questi ultimi qual è il loro «posto»⁹⁶. Questi conflitti si ripercuotono anche sul giornalismo: quando scrivono per esempio su tematiche legate all'identità, come la religione o la parità di genere, i giornalisti svizzeri sono criticati dal pubblico con particolare frequenza⁹⁷. Inoltre le società polarizzate sono considerate terreno fertile per i discorsi d'odio⁹⁸ e Internet costituisce un'estensione del campo di battaglia per la conquista di risorse, potere e chance di sopravvivenza. I troll dell'ex Reconquista Germanica associati all'Alternativa per la Germania (Alternative für Deutschland, AfD) sono soltanto un esempio dell'impiego strategico dei discorsi d'odio online per scopi politici.

Queste strutture sociali possono riflettersi in *singoli eventi*, che a loro volta scatenano discorsi d'odio razzisti online. La scena pubblica dei social media consente di produrre e consumare discorsi d'odio razzisti in massa, in particolare in caso di eventi ben precisi. Analisi estese di commenti online evidenziano che nello spazio digitale gli attentati contro persone amplificano i discorsi d'odio contro i gruppi ritenuti colpevoli. Ciò si spiega tra l'altro con il fatto che gli attentati suscitano sensazioni di minaccia e insicurezza, che a loro volta alimentano l'ostilità nei confronti degli «stranieri». L'appartenenza sociale delle vittime sembra tuttavia un fattore determinante: su Twitter, dopo gli attentati islamisti – come quello durante un concerto alla Manchester Arena nel 2017, con 23 morti – il numero di discorsi d'odio è cresciuto (come anche, ed è interessante notarlo, le contronarrazioni che si opponevano alla condanna indiscriminata della comunità musulmana). Nel caso degli attentati motivati dall'odio verso i musulmani – come l'uccisione di due indiani scambiati erroneamente per iraniani in Kansas, negli Stati Uniti, nel 2017 – non è invece stato rilevato alcun incremento dei discorsi d'odio⁹⁹. I discorsi d'odio online possono anche aumentare durante altri eventi politici, come le elezioni¹⁰⁰, pur riappiattendosi relativamente in fretta¹⁰¹.

Questi studi – tutti realizzati al di fuori della Svizzera – danno una prima impressione della complessità dei motivi che spiegano i discorsi d'odio razzisti online. Per poter orientare gli interventi sui gruppi target, in Svizzera occorrerebbe analizzare in particolare la correlazione con fattori sociali (il reddito, il livello d'istruzione, la posizione sociale ecc.) ed eventi sociali (p. es. le elezioni).

⁹⁴ Il comportamento morale è infatti spesso appreso attraverso la formazione. Oggi l'educazione ai media digitali, che insegna il comportamento empatico-morale in Internet e la valutazione dell'impatto del proprio comportamento, è spesso parte del programma scolastico. Le persone che in futuro beneficerebbero di meno formazione in questo campo saranno quindi meno competenti e di conseguenza più vulnerabili di fronte alla diffusione di discorsi d'odio online (benché anche l'educazione ai media non costituisca un «vaccino contro i discorsi d'odio»). Il fatto che tra gli adulti attuali non si rilevi praticamente alcun effetto potrebbe essere dovuto al fatto che quasi tutta questa generazione non ha ricevuto alcuna educazione ai media digitali nel corso della sua formazione; da solo, il livello di formazione difficilmente può influenzare la predisposizione alla diffusione di discorsi d'odio.

⁹⁵ Cfr. in proposito anche la dipendenza dal Paese nei contenuti concreti dei discorsi d'odio razzisti in sCAN 2018a.

⁹⁶ Quent 2018: pag. 49 segg.

⁹⁷ Stahel, NZZ, 7 maggio 2019: [Gehässige Leserreaktionen können die Qualität der Berichterstattung auch erhöhen.](#)

⁹⁸ Intervista con Jonathan Birdwell (Institute of Strategic Dialogue), aprile 2020.

⁹⁹ Olteanu et al. 2018: pag. 228.

¹⁰⁰ Lim 2017.

¹⁰¹ Siegel et al. 2019: pag. 32.

5.3 Principali piattaforme

Esistono molti spazi online che consentono di diffondere idee razziste. Vi sono tuttavia alcune piattaforme particolarmente importanti, o perché ampiamente utilizzate nella società (p. es. Facebook) o in quanto piattaforme alternative. Queste ultime sono state create appositamente per scopi estremistici o tollerano i discorsi d'odio in virtù di una generosa interpretazione del concetto di libertà di espressione oppure sono state nel frattempo destinate a scopi diversi da quelli, apolitici, originari (p. es. videogiochi)¹⁰². Nel seguito sono presentate separatamente le principali piattaforme.

In generale, tuttavia, nella pratica i diffusori di discorsi d'odio circolano su più piattaforme e si collegano, interconnettono e ispirano al di là delle piattaforme. Attori che operano strategicamente, come il gruppo tedesco antislamico PEGIDA, il Movimento identitario (Identitäre Bewegung, IB) etno-nazionalista, con sede in Europa, o la rete internazionale terroristica di destra Divisione armi nucleari (Atomwaffen Division, AWD), si servono di Internet per diffondere transnazionalmente propaganda e disinformazione, osannare gli attentatori, ordire attacchi coordinati contro politici e avviare campagne di memi (cfr. cap. 5.4.2)¹⁰³. Questi gruppi sono connessi a livello nazionale e internazionale e l'interconnessione è favorita in caso di lingua comune, come negli Stati Uniti e nel Regno Unito (cfr. fig. 4). Attingono informazioni da un ampio panorama mediatico populista di destra, che scredita i mezzi d'informazione considerati *mainstream* e legittima i discorsi d'odio¹⁰⁴.

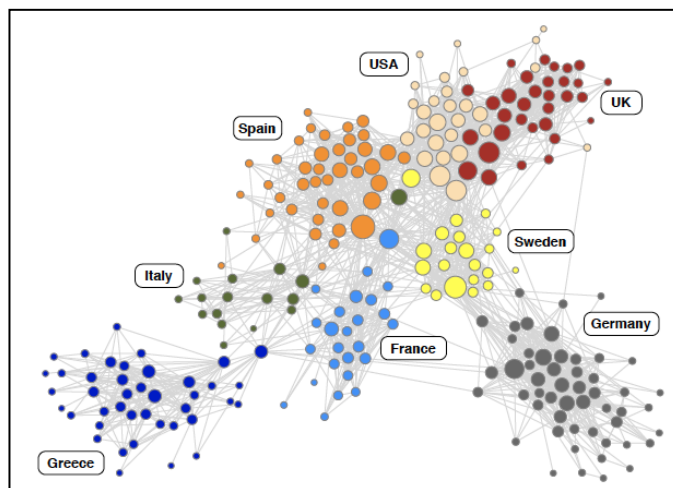


Fig. 4: interconnessione di profili di gruppi di estrema destra di otto Paesi sulla piattaforma Twitter (è tracciata una linea di collegamento quando il profilo Twitter di un gruppo menziona quello di un altro nei suoi tweet o ne condivide i tweet).

Siti web e blog

I siti web e i blog sono importanti centri d'informazione per le idee razziste, che in genere sono comunicate apertamente, per esempio su [The Daily Stormer](#) e [Vanguard News Network](#), o, in forma attenuata, su siti web populistici di destra, come [Breitbart](#) o il tedesco [PI News](#). Siti web paragonabili in Svizzera, che rasentano o difendono il razzismo o il populismo, sono [Hammerschweizer.ch](#) (Svizzera tedesca) o [LesObservateurs.ch](#) (Svizzera romanda). In Ticino benché siano constatate diffuse tendenze xenofobe (il che può essere legato alla frontiera con l'Italia)¹⁰⁵, i simpatizzanti sembrano informarsi

¹⁰² Guhl et al. 2020: pag. 5.

¹⁰³ Guhl et al. 2020: pag. 7.

¹⁰⁴ Guhl et al. 2020: pag. 7.

¹⁰⁵ Haymoz et al. 2019: pag. 11.

soprattutto su siti web italiani¹⁰⁶, per esempio quelli del movimento di estrema destra [Casapound Italia](#) (che gestisce anche un canale Twitter e Telegram) o della [Comunità militante dei Dodici Raggi](#) (Do.Ra.), una delle maggiori e meglio organizzate comunità nazionalsocialiste italiane. Negli ultimi 20 anni, questi siti web hanno registrato una forte crescita, in particolare in seguito alla diffusione di massa di Internet¹⁰⁷. Il loro reale numero è tuttavia difficile da stabilire, dal momento che chiudono o cambiano nome spesso¹⁰⁸. Per acquisire visibilità, interagiscono intensamente linkandosi e citandosi a vicenda. Si rivolgono a gruppi target di ampiezza variabile. Attribuiscono per esempio importanza alle donne, sottolineandone il ruolo nella costituzione della famiglia, non importa se ariana o jihadista¹⁰⁹.

Forum

I forum sono invece più partecipativi: gli utenti possono infatti postare i loro contributi razzisti. Il più vecchio è il forum di estrema destra [Stormfront](#), messo online negli Stati Uniti nel 1995, dove oltre ad ascoltare la radio, chattare e leggere blog si può principalmente postare contenuti e dialogare. I contenuti vertono sulla «superiorità bianca» e comprendono cosiddetti «thread», ovvero discussioni separate, per esempio su «storia e revisionismo». Forum più recenti sono [Gab](#), [Reddit](#) e [8chan](#), che hanno seguaci in tutto il mondo¹¹⁰. Qui il razzismo si concentra spesso in singoli thread, per esempio il popolare thread pro-Trump [the donald](#) (su Reddit) o il «politicamente scorretto /pol/» (su [4chan](#)). Il razzismo può tuttavia anche accompagnare dibattiti, come quello sui test genetici per dimostrare la propria «purezza genetica»¹¹¹. Questi forum sono prevalentemente anonimi, spesso umoristici e meno controllati delle piattaforme social affermate. Qui razzismo e movimenti misogini come quello degli *incel*¹¹² vanno a braccetto. Molti di questi forum si autodefiniscono pagine di estrema destra, attraendo così persone che ne condividono le convinzioni e si aspettano di trovarvi «porti sicuri» in cui poter esprimersi senza temere sanzioni. Una volta che circolano su questi forum, gli utenti s'imbattono velocemente in numerosi link verso siti estremisti, il che ne favorisce l'ulteriore radicalizzazione¹¹³.

Per finire possono essere sfruttati per diffondere razzismo anche forum che non sono stati creati per propagare discorsi d'odio, bensì, per esempio, per parlare di sport¹¹⁴.

Piattaforme social affermate

Il social network Facebook, la piattaforma audiovisiva YouTube o il microblog Twitter: è su queste piattaforme affermate che naviga la maggior parte della popolazione svizzera. Qui il cittadino medio s'imbatte, in genere casualmente, anche in discorsi d'odio razzisti¹¹⁵. Con le loro strutture interconnesse, queste piattaforme sono infatti predestinate a diffondere contenuti su scala globale, il che le rende particolarmente attrattive per gli attori razzisti in cerca di visibilità, che propagano, per esempio su Twitter, contenuti che poi sono ripresi da giornalisti¹¹⁶. È così che i media hanno riferito della campagna su Twitter «#120db» del Movimento identitario radicale di destra: quest'ultimo ha sfruttato la narrativa dei diritti femminili per protestare contro la cosiddetta «violenza dei migranti» contro le donne e inveire contro gli immigrati di sesso maschile¹¹⁷. Altri autori caricano su YouTube video musicali razzisti che gli

¹⁰⁶ Informazione del Dipartimento delle istituzioni (DI), Piattaforma di prevenzione della radicalizzazione e dell'estremismo violento, Bellinzona.

¹⁰⁷ Winter 2019: pag. 40; Perry e Olsson 2009: pag. 188.

¹⁰⁸ Perry e Olsson 2009: p. 188.

¹⁰⁹ Musial 2017.

¹¹⁰ Hine et al. 2016.

¹¹¹ Mittos et al. 2019.

¹¹² «Incel» è l'acronimo di «Involuntary Celibate», ossia «celibato involontario». I membri di questa comunità online, per lo più uomini, sostengono di essere stati privati del diritto a rapporti sessuali. Sono seguaci dell'ideologia della virilità egemonica e manifestano apertamente odio nei confronti delle persone sessualmente attive, in particolare delle donne. Cfr. Pfeiffer, Belltower News, 25 marzo 2020: [Hate Speech in der Incel-Szene](#).

¹¹³ Marwick e Lewis 2017.

¹¹⁴ Cleland 2013.

¹¹⁵ Celik 2019; Reichelmann et al. 2020: pag. 6.

¹¹⁶ Marwick e Lewis 2017: pag. 26.

¹¹⁷ Jetzt, 20 febbraio 2018: [Rechtsextreme «Feministinnen» stören die Berlinale](#).

utenti passano in rassegna cliccando e postandovi sotto commenti razzisti. Su queste piattaforme sono presenti anche influencer che fungono da ponte tra *mainstream* ed estremismo. Brillano nell'accalappiare l'attenzione rafforzando la visibilità delle loro idee mediante l'uso strategico delle piattaforme¹¹⁸. Stefan Molyneux, un commentatore attivo su YouTube, per esempio, diffonde con successo tesi razziste in veste scientifica. I video sul suo canale YouTube, che portano titoli come «Human Biodiversity and Criminality», sono già stati visualizzati quasi 300 milioni di volte¹¹⁹.

Un'indagine internazionale¹²⁰ rivela che la piattaforma su cui gli utenti vedono la maggior quantità di discorsi d'odio è Facebook – il che potrebbe essere correlato alla sua posizione di social network più gettonato (cfr. fig. 5). Gruppi in parte pubblici e in parte privati inveiscono ad esempio contro i rifugiati e i musulmani: in Germania si tratta dei gruppi *Nein zum Heim* e in Svizzera di gruppi di stampo nazionalista come il gruppo Facebook *New Swiss Journal*, dove anche a chi si adopera a sostegno dei rifugiati è attribuita l'etichetta di «traditore della patria».

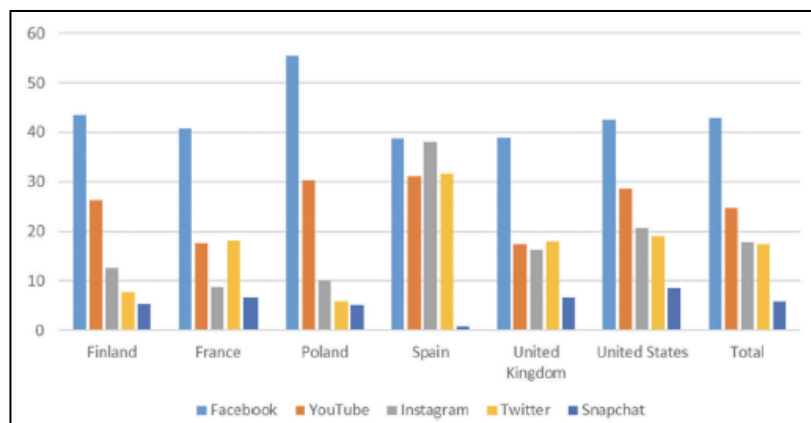


Fig. 5: quota di persone in età compresa tra i 18 e i 25 anni che riferisce di aver visto discorsi d'odio sulle varie piattaforme Internet nei tre mesi precedenti l'indagine in sei Paesi (N= 2592).

Videogiochi

Anche il carattere competitivo dei videogiochi (*online games*), la forte identificazione di certi giocatori e una cultura del gioco specifica possono favorire i discorsi d'odio¹²¹. I videogiochi sono ormai molto diffusi e diversificati per quanto riguarda i terminali, i generi, i giocatori e le modalità di gioco (single/multiplayer). Benché la gamma di interazioni tra i giocatori sia ampia, secondo gli esperti da sempre i videogiochi rappresentano anche uno spazio dinamico per il razzismo¹²². Gruppi d'odio hanno iniziato molto presto a produrre giochi di odio (*hate games*), ossia versioni d'odio di popolari giochi online destinate a intrattenere e attrarre nuovi membri. I giochi di odio riproducono stereotipi e pregiudizi, confondono la realtà e la fantasia e presentano in genere «soluzioni» violente. Vi è poi la grande comunità del gaming, connessa su scala mondiale¹²³. Alcuni ricercatori equiparano questo gruppo molto ampio di giocatori a una comunità culturale, in cui domina la conoscenza di simboli, significati e pratiche comuni. Una diffusa pratica escludente è quella del *trash talk*, ossia degli insulti verbali – anche razzisti – tra i giocatori¹²⁴. Questa pratica è emersa in modo particolarmente evidente nella campagna #Gamergate. Nell'agosto del 2014, diverse donne che avevano criticato la misoginia e il razzismo nella cultura del gaming sono state coperte di odio online su 4chan, Reddit e Twitter.

¹¹⁸ Marwick e Lewis 2017: pag. 1.

¹¹⁹ Southern Poverty Law Center, 19 aprile 2018: [McInnes, Molyneux, and 4chan: Investigating pathways to the alt-right](#).

¹²⁰ Reichelmann et al. 2020.

¹²¹ Breuer 2017: pag. 107; Banaszczuk 2019.

¹²² P. es. Ortiz 2019.

¹²³ Banaszczuk 2019.

¹²⁴ Ortiz 2019.

E-mail e chat

Sono possibili attacchi razzisti anche su canali privati. Le e-mail consentono di attaccare e intimidire direttamente individui specifici. Questi attacchi non sono rari. In Svizzera, nel 2017 l'80 per cento dei giornalisti che avevano riferito di essere stati bersaglio di attacchi (razzisti e/o non razzisti) erano stati attaccati attraverso canali online non pubblici¹²⁵. Tra questi rientrano anche le chat. Hanno destato scalpore casi in cui «chat di classe» veicolavano memi¹²⁶ razzisti¹²⁷ e il gruppo chat WhatsApp FC NSDAP¹²⁸, in cui adolescenti condividevano «barzellette su ebrei, campi di concentramento e saluti nazisti». Questa comunicazione non pubblica o parzialmente pubblica su WhatsApp (limite della chat fissato a 256 membri) o Telegram (fino a 200 000 membri) assumerà presumibilmente sempre più rilievo per il razzismo: essendo praticamente impossibile controllarla dall'esterno, infatti, attrae coloro che sono banditi dalle piattaforme social affermate¹²⁹.

5.4 Infrastruttura online che favorisce il razzismo

Il sistema delle piattaforme descritto offre una struttura di opportunità, che favorisce i discorsi d'odio e il razzismo. Le sue condizioni comunicative e architettoniche attenuano gli ostacoli che limitano i discorsi d'odio nel contesto reale¹³⁰. Non è quindi una sorpresa che gli utenti riferiscano di osservare sui social media una maggior presenza di discorsi d'odio che nella vita analogica¹³¹. Il presente capitolo esplora questo ambiente digitale, che non va tuttavia visto come una causa, bensì come un tassello di una spiegazione più complessa (cfr. cap. 5.2).

5.4.1 Condizioni comunicative

Le condizioni ambientali determinano il modo in cui la gente comunica. Pur non offrendo condizioni comunicative sostanzialmente nuove, di qualità esclusiva (l'anonimato per esempio è sempre esistito), lo spazio digitale ne potenzia delle più disparate, che, combinate, finiscono con il promuovere l'odio¹³². Tra di esse figurano le seguenti:

Conflitti provocati dallo scontro tra spazi. In Internet «si scontrano mondi quotidiani divergenti: (...) femministe contro maschilisti e liberali di città contro nazisti di campagna»¹³³. Persone con un retroterra etnico, sociale o culturale differente entrano in contatto da sempre. Con l'interconnessione digitale globale, tuttavia, questi contatti si sono intensificati, mentre i valori e le norme di ogni singolo pubblico non si sono mai dissolti. Per questi pubblici, i contenuti archiviati in eterno in Internet sono raggiungibili con un semplice clic, assieme agli stili di vita più curiosi, opposti e «amoralisti» di altri pubblici. Il materiale di cui indignarsi non manca. Lo scontro sociale e geografico tra questi spazi può così provocare conflitti dinamici e incontrollabili¹³⁴, che si sfogano anche in discorsi d'odio razzisti.

Facilità di accesso. La gente può fornire un contributo allo spazio pubblico digitale in qualsiasi momento, da qualsiasi luogo e praticamente gratuitamente (a patto di disporre dell'accesso a Internet, delle competenze necessarie e di un'infrastruttura digitale dipendente dal Paese). Si tratta di una situazione in contrasto con il passato, dove i custodi dei media tradizionali controllavano la messa a disposizione

¹²⁵ Stahel 2020. Dati non pubblicati, rilevati nel 2017.

¹²⁶ I memi sono unità d'informazione – spesso una combinazione di testo e immagine – tipicamente con una nota umoristica, la cui influenza aumenta con la diffusione. Un esempio è il personaggio dei fumetti Pepe the Frog, inserito dalla Anti-Defamation League nella sua banca dati dei simboli razzisti online. Cfr. Spiegel.de, 28 settembre 2016: [Freundlicher Frosch wird Hasssymbol](#).

¹²⁷ 20min, 30 dicembre 2019: [Primarschüler machen sich über Dunkelhäutige lustig](#).

¹²⁸ 20min, 19 marzo 2019: [«Rechtsextreme Chats gibt es an jeder Schule»](#).

¹²⁹ Guhl et al. 2020 : pag. 18.

¹³⁰ Brown 2018: pag. 301.

¹³¹ Barnidge et al. 2019.

¹³² Brown 2018 : pag. 306.

¹³³ Quent 2018: pag. 48.

¹³⁴ Marwick 2010.

dell'informazione secondo direttive etiche. Le basse barriere di accesso a Internet avevano suscitato la speranza di una democratizzazione del discorso pubblico già 20 anni fa. Una speranza che si è in parte effettivamente realizzata – assieme al rovescio della medaglia. A esprimersi sono infatti anche utenti le cui opinioni si scostano dalle norme comunicative correnti in quanto latrici di contenuti infondati, penalmente perseguibili o di «cattiva qualità». Anche i gruppi d'odio hanno approfittato della nascita di questo «culto dei dilettanti»¹³⁵: nel 2008, negli Stati Uniti un capo del Ku Klux Klan ha dichiarato: «We don't really need the media any more (...) the only thing we need is the Internet» («Non abbiamo veramente più bisogno dei media (...) l'unica cosa di cui abbiamo bisogno è Internet»)¹³⁶.

Pubblicazione immediata. Nell'era predigitale, per raggiungere un ampio pubblico con idee razziste bisognava in genere aspettare che venissero stampati dispendiosi volantini o organizzata una manifestazione. Vi era quindi uno scarto temporale tra l'idea e la pubblicazione: si era costretti a dormirci su e magari il giorno dopo si vedevano le cose con più calma. Oggi invece questo scarto temporale si è nettamente accorciato. Attraverso i social media, il «bar interiore» è reso accessibile al pubblico senza alcun filtro nel giro di pochi secondi. È quindi ipotizzabile che a lungo termine i social media favoriranno in modo particolare l'esternazione spontanea, emotiva e non filtrata di discorsi d'odio razzisti¹³⁷.

Distanza e impressione d'invisibilità. Ovviamente già nell'era predigitale si potevano esprimere idee razziste a distanza e con una buona probabilità di restare invisibili (p. es. in lettere, volantini, grandi manifestazioni). Ma nella comunicazione tramite Internet la distanza da chi ci sta di fronte è connaturata al mezzo. Parlante e pubblico sono separati dalla tecnica e spesso interagiscono asincronicamente. Lo scambio è più asettico e impersonale¹³⁸. La mimica e i gesti, che nel mondo analogico tengono in parte in scacco i comportamenti antisociali, sono fuori gioco. È praticamente impossibile vedere se l'interlocutore è ferito sul piano emotivo e reagire di conseguenza, il che rende le persone meno empatiche. Nelle chat online, le espressioni aggressive diminuiscono se è attivato il contatto visivo (p. es. attraverso la webcam)¹³⁹. Più informazioni (non verbali) sono disponibili sull'interlocutore, più si riduce la superficie su cui proiettare i propri stereotipi e pregiudizi. In Internet si comunica insomma in modo più disinibito e di conseguenza si censurano meno anche le idee razziste¹⁴⁰.

Anonimato/nome reale. Il fatto che l'anonimato rafforzi i discorsi d'odio dipende dal contesto. In linea di principio, gli utenti possono essere tecnicamente anonimi: l'identificazione di fatto della persona è esclusa mediante pseudonimi, dati personali falsi o l'occultamento dell'indirizzo IP. Oppure possono essere socialmente anonimi, nel senso che altri utenti li percepiscono come privi di un'identità individuale, malgrado un eventuale nome reale¹⁴¹. L'anonimato può favorire i discorsi d'odio. È infatti più difficile sanzionare gli autori. Gli utenti socialmente anonimi si orientano inoltre maggiormente al gruppo. Se in un forum la norma prevalente è di natura razzista, anche i neofiti anonimi commenteranno in termini più razzisti¹⁴². Al tempo stesso, i diffusori di discorsi d'odio online possono anche preferire nomi reali, in particolare se cercano riconoscimento e fiducia, sono moralmente convinti delle loro azioni o sono sostenuti da una community. Gli studi in materia rilevano quindi correlazioni differenti tra l'anonimato e l'aggressione o i discorsi d'odio¹⁴³. Nella Corea del Sud, l'obbligo di utilizzare il nome reale in Internet ha per esempio ridotto i discorsi d'odio soltanto in relazione a gruppi specifici¹⁴⁴.

Debolezza dei meccanismi di sanzionamento. Attualmente, la probabilità di essere puniti per idee razziste in Internet può essere ritenuta bassa. Se sono tecnicamente anonimi, gli autori possono eventualmente proteggersi da un'identificazione da parte delle autorità penali. Siccome le persone o i gruppi aggrediti sono fisicamente assenti, gli autori possono sottrarsi a possibili conflitti o atti di vendetta tramite il «logout». Manca inoltre la presenza diretta delle autorità: questo ruolo è attribuito generalmente ad altri utenti di Internet. I loro eventuali sguardi di riprovazione non sono visibili agli

¹³⁵ Keen 2007.

¹³⁶ Garland, Cult Education Institute, 27 marzo 2008: [Klan's new message of cyber-hate](#).

¹³⁷ Brown 2018: pag. 306.

¹³⁸ Brown 2018: pag. 298 segg.

¹³⁹ Lapidot-Lefler e Barak 2012.

¹⁴⁰ Suler 2004.

¹⁴¹ Hayne e Rice 1997 : pag. 432.

¹⁴² Christopherson 2007 : pag. 3048.

¹⁴³ Mondal et al. 2018; Rost, Stahel e Frey 2016.

¹⁴⁴ Cho, Kim e Acquisti 2012.

autori. Inoltre spesso gli utenti non condannano i commenti razzisti sulle pagine pubbliche di Internet con commenti contrari. E se gli utenti non sono chiamati ad assumersi le proprie responsabilità, le remore psicologiche e sociali si abbassano. Può anche sorgere l'impressione di uno spazio senza legge. Entrambi questi elementi favoriscono l'espressione incontrollata di idee razziste¹⁴⁵.

Attenzione attraverso la semplificazione e la drammatizzazione. In Internet, i discorsi d'odio rientrano tra i contenuti spesso premiati con una maggior attenzione. In uno spazio Internet traboccante di informazioni, la competizione per l'attenzione del pubblico è infatti generalmente agguerrita¹⁴⁶. Affinché tweet, blog e post sui social media siano letti, il loro valore d'intrattenimento è fondamentale. I conflitti, gli scandali o le infrazioni alle norme giuridiche e morali sono ideali per far leva sugli umori e suscitare emozioni. Ciò vale anche per i contributi viscerali, estremi, semplicistici e polarizzanti. Alla stessa stregua di semplici voci di corridoio o teorie complottiste, i discorsi d'odio razzisti corrispondono a questo schema di comunicazione semplificato e promettono più «condividi» e «mi piace»¹⁴⁷.

5.4.2 Architettura delle piattaforme online

Alla base delle condizioni comunicative analizzate vi sono architetture informatiche che determinano le possibilità di agire degli utenti e le restrizioni cui sono soggetti. Ciò può avvenire in modo palese (p. es. mediante opzioni di anonimato) o non palese (p. es. l'ordine in cui l'algoritmo presenta i contenuti su un sito web). Gli utenti possono anche essere premiati per il loro comportamento: gli algoritmi possono per esempio diffondere maggiormente i loro contenuti a seconda della formulazione scelta. Ma sono anche puniti: i loro contributi possono per esempio essere cancellati mediante opzioni di segnalazione. Conformemente a questa visione «materialistica», il razzismo nasce dalla complessa interazione sociotecnologica tra essere umano, funzionamento delle tecnologie e intenzioni dei designer della piattaforma¹⁴⁸. Sempre più spesso i ricercatori osservano che l'ambiente tecnologico è un humus ideale per il razzismo¹⁴⁹. Qui di seguito sono presentati alcuni fattori centrali:

Molteplici elementi tecnici di diffusione. I social media mettono a disposizione un ampio ventaglio di elementi – tradizionalmente testi, commenti, immagini, fotografie, simboli, video e musica – che consentono di produrre e far circolare idee razziste. Le regole delle piattaforme, come il limite di 280 caratteri su Twitter, possono indirizzare in modo raffinato la comunicazione, per esempio spingendo verso la semplificazione e ostacolando le argomentazioni più lunghe. In linea di principio, la compresenza di questi elementi aiuta a trasmettere al meglio e in modo convincente le idee razziste e le varie sfaccettature del razzismo: i testi aiutano a trasmettere l'ideologia razzista e i video toccanti ad alimentare le paure. Secondo la teoria della ricchezza dei media è il caso in particolare in presenza di contenuti personalizzati, scambi interattivi (p. es. nelle chat room) e stimoli combinati (p. es. video su YouTube)¹⁵⁰. La ricca offerta di formati consente di adattare velocemente e facilmente i contenuti razzisti al gruppo target, per esempio videogiochi e musica per bambini e adolescenti¹⁵¹. Anche i «mi piace» o la condivisione di [collegamenti ipertestuali](#) a siti web razzisti possono servire alla diffusione¹⁵². Sono strumentalizzati addirittura gli hashtag (#): l'hashtag della campagna #MAGA («Make America Great Again») del presidente degli Stati Uniti Donald Trump è stato sfruttato da gruppi razzisti per organizzarsi ed esprimersi ovunque e in qualsiasi momento¹⁵³. Per finire, anche elementi di controllo (p. es. il pulsante per la segnalazione di discorsi d'odio) possono decidere tra discorso d'odio sì o no.

Algoritmi. Gli algoritmi possono riprodurre il razzismo¹⁵⁴. Alcuni motori di ricerca forniscono per esempio risultati che ricalcano il precedente comportamento (razzista) dell'utente, confermandolo nelle sue opinioni (razziste). Se nel contesto di una crescente personalizzazione dei contenuti, questa logica

¹⁴⁵ Kaspar 2017: pag. 68.

¹⁴⁶ Marwick 2015: pag. 138.

¹⁴⁷ Hine et al. 2016.

¹⁴⁸ Murthy e Sharma 2019: pag. 196.

¹⁴⁹ P. es. Matamoros-Fernández 2017.

¹⁵⁰ Daft e Lengel 1986; Bliuc et al. 2018: pag. 81.

¹⁵¹ Cfr. [Stormfront for Kids](#).

¹⁵² Ben-David e Matamoros-Fernández 2016.

¹⁵³ Eddington 2018.

¹⁵⁴ Daniel 2018.

algoritmica produca le tanto citate «bolle di filtraggio»¹⁵⁵, ossia camere dell'eco create dagli algoritmi, è tuttavia un punto su cui le opinioni divergono. Nei dibattiti sui social media durante le elezioni in Indonesia, Lim¹⁵⁶ ha constatato che da soli gli algoritmi non formano «enclave algoritmiche». Utenti e algoritmi s'influenzano piuttosto a vicenda nell'ordinare, classificare e gerarchizzare persone, informazioni e preferenze politiche. Il razzismo trova poi sfogo nelle enclave risultanti. Un altro punto su cui le opinioni divergono è se gli algoritmi di raccomandazione dissuadano o motivino gli utenti ad accedere a contenuti estremisti¹⁵⁷. Manca ancora anche un giudizio definitivo sull'influsso concreto degli algoritmi sul razzismo.

Struttura a rete. Le reti sono sfruttate per diffondere meglio i discorsi d'odio e coordinare «campagne d'odio». Conformemente alla logica di fondo «precedenza alla libertà sul controllo», Internet elimina le frontiere territoriali attraverso un'interconnessione su larga scala¹⁵⁸. I gruppi d'odio sfruttano la situazione per raggiungere individui geograficamente e socialmente isolati e compensare la massa critica che manca localmente, contribuendo così a una «subcultura razzista globale»¹⁵⁹. Grazie alla fitta interconnessione, i loro contenuti si diffondono meglio. Un'analisi¹⁶⁰ di 21 milioni di post di circa 340 000 utenti nel forum Gab è giunta alla conclusione che i contenuti di coloro che propagano discorsi d'odio si diffondono in modo più rapido ed esteso e raggiungono un pubblico più ampio (rispetto a quelli di coloro che non propagano discorsi d'odio). A loro volta, queste fitte reti d'odio favoriscono la polarizzazione, dal momento che gli scettici le abbandonano progressivamente. Nelle camere dell'eco che si creano, le persone con idee affini sono sempre meno esposte a opinioni alternative¹⁶¹. Sono inoltre avviate «campagne d'odio»: non tutto ciò che sembra un'ondata d'indignazione spontanea del grande pubblico lo è veramente. Basti pensare alle «fabbriche di troll» di estrema destra¹⁶². In passato, per esempio, la rete di Reconquista Germanica contava parecchie migliaia di membri e utilizzava troll per disturbare (e sabotare) selettivamente discorsi online. Tali «eserciti di troll» agiscono anche a nome di attori politici affermati, come dimostra il caso dell'AKP, il partito della giustizia e dello sviluppo al governo in Turchia¹⁶³. Non è noto in che misura vi siano reti organizzate anche in Svizzera. Le esperienze raccolte suggeriscono tuttavia l'esistenza di interconnessioni¹⁶⁴.

Manipolazione metrica. Con l'apertura pluralistica del panorama mediatico, anche reti razziste possono ottimizzare i loro contenuti in termini numerici¹⁶⁵. Come alcuni pubblicitari e attori politici, comprano «like» («mi piace») oppure utilizzano falsi profili (ossia profili con dati falsi) o contenuti generati automaticamente mediante programmi informatici (bot). Nella concorrenza digitale lottano per conquistare follower e visibilità, al fine di essere efficaci sul piano politico. Secondo Fielitz e Marcks¹⁶⁶, questa possibilità fa il gioco del pragmatismo radicale dei gruppi estremisti e razzisti, che si sentono meno vincolati a direttive etiche e privilegiano una relazione strumentale con la verità – che può anche essere «nascosta» se ciò garantisce influenza e potere. I falsi profili e i bot sono impiegati soprattutto in campagne d'odio per manipolare discorsi sul piano numerico ed emarginare le opinioni divergenti. L'impiego di numerosi falsi profili può inoltre rendere «trendy» gli hashtag di stampo razzista – aumentando la probabilità che siano accolti dai media tradizionali e di riflesso visibili su tutte le piattaforme. Una minoranza rumorosa può così spacciarsi per un'ampia maggioranza. Un'analisi di migliaia di commenti e oltre un milione di «mi piace» su Facebook¹⁶⁷ è giunta alla conclusione che il cinque per cento più attivo degli utenti aveva generato il 50 per cento dei «mi piace» di contributi d'odio

¹⁵⁵ Pariser 2011.

¹⁵⁶ Lim 2017: pag. 422.

¹⁵⁷ Per un esempio di studio, cfr. Ledwich e Zaitsev 2019.

¹⁵⁸ Klein 2017.

¹⁵⁹ Perry e Olsson 2009: pag. 185; Fielitz e Marcks 2019: pag. 9.

¹⁶⁰ Mathew et al. 2019.

¹⁶¹ Marwick e Lewis 2017; pag. 18; Sunstein 2000.

¹⁶² Kreissel et al. 2019; Hine et al. 2016.

¹⁶³ Bulut e Yörük 2017.

¹⁶⁴ L'associazione #Netzcourage osserva su piattaforme social una dozzina di gruppi riferiti alla Svizzera con da 800 a 2000 utenti, che in parte si sovrappongono. Intervista con Jolanda Spiess-Hegglin, febbraio 2020.

¹⁶⁵ Fielitz e Marcks 2019: pag. 12.

¹⁶⁶ Fielitz e Marcks 2019: pag. 12.

¹⁶⁷ Kreissel et al. 2019: pag. 12.

e l'uno per cento più attivo addirittura il 25 per cento. E siccome la gente si accoda volentieri alla maggioranza, tali opinioni manipolate possono distorcere il processo di formazione delle opinioni.

Manipolazione contenutistica. Oltre alle metriche, le notizie sono manipolate anche sul piano dei contenuti: basti pensare alle fake news e alle teorie complottiste. Queste molle del discorso d'odio offrono soluzioni semplici a problemi complessi, facendo ricorso a stereotipi sprezzanti. In particolare, le fake news mirano a seminare l'incertezza e a disorientare¹⁶⁸. Circolano soprattutto in un sistema di notizie alternative, creato da ambienti di estrema destra. Possono anche raggiungere ampie fasce della società, come dimostra la commistione di fake news e discorsi d'odio razzisti contro i cinesi nel clima da fortezza assediata suscitato dal coronavirus (cfr. l'esempio nella fig. 6)¹⁶⁹. Chiunque può generare autonomamente contenuti manipolati, per esempio video complottisti di registi amatoriali su YouTube. Pur potendo essere tanto inventati quanto reali, i contributi manipolati sul piano dei contenuti sono sensibilmente distorti in una determinata direzione e spesso corroborati da statistiche false¹⁷⁰. Costrutti sociali (p. es. in Germania lo «straniero che dà fastidio»), personalizzazione (p. es. slogan come «la Merkel se ne deve andare») e moralizzazione coltivano un'«affettività aggressiva»¹⁷¹. Strategicamente, la cosa ha senso: pertinenti studi evidenziano che sui social media i contributi si diffondono maggiormente e più rapidamente se contengono parole che toccano il sentire morale e l'emotività o informazioni false¹⁷². Siccome generano più clic, sono potenzialmente prioritizzati dagli algoritmi, poiché in base alla logica della monetizzazione delle interazioni sociali promettono maggiori entrate alle imprese che operano in Internet.



Fig. 6: commento pubblicato online su 20min.ch nel marzo del 2020 (segnalato alla GRA).

Camuffamento. Le idee razziste possono circolare senza essere riconosciute immediatamente in quanto tali dagli utenti o dagli algoritmi di riconoscimento automatico. La nuova veste mira a evitare che questi contenuti siano associati all'estremismo di destra tradizionale e a farli sembrare più convincenti in una società a elevata competenza mediale. Movimenti come lo Stato islamico (ISIS) o il movimento alt-right sono passati dalla «propaganda nera» del discorso d'odio diretto alla «propaganda grigia», ossia a un discorso d'odio avvolto in un linguaggio in codice e ispirato a una cultura digitale satirico-ironica¹⁷³. A essere particolarmente popolari sono i memi¹⁷⁴, che attraggono il pubblico sul piano emotivo e trasmettono informazioni velocemente. I «pony nazisti» (fig. 7 e 8) sono un esempio di come si possano combinare contenuti attrattivi per destinatari ignari (in questo caso i bambini) con simboli nazisti e far sì che vengano inconsapevolmente ridiffusi da questi stessi destinatari¹⁷⁵. Elementi umoristici permettono inoltre di diffondere come se fossero barzellette contenuti non accettati¹⁷⁶. L'umorismo serve ad ampliare il ventaglio di idee accettate: «If you make racism or anti-Semitism funny, you can subvert the cultural taboo. Make people laugh at the Holocaust – you've opened a space in which history and fact become worthless» («Se rendete divertente il razzismo o l'antisemitismo potete

¹⁶⁸ Fielitz e Marcks 2019: pag. 11.

¹⁶⁹ Priebe, FAZ, 3 febbraio 2020: [Wie Rassisten das Coronavirus für sich nutzen.](#)

¹⁷⁰ Lanzke et al. 2013: pag. 25 segg.

¹⁷¹ Fielitz e Marcks 2019: pag. 10.

¹⁷² P. es. Brady et al. 2017; Vosoughi, Roy e Aral 2018.

¹⁷³ Klein 2017: pag. 27.

¹⁷⁴ Alcuni gruppi promuovono espressamente la «guerra memetica» anche quale strategia, come si può leggere nel «manuale strategico» Generation D del 18 maggio 2017: [Das Shitposting 1x1.](#)

¹⁷⁵ Winter 2019: pag. 51.

¹⁷⁶ Dittrich et al. 2020: pag. 42.

sovertire il tabù culturale. Fate ridere la gente dell'Olocausto – avete aperto uno spazio in cui la storia e i fatti non hanno più valore»), ha spiegato Keegan Hankes, ricercatore del Southern Poverty Law Center¹⁷⁷. Dove poi le direttive comunitarie (p. es. di Facebook) vietano il discorso d'odio, ma consentono l'umorismo, memi razzisti-umoristici possono circolare meglio. Altre strategie mirano anzitutto ad aggirare gli algoritmi di riconoscimento. Per «marchiare» singole persone come ebrei, si utilizza la consuetudine antisemita di scrivere i nomi ebraici tra parentesi triple, ossia (((nome))). Sono utilizzati anche codici come 88 per «Heil Hitler» (HH) ed emoji distorti¹⁷⁸. Nelle sue direttive interne sulla moderazione¹⁷⁹ Facebook elenca gli emoji che possono celare discorsi d'odio (fig. 9). Con il termine «dog whistling» razzista, infine, si descrive una lingua subliminale, decodificabile soltanto nel contesto e da persone iniziate (p. es. «ci vorrebbe una visita a domicilio»)¹⁸⁰. Senza il contesto, queste affermazioni sembrano innocue ed è praticamente impossibile perseguirle.

¹⁷⁷ Citato in Reitman, Rolling Stones, 2 maggio 2018: [All-American Nazis](#).

¹⁷⁸ Dittrich et al. 2020: pag. 40.

¹⁷⁹ Fisher, New York Times, 27 dicembre 2018: [Inside Facebook's secret rules for global political speech](#).

¹⁸⁰ Dittrich et al. 2020: pag. 37.

6 CONSEGUENZE NEGATIVE DELL'ODIO RAZZISTA ONLINE

6.1 Osservazione preliminare sull'amplificazione digitale

Non è facile rilevare gli effetti, a breve e a lungo termine, che il razzismo online ha sulle vittime, sul pubblico e sulla società nel suo complesso. Gli studi disponibili suggeriscono tuttavia che sono paragonabili a quelli del razzismo offline. Lo spazio digitale può addirittura amplificare le conseguenze negative, in quanto accresce le probabilità di un concorso di fattori potenzialmente lesivi¹⁸⁵. Le aggressioni fisiche, per esempio, possono provocare uno stress maggiore per la loro immediatezza, ma la violenza in rete raggiunge più rapidamente un pubblico molto vasto, il che aumenta il senso di vergogna delle persone attaccate. Secondo i giovani, in Svizzera il fenomeno del (cyber-)mobbing pubblico (rispetto a quello non pubblico) è particolarmente grave¹⁸⁶. Le persone prese di mira in Internet si trovano rapidamente di fronte a una serie di messaggi d'odio in crescita esponenziale e difficilmente possono sottrarsi nel tempo o nello spazio. Mentre i discorsi d'odio offline (p. es. in graffiti, articoli o nelle interazioni tra persone) alla fine scompaiono per effetto della riverniciatura delle pareti, dello smaltimento dei giornali o dell'attenuazione dei ricordi, i messaggi postati in rete sono permanenti e possono essere letti sempre e ovunque. Non appena vengono pubblicati in Internet, sono praticamente impossibili da cancellare in modo definitivo, poiché possono essere salvati e pubblicati su altre piattaforme in qualsiasi momento. Questo carattere permanente può essere traumatico per le vittime, che hanno la sensazione di non riuscire più a controllare la situazione.

6.2 Vittime

I discorsi d'odio online, di matrice razzista e non, possono provocare in chi ne è vittima non soltanto emozioni dolorose, ma anche sintomi fisici. In Gran Bretagna, musulmani oggetto di questo tipo di messaggi hanno per esempio dichiarato di sentirsi più vulnerabili e insicuri e di reagire con crisi di depressione e ansia e adattandosi alla società maggioritaria (p. es. rinunciando al velo e alla barba), nel timore di essere aggrediti anche fisicamente¹⁸⁷. In un sondaggio condotto in Germania¹⁸⁸, due terzi delle persone direttamente interessate da commenti d'odio in rete hanno dichiarato di averne sofferto e di patire una serie di problemi psicologici (demoralizzazione, svogliatezza, paura e ansia, attacchi di depressione e disagio con sé stessi). Sono stati menzionati anche effetti correlati, quali problemi sul lavoro o nell'istituzione formativa frequentata. Secondo studi svolti negli Stati Uniti, queste conseguenze si verificano indipendentemente da eventuali, contemporanee aggressioni offline, motivo per cui si ritiene che siano realmente riconducibili agli attacchi subiti in rete¹⁸⁹. In un sondaggio australiano condotto su 103 persone direttamente colpite dal razzismo online è stata riscontrata una gamma di reazioni emotive molto più ampia: il 48 per cento ha provato rabbia e frustrazione, il 38 per cento disgusto, il 25 per cento divertimento, il 18 per cento senso d'impotenza o depressione, il 13 per cento vergogna e il 12 per cento solidarietà con gli autori¹⁹⁰. Una piccola percentuale (10 %) ha riferito di problemi di stomaco, mal di testa e palpitazioni. E nei fatti, come hanno reagito queste persone alle violenze in rete? Due terzi hanno affermato di essersi difesi attivamente, per esempio postando un contro-commento, segnalando il post o bloccandone gli autori. Una persona su sette ha ignorato i messaggi.

Tra gli effetti a più lungo termine si segnala, per esempio, un calo della performance o la disconnessione digitale. In un sondaggio sul lungo periodo¹⁹¹ condotto tra giovani americani, coloro che hanno osservato un aumento del razzismo online contro la propria persona hanno riferito di un concomitante calo motivazionale a scuola. La reazione di «silenziarsi», ossia di pubblicare meno contenuti in Internet e ridurre così la propria presenza in rete, è invece diffusa soprattutto tra le personalità (semi)pubbliche.

¹⁸⁵ Brown 2018: pag. 306 segg.

¹⁸⁶ Sticca e Perren 2013.

¹⁸⁷ Awan e Zempi 2015: pag. 37.

¹⁸⁸ Geschke et al. 2019: pag. 27.

¹⁸⁹ Tynes et al. 2018.

¹⁹⁰ Jakubowicz et al. 2017 : pag. 77-79.

¹⁹¹ Tynes, Torro e Lozada 2019.

Non c'è da stupirsi, dato che questo gruppo di persone è colpito più degli altri dai discorsi d'odio online¹⁹². Gli scienziati che diffondono attivamente i risultati delle loro ricerche sui media sociali riferiscono, per esempio, di essere oggetto di attacchi razzisti in rete e di dover lottare contro la tentazione di disconnettersi dai social media¹⁹³. Anche i giornalisti attivi in Svizzera dichiarano di tendere maggiormente a evitare il pubblico in risposta agli attacchi online, siano essi razzisti o meno¹⁹⁴. Questi risultati suggeriscono che i discorsi d'odio online possono limitare le persone nell'esercizio dei loro diritti umani, come la libertà di espressione, la libertà di religione o la sicurezza personale.

Infine, vi sono indizi secondo cui i discorsi d'odio in rete arrecherebbero maggiori sofferenze a determinati gruppi della popolazione, in particolare i giovani, le donne e le persone con una combinazione di caratteristiche minoritarie¹⁹⁵. I giovani, quindi, non soltanto sarebbero più spesso coinvolti in questi attacchi, ma ne patirebbero anche di più. Non è invece ancora del tutto chiaro se la maggior gravità dell'impatto percepito dalle donne sia dovuto alla frequenza della violenza in rete o alla sua natura oppure ad altri fattori¹⁹⁶.

6.3 Pubblico e società nel suo complesso

Il razzismo online può avere un impatto anche su chi vi assiste da fuori – vale a dire sul pubblico degli internauti – e, a lungo termine, sulla società nel suo complesso. I discorsi d'odio sia offline che online possono acutizzare le tensioni sociali e sono sempre stati considerati un segnale premonitore di instabilità politica e violenza. Lo dimostrano studi sull'impatto della radio sul livello di violenza durante il genocidio in Ruanda o sull'effetto della propaganda radiofonica, nella Germania nazista, sulla violenza antisemita¹⁹⁷. I discorsi d'odio fungerebbero da terreno fertile per l'aggressione fisica e l'intimidazione, in quanto le calunnie intrise di stereotipi così veicolate sarebbero propedeutiche alla disumanizzazione e allenterebbero i freni che trattengono dal compiere reati¹⁹⁸. Tale «preparazione» potrebbe portare al risultato che i discorsi d'odio razzisti, se diffusi in Internet e raramente segnalati o sanzionati, diventino socialmente accettabili¹⁹⁹. Il pubblico potrebbe avere l'impressione che, sostanzialmente, tutti quanti odiano e che per questo non si viene puniti. In tal modo i discorsi d'odio verrebbero percepiti come la norma, il che a sua volta spingerebbe le persone a diffondere messaggi d'odio in prima persona.

I discorsi d'odio online possono dunque diventare «contagiosi», nel senso che possono indurre il pubblico a pensare e agire in modo più ostile. È un meccanismo visibile negli esperimenti in cui le discussioni vengono «prefabbricate»: ad alcuni partecipanti vengono presentati commenti d'odio e ad altri commenti neutrali²⁰⁰. Il risultato è che, nei partecipanti del primo gruppo, la discussione lascia pregiudizi più forti e opinioni più polarizzate sul tema. Il primo gruppo valuta inoltre il pubblico in generale come più polarizzato, il che a sua volta può indurre le singole persone a esprimere opinioni personali più polarizzate. Un'altra caratteristica dei partecipanti del primo gruppo è la minor propensione a donare denaro ai rifugiati. Anche altri esperimenti hanno mostrato effetti negativi a lungo termine: un esperimento polacco, per esempio, ha mostrato che i partecipanti che hanno osservato regolarmente discorsi d'odio online diretti contro gruppi minoritari (p. es. musulmani) si sono «desensibilizzati» ad attacchi di questo tipo, iniziando anche a prendere interiormente le distanze da questi gruppi, ad apprezzarli di meno e ad avere più pregiudizi nei loro confronti²⁰¹. I discorsi d'odio, la giustizia *fai da te* e la disinformazione in rete sono in ultima analisi strettamente collegati. In Internet possono circolare

¹⁹² Isbister et al. 2018.

¹⁹³ Barlow e Awan 2016.

¹⁹⁴ Stahel e Schoen 2019.

¹⁹⁵ P. es. Bucchianeri et al. 2014; Geschke et al. 2019.

¹⁹⁶ Celik 2019; Duggan 2017; Stahel e Schoen 2019: secondo questi studi, per esempio, uomini e donne sarebbero vittima in ugual misura di attacchi online. Fox e Tang 2017: secondo questo studio, invece, le maggiori sofferenze patite dalle donne sarebbero imputabili ad attacchi sessisti.

¹⁹⁷ Gagliardone et al. 2016; Yanagizawa-Drott 2014; Adena et al. 2015.

¹⁹⁸ Humanrights.ch consultato il 6 febbraio 2017: [Hassreden: Die Grenzen der Meinungsäußerungsfreiheit](#).

¹⁹⁹ Questa teoria riguarda chi è spettatore di attacchi online: Jakubowicz et al. 2017: pag. 79; Landesanstalt für Medien NRW 2018: pag. 4.

²⁰⁰ Esperimenti emblematici sono descritti nei seguenti studi: Anderson et al. 2014; Hsueh et al. 2015; Ziegele, Koehler e Weber 2018.

²⁰¹ Soral, Michal e Winiewski 2018.

sospetti non provati e false accuse accompagnati da incitamenti alla violenza (p. es. impiccagioni, incendi dolosi). Nella realtà sono noti innumerevoli casi di persone stigmatizzate la cui reputazione è stata distrutta per sempre, nonostante le prove della loro innocenza. Tutti questi elementi contribuiscono al dilagare di un clima di pensiero e di discussione nefasto.

I discorsi d'odio online possono inoltre costituire un rischio per la diversità dell'opinione pubblica digitale: il fenomeno della disconnessione digitale non tocca soltanto le vittime della violenza in rete, ma anche il pubblico in generale. In un sondaggio tedesco²⁰², quasi la metà degli utenti di Internet ha dichiarato di partecipare con minore frequenza a discussioni politiche online a causa di (minacciosi) commenti d'odio. Il 15 per cento ha dichiarato di avere già disattivato o cancellato una volta il proprio profilo Internet a causa dei discorsi d'odio.

È interessante notare, tuttavia, che le probabilità di disconnettersi dai media sociali non sono uguali per tutti i gruppi di persone: in un altro sondaggio²⁰³, gli elettori dell'Alternativa per la Germania hanno dichiarato di ritirarsi molto meno spesso, rispetto agli elettori di altri partiti, dalle discussioni online a causa dei commenti d'odio. Quest'indicazione può suggerire quali siano i gruppi di persone che sceglieranno di ritirarsi dalle discussioni pubbliche se sul lungo periodo l'odio in rete dovesse aumentare e quali invece no.

Infine, i discorsi d'odio online sono legati ai crimini d'odio offline, anche se i nessi causali non possono essere stabiliti con chiarezza. Di diversi attentatori si sa che si sono radicalizzati in Internet, per esempio l'assassino 21enne che ha sparato a nove afroamericani in una chiesa di Charleston (USA) nel 2015: come sua prima ispirazione razzista ha indicato il [Council of Conservative Citizens](#) (Consiglio dei cittadini conservatori), un'organizzazione suprematista bianca americana²⁰⁴. Oppure Anders Breivik, un estremista di destra e islamofobo membro di [Stormfront.org](#), che ha ucciso 77 persone in Norvegia. Anche gli autori degli attacchi di estrema destra compiuti a Pittsburgh e San Diego e a Christchurch in Nuova Zelanda nel 2018 e nel 2019 frequentavano subculture digitali di estrema destra²⁰⁵. I risultati di studi scientifici, ricchi di dati, avvalorano queste osservazioni. Uno studio tedesco di lungo periodo mostra per esempio che la crescita dei sentimenti ostili ai rifugiati sulla pagina Facebook del partito Alternativa per la Germania va di pari passo con incendi dolosi e aggressioni fisiche contro i richiedenti l'asilo²⁰⁶. Per chiarire gli effetti di Internet sui crimini d'odio offline negli Stati Uniti, tra il 1999 e il 2008 è stato effettuato un confronto tra numero di provider a banda larga e numero di crimini d'odio razzista²⁰⁷, dal quale è emerso che l'aumento dell'accesso alla banda larga in una determinata regione ha comportato anche un aumento dei crimini d'odio, mentre altri reati come il furto con scasso o l'omicidio sono rimasti stabili. Un fenomeno, questo, più frequente nelle aree con una maggiore segregazione sociale e in cui gli utenti cercavano più spesso parole razziste in Internet. Secondo gli autori, la rete aumenterebbe la propensione *in nuce* ad atti razzisti offline, perché in Internet l'odio si diffonde più rapidamente e le persone che la pensano allo stesso modo si interconnettono più facilmente. Infine, anche il cosiddetto «dog whistling» (cfr. cap. 5.4.2) praticato da personaggi pubblici può incitare ai crimini d'odio. Da uno studio è emerso che l'aumento dell'attività su Twitter (in particolare del «dog whistling») dei membri della Camera dei Rappresentanti degli Stati Uniti che hanno il maggior numero di follower nazionalisti ha comportato anche un aumento del numero di crimini d'odio offline²⁰⁸. Secondo lo studio, i politici possono incitare sottilmente i loro seguaci alla violenza, confermandone con la loro retorica l'ideologia estremista e dando così potere a queste comunità.

Per le vittime e per la società nel suo complesso, le conseguenze a breve e lungo termine dei discorsi d'odio online sono pesanti e in futuro andrebbero studiate anche nel contesto svizzero.

²⁰² Geschke et al. 2019: pag. 28.

²⁰³ Eckes et al. 2018: pag. 6.

²⁰⁴ Southern Poverty Law Center, 21 giugno 2015: [The Council of Conservative Citizens: Dylann Roof's gateway into the world of white nationalism](#).

²⁰⁵ Guhl et al. 2020: pag. 7.

²⁰⁶ Müller e Schwarz 2019.

²⁰⁷ Chan, Ghose e Seamans 2013.

²⁰⁸ Chyzh, Nieman e Webb 2019.

7 CONTROMISURE, ATTUAZIONE ED EFFICACIA

Gli effetti sempre più visibili del razzismo online spingono i decisori, i gestori di social media, i ricercatori, le ONG e la società civile di molti Paesi a sviluppare contromisure. Questo approccio attivo risponde alla grande necessità di agire percepita da un'ampia fetta di popolazione in relazione ai discorsi d'odio online, rilevata quanto meno da alcuni sondaggi condotti in Germania²⁰⁹. Le contromisure si articolano sui due assi prevenzione e intervento, anche se non sempre è possibile una chiara delimitazione. Per esempio, la contronarrazione nel senso lato (cfr. cap. 7.5.4) può essere utilizzata a lungo termine per creare una narrazione non violenta. Allo stesso tempo, può contrapporsi apertamente e selettivamente a specifici messaggi d'odio.

Questo capitolo presenta le contromisure messe in campo da attori e istituzioni rilevanti, in Svizzera e all'estero (in particolare nei Paesi vicini). In primo luogo illustra come il tema dei discorsi d'odio online viene affrontato dal sistema giuridico, dai gestori di social media, dai media tradizionali e dalla ricerca. Questi attori formano un ecosistema strutturale, all'interno del quale alcuni di essi svolgono funzioni di controllo. Capire qual è il loro margine di manovra in relazione ai discorsi d'odio in Internet è fondamentale per poter analizzare in chiave critica le misure che la popolazione civile può adottare e la loro efficacia²¹⁰. La panoramica qui proposta non ha alcuna pretesa di completezza. Le importanti questioni di responsabilità a monte delle discussioni sono soltanto accennate e non trattate in modo esaustivo. Infine, si solleva il problema fondamentale dell'attualità e della flessibilità: spesso le soluzioni sono in ritardo rispetto alla minaccia rappresentata dai discorsi d'odio sia sul piano quantitativo che su quello della velocità di diffusione, visto che gli autori dei messaggi si adattano con agilità a Internet²¹¹.

7.1 Legislazione e giurisprudenza in materia di discorsi d'odio

Il presente sottocapitolo fornisce una breve panoramica del quadro giuridico in cui si inseriscono la legislazione e la giurisprudenza sui discorsi d'odio di stampo razzista e illustra le altre problematiche legate alla loro diffusione digitale. Una descrizione precisa della situazione giuridica in Svizzera sarà pubblicata ancora nel 2020 sul sito Internet del SLR. Inoltre, la [Guida giuridica on-line](#)²¹² del SLR contiene informazioni di base sulle norme e sui rimedi giuridici applicabili in caso di discriminazione razziale.

La legislazione sui discorsi d'odio di stampo razzista deve sempre considerare due diritti umani fondamentali che si pongono in un rapporto antitetico: la libertà di espressione e la protezione contro la discriminazione. Tutte le convenzioni internazionali sui diritti umani²¹³ lasciano un certo margine di manovra nella ponderazione di questi due diritti, in modo che ogni Paese possa tener conto delle proprie specificità. Per esempio, nell'area anglosassone la libertà di espressione ha un peso maggiore rispetto all'Europa, mentre l'Europa tende ad essere più tollerante verso contenuti di tipo sessuale o persino sessista di quanto non lo siano gli Stati Uniti.

Pertanto, i singoli Stati decidono in modo relativamente autonomo quali affermazioni (in parole, gesti, immagini o suoni) considerano così gravi da dover essere perseguite penalmente, quali affermazioni possono portare almeno all'avvio di un'azione civile e quali responsabilità i media o i gestori di social media devono assumersi per i contenuti che pubblicano, moderano o diffondono.

Tuttavia, non soltanto la legislazione, ma anche la sua interpretazione nell'applicazione concreta dipende dal contesto temporale e culturale. Contenuti che vent'anni fa sembravano ancora accettabili oggi possono rappresentare una grave offesa alla sensibilità sociale. Viceversa, il discorso razzista o l'incitamento all'odio contro singoli gruppi di popolazione danno prova di una grande capacità di adattamento e trovano sempre nuovi modi e mezzi per continuare a diffondersi passando tra le maglie

²⁰⁹ Geschke et al. 2019: pag. 32.

²¹⁰ Per una breve valutazione delle misure finora adottate in Germania, cfr. Dittrich et al. 2020: pag. 74.

²¹¹ Baldauf, Ebner e Guhl 2018.

²¹² Percorso: www.frb.admin.ch > Diritto e consulenza > [Guida giuridica](#).

²¹³ Rilevanti per la Svizzera sono in particolare le convenzioni delle Nazioni Unite sui diritti umani (Dichiarazione universale dei diritti dell'uomo, Patto internazionale relativo ai diritti economici, sociali e culturali [Patto ONU I], Patto internazionale relativo ai diritti civili e politici [Patto ONU II]) e la Convenzione europea dei diritti dell'uomo.

della legislazione²¹⁴. Pertanto, quando si valuta un contenuto problematico occorre sempre considerare il contesto in cui viene diffuso. E nel caso specifico delle forme codificate, implicite o subdole del discorso d'odio di stampo razzista, non esistono ricette universalmente valide.

In Svizzera la protezione contro la discriminazione razziale è sancita dalla Costituzione federale (art. 8 Cost.). Questo diritto fondamentale si concretizza soprattutto nella norma penale contro il razzismo, che vieta l'espressione in pubblico di contenuti razzisti di una certa gravità (art. 261^{bis} CP). Nel caso di esternazioni proferite nella sfera privata o di esternazioni la cui gravità e il cui carattere razzista sia difficile da provare, si può configurare un delitto contro l'onore perseguibile penalmente (art. 173, 174 o 177 CP). Sul piano del diritto privato, le persone vittima di messaggi d'odio possono invocare la protezione della personalità (art. 28 e seg. CC) e intentare per esempio un'azione legale per chiedere la rimozione di un contenuto. Nel caso di messaggi d'odio pubblicati nei media tradizionali, ci si può appellare agli organi di autoregolamentazione istituzionalizzati, come il Consiglio svizzero della stampa e l'Autorità indipendente di ricorso in materia radiotelevisiva (art. 4 cpv. 1 LRTV)²¹⁵.

Questi rimedi giuridici sono noti e la loro applicazione all'ambiente digitale è in costante evoluzione. La giurisprudenza è chiamata per esempio a stabilire cosa debba essere considerata un'esternazione pubblica e cosa no in Internet, come nel caso dei gruppi chiusi su Facebook o WhatsApp²¹⁶. Vista la confusione di ruoli tra autori, produttori, diffusori e utenti, occorre chiarire chi deve assumersi quale responsabilità per i contenuti razzisti e la loro diffusione. Se questi sono diffusi in formato digitale attraverso i media tradizionali, occorre considerare anche le disposizioni del diritto dei media e trasporle al contesto digitale. Non bisogna infine dimenticare l'importanza del diritto internazionale, dato che gli utenti che pubblicano commenti si trovano spesso in un Paese diverso da quello in cui ha sede il mezzo di comunicazione digitale o il gestore di social media. Questo può rappresentare un problema considerevole per l'applicazione della legge.

Come nel caso dei discorsi d'odio offline, spesso non è possibile adire le vie legali contro i discorsi d'odio di stampo razzista pubblicati online. Come già accennato, in molti casi il contenuto è formulato in modo tale da non poter essere qualificato in modo sufficientemente chiaro come discriminazione razziale penalmente rilevante. Inoltre, gli ostacoli per avviare una procedura civile sono considerevoli.

Vi è però sempre la possibilità di segnalare i contenuti razzisti pubblicati in Internet. Oltre che alle autorità di ricorso per i media tradizionali, i contenuti possono essere segnalati anche ai gestori di social media. Oggi la trasparenza dei criteri di verifica (che spesso si basano sul contesto e la giurisprudenza del Paese in cui il gestore ha la propria sede) e la qualità del controllo parzialmente automatico devono ancora essere sottoposte a un esame critico (cfr. cap. 7.2). A questo si aggiunge che i gestori non reagiscono sistematicamente a tutte le segnalazioni o non possono farlo vista la grande quantità di contenuti e segnalazioni. Per questo alcune piattaforme hanno messo a punto un sistema di segnalatori attendibili (*trusted flaggers*), in base al quale le segnalazioni provenienti da servizi o organizzazioni che hanno questo statuto sono trattate in via prioritaria. Sul suo sito, il SLR offre un elenco dei centri di segnalazione e dei segnalatori attendibili in Svizzera ([Link](#)²¹⁷).

7.2 Gestori di social media

Facebook, Twitter e altri gestori di reti sociali possono arginare i discorsi d'odio di stampo razzista in primo luogo moderando i contenuti pubblicati. La moderazione permette di ridurre la probabilità che gli utenti abituali siano esposti inavvertitamente a discorsi d'odio²¹⁸. Le piattaforme possono rimuovere i contenuti che violano le loro direttive (cfr. cap. 3.1) o bloccarli in determinati Paesi in virtù della legislazione nazionale: Facebook, per esempio, vieta e blocca contenuti nazisti in Germania, ma li autorizza (e ne permette la pubblicazione) negli Stati Uniti. Le piattaforme possono anche bloccare

²¹⁴ Un esempio tipico è l'espressione «Get back to work!» che si è diffusa durante la guerra civile del 1994 in Ruanda come un chiaro appello al genocidio contro i Tutsi: cfr. Schabas 2001: pag. 161.

²¹⁵ In linea di principio, in virtù della libertà e dell'indipendenza sancite dalla Costituzione federale (art. 17 Cost.), i media sono tenuti a garantire il rispetto della Dichiarazione dei doveri dei giornalisti.

²¹⁶ Refaeil e Wiecken 2018.

²¹⁷ Percorso: www.frb.admin.ch > Campi d'attività > Media e Internet > Segnalazione.

²¹⁸ Siegel 2020: pag. 27 segg.

singoli profili o intere community. Per esempio, nel giugno del 2020 Facebook ha cancellato circa 190 profili social che erano collegati a gruppi di suprematisti bianchi e pianificavano attività concrete. In questo contesto, Twitter ha cancellato alcuni profili di estrema destra che avevano molti follower²¹⁹. Per identificare i discorsi d'odio, le piattaforme utilizzano algoritmi con software di riconoscimento di parole e/o affidano agli utenti e agli organismi di reclamo il compito di segnalare contenuti problematici. La maggior parte dei gestori di social media come Facebook, YouTube, Twitter, Instagram, Snapchat o TikTok hanno integrato una funzione di segnalazione²²⁰. Moderatori per lo più interni decidono sulla base di direttive se cancellare o bloccare i contenuti. Per promuovere la rimozione di messaggi che istigano alla violenza e all'odio, nel 2016 la Commissione europea ha pubblicato un codice di condotta²²¹ facoltativo in collaborazione con Facebook, Twitter, YouTube e Microsoft (*Code of Conduct on Countering Illegal Hate Speech Online*). In virtù di questo codice, le piattaforme che l'hanno sottoscritto (alle quali nel frattempo se ne sono aggiunte altre) sono tenute a rimuovere i discorsi d'odio (secondo la definizione dell'UE).

Qual è il vantaggio di un disciplinamento da parte dei gestori di social media? Secondo Brown²²², i gestori sono paradossalmente sia i maggiori diffusori che i migliori regolatori dei discorsi d'odio. In virtù del loro ruolo centrale, hanno la flessibilità necessaria per affrontare meglio di chiunque altro le attuali forme di discorso d'odio. Grazie al contatto costante con i contenuti pubblicati, i moderatori sono in grado di riconoscere le nuove forme in cui il razzismo online si manifesta, che possono poi essere recepite rapidamente nelle direttive dei rispettivi social media. In teoria i contenuti possono essere inoltre cancellati velocemente dopo la pubblicazione, anche perché le aziende – a differenza dei tribunali – non devono giustificare la rimozione all'opinione pubblica. Brown aggiunge un argomento morale: chi inventa o mantiene una tecnologia che promuove (anche non intenzionalmente) azioni dannose o illegali deve provvedere all'introduzione di misure per prevenirle o limitarle, in particolare se trae vantaggio da questa tecnologia.

Come funziona la moderazione operata sulle piattaforme? In generale, questi meccanismi sono sempre più apprezzati nonostante vi sia ancora un margine di miglioramento. Dittrich et al.²²³ rilevano che il modo in cui le piattaforme affrontano la problematica dei discorsi d'odio online è migliorato in termini di linee guida e possibilità di segnalazione (p. es. segnalatori attendibili), implementazione delle linee guida, rafforzamento della contronarrazione (p. es. attraverso la pubblicità) e condizioni di formazione e di lavoro migliori per i moderatori. Anche l'UE valuta positivamente²²⁴ l'attuazione del codice di condotta: in media i gestori hanno trattato l'89 per cento dei contenuti segnalati entro 24 ore e ne hanno rimosso il 72 per cento (un risultato considerato soddisfacente). Il feedback agli utenti e ai segnalatori attendibili potrebbe però essere intensificato.

Non è ancora chiaro in che misura la moderazione possa effettivamente ridurre i discorsi d'odio online. Ricercatori, giornalisti e politici attirano regolarmente l'attenzione sui punti critici. Se, per esempio, la percentuale di contenuti rimossi rispetto ai contributi segnalati è nota, spesso non si conosce la proporzione di quelli segnalati rispetto al totale²²⁵. Anche il riconoscimento automatico e algoritmico è stato criticato perché può classificare come discorsi d'odio contenuti neutrali o contronarrazioni. Non tiene inoltre sufficientemente conto dei contesti e delle lingue²²⁶ locali – l'esempio dello svizzero tedesco è emblematico. Per questi e altri motivi, è stata chiesta ripetutamente²²⁷ a livello internazionale maggiore trasparenza, in particolare per quanto riguarda i moderatori (competenza, numero ecc.), visto anche che in passato le loro cattive condizioni di lavoro e lo stress psicologico cui sono esposti sono stati più

²¹⁹ Klepper, ABC News del 6 giugno 2020. [Facebook removes nearly 200 accounts tied to hate groups.](#)

²²⁰ Dato che i sistemi di segnalazione sono molto diversi tra loro, la Fondazione contro il razzismo e l'antisemitismo GRA nel suo tool di segnalazione rimanda direttamente ai servizi di segnalazione dei principali gestori (p. es. Facebook, Twitter); percorso: [www.gra.ch > Bildung > Hate speech > Vorfall melden.](#)

²²¹ Jourová 2019a.

²²² Brown 2018 : pag. 310.

²²³ Dittrich et al. 2020.

²²⁴ Jourová 2019a; Jourová V. 2019b. Per una valutazione simile, anche se un po' più critica, cfr. sCAN 2019.

²²⁵ Siegel 2020.

²²⁶ Siegel 2020: pag. 27.

²²⁷ Per maggiori informazioni cfr. INACH, 19 novembre 2019: [The state of policy on cyber hate in the EU.](#)

volte oggetto di critiche²²⁸. È stata inoltre auspicata un'attuazione più trasparente delle linee guida per la moderazione (p. es. tenendo conto delle leggi nazionali).

I primi risultati degli studi scientifici condotti suggeriscono che la rimozione e il blocco dei contenuti sono efficaci a lungo termine, ma possono avere anche effetti collaterali. La decisione di Reddit di bloccare singoli *hate threat* ha indotto molti utenti ad abbandonare il sito. Parallelamente, gli utenti rimasti hanno pubblicato un numero nettamente inferiore di messaggi d'odio e hanno interagito meno²²⁹. I discorsi d'odio non si sono quindi semplicemente spostati su altri *thread*. Inoltre, si osserva soltanto in misura limitata la spesso paventata migrazione verso piattaforme meno moderate (*deplatforming*)²³⁰. Questo potrebbe essere riconducibile al fatto che gli utenti bloccati registrano un calo di follower ogni volta che creano un nuovo profilo. In linea di principio, il blocco potrebbe quindi essere utile, a condizione che sia solidamente motivato e non presti il fianco ad accuse di censura. Ciò non toglie che vi sono effetti collaterali dannosi da non sottovalutare. Per esempio, il *deplatforming* può spingere gli utenti a passare su piattaforme più estreme, dove sono meno confrontati a opinioni moderate («camere dell'eco»), favorendo così la radicalizzazione²³¹.

7.3 Media tradizionali in Svizzera

I media tradizionali riflettono i dibattiti in corso nella società e contribuiscono ad alimentarli²³². Sono quindi decisivi anche per la prevenzione o il rafforzamento dei discorsi d'odio razzisti, con i contenuti che pubblicano e le modalità con cui gestiscono la propria community, ossia i propri utenti (per gli aspetti giuridici relativi alle aziende mediatiche cfr. cap. 7.1). I membri della community commentano i contenuti e gli articoli nelle apposite colonne dei siti e sui profili social dei media. In questi due «luoghi» si applicano regole di pubblicazione e possibilità di moderazione diverse. Sui siti dei media gli utenti devono spesso indicare il loro indirizzo e-mail o il loro nome, mentre sui profili social possono pubblicare commenti in forma anonima utilizzando un falso profilo. In questo secondo caso i loro post sono collegati ai loro profili e alle loro reti social, il che influisce sul tenore dei commenti. Sui propri siti i media possono inoltre controllare i contenuti prima della pubblicazione e se del caso chiudere la rubrica dei commenti per limitare i messaggi d'odio. Sui social media entrambe le cose sono più difficili.

Il tipo di contenuti pubblicati e le modalità di gestione dei post della community variano da un media tradizionale all'altro. Attraverso la presentazione e il tono dei contenuti che pubblicano, i media possono per esempio incoraggiare l'emarginazione di determinati gruppi sociali. Così come i personaggi pubblici possono incitare velatamente ai discorsi d'odio²³³, anche i media possono, con il loro modo di presentare le notizie, confermare schemi di pensiero escludente fornendo indirettamente una legittimazione a sostenerli pubblicamente. Un elemento da non ignorare è la diversa concezione di community²³⁴. I media si distinguono per l'intensità della comunicazione unilaterale o bilaterale, cioè per l'intensità con cui moderano la propria community. Per prevenire i discorsi d'odio, il giornalista e imprenditore svizzero Hansi Voigt²³⁵ raccomanda una presenza attiva dei moderatori nelle rubriche dei commenti in modo da evitare che le discussioni sfuggano di mano e da ridurre gli spazi di diffusione. Voigt menziona inoltre la possibilità di far valutare e/o moderare anche dalla community i contenuti pubblicati dai suoi membri, in modo da coinvolgerla attivamente nell'impostazione del clima di discussione. Teoricamente, questi approcci sono promettenti perché gli utenti si sentono ascoltati e capiscono di muoversi in un ambiente in cui vigono regole di comunicazione: la presenza di un interlocutore reale riduce la disinibizione digitale²³⁶. Queste misure possono avere un impatto anche sui giornalisti e sul loro lavoro, visto che a volte anche in Svizzera sono bersaglio di attacchi e commenti d'odio da parte del pubblico²³⁷. Da studi

²²⁸ Newton, The Verge del 25 febbraio 2019: [The Trauma Floor](#)

²²⁹ P. es. Chandrasekharan et al. 2017.

²³⁰ Guhl et al. 2020 : pag. 11.

²³¹ Marwick und Lewis 2017: S. 18; Sunstein 2000.

²³² Questo sottocapitolo si basa in particolare sul rapporto del Servizio per la lotta al razzismo 2019: pag. 55 segg.

²³³ Chyzh et al. 2019.

²³⁴ Le affermazioni riportate si basano tra l'altro su un'intervista con Hansi Voigt (dasnetz.ch; bajour), aprile 2020.

²³⁵ Intervista con Hansi Voigt (dasnetz.ch; bajour), aprile 2020.

²³⁶ Suler 2004.

²³⁷ Stahel e Schoen 2019.

condotti²³⁸ emerge che i giornalisti sentono la pressione di dover essere presenti online. Tuttavia, l'esposizione pubblica che ne deriva aumenta il rischio di subire attacchi. D'altra parte, i giornalisti non sempre ricevono il necessario sostegno dalla testata o dall'agenzia di stampa per la quale lavorano. Prevenire o ridurre i discorsi d'odio può quindi non soltanto promuovere un clima pacato nella copertura mediatica e nei dibattiti pubblici, ma anche permettere ai giornalisti di svolgere serenamente il proprio lavoro.

7.4 Attività di ricerca in Svizzera

La ricerca, che sia condotta da università, scuole universitarie professionali, think tank o altre istituzioni, svolge un ruolo fondamentale nell'acquisizione di conoscenze sul razzismo online. La ricerca scientifica sui discorsi d'odio online in generale è cresciuta in modo esponenziale, di pari passo con l'attenzione dei media per l'argomento. Se nel 2005 gli articoli pubblicati a livello internazionale erano una cinquantina, nel 2018 se ne contavano 550 (cfr. grafico 10)²³⁹.

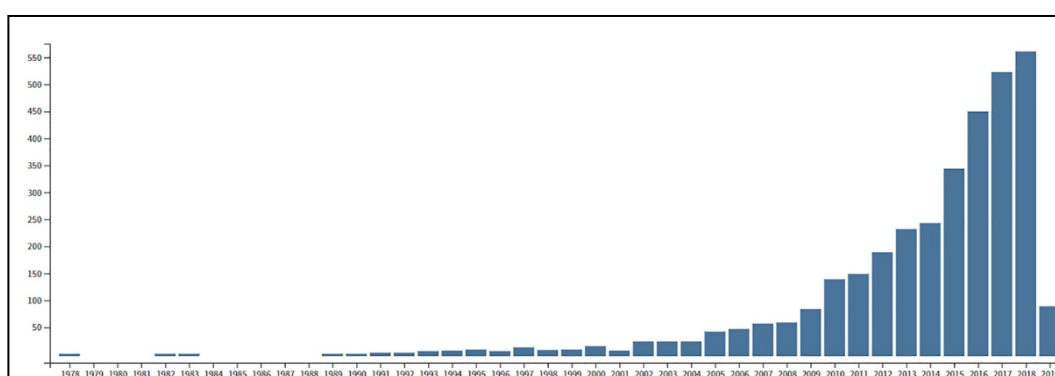


Grafico 10: evoluzione del numero di pubblicazioni sull'odio online dal 1978 al 2019 (marzo).

I lavori di ricerca si focalizzano sulla descrizione e la spiegazione dei discorsi d'odio online in generale e di quelli razzisti in particolare. Si parte dalla definizione dei concetti rilevanti che per la loro complessità richiedono un approccio multidisciplinare (cfr. cap. 3.1 e 3.2). In questo contesto, vengono esaminati in particolare i seguenti aspetti: con che frequenza e in che forma (p. es. aperta o camuffata) si manifestano i discorsi d'odio? Contro chi sono rivolti? Che impatto hanno sulla società? Per ottenere primi risultati, i ricercatori utilizzano metodi qualitativi, come le interviste con le vittime e gli autori. Da alcuni anni, realizzano inoltre sempre più sondaggi per scoprire chi è coinvolto in qualità di vittima, spettatore e autore, con quale frequenza e perché. Rispetto a queste autovalutazioni, parzialmente distorte dalla tendenza a dare le risposte auspicate dalla società, gli approcci più recenti si prestano meglio per osservare direttamente il comportamento online di milioni di internauti, perché consentono di raccogliere in modo relativamente semplice ingenti quantità di dati su un lungo periodo di tempo (nel rispetto delle pertinenti direttive sulla protezione dei dati). Inoltre, oggi si analizza in che modo le infrastrutture tecniche e l'architettura delle piattaforme possono promuovere i discorsi d'odio. Infine, i ricercatori conducono esperimenti per capire meglio le conseguenze dei discorsi d'odio online, alternando situazioni con e senza discorsi d'odio per studiarne gli effetti sul pensiero e sull'azione dei partecipanti.

Sebbene molti di questi risultati possano essere trasposti alla Svizzera, non esistono studi che analizzino esplicitamente i discorsi d'odio razzisti online nel contesto svizzero. In Svizzera non sono ancora stati realizzati studi di ampio respiro sui discorsi d'odio online in generale e su quelli razzisti in particolare anche se diverse istituzioni si sono occupate di discorsi d'odio e di comportamenti aggressivi in Internet – per esempio l'Università di Zurigo, la Scuola universitaria di scienze applicate di Zurigo

²³⁸ P. es. Wolfe 2019.

²³⁹ Waqas et al. 2019: pag. 5.

(ZHAW), l'Alta scuola pedagogica vodese (HEP Vaud) o l'Università della Svizzera italiana (USI)²⁴⁰. Sarebbe prioritaria un'indagine sulla popolazione che identificasse le vittime, gli spettatori e gli autori dei discorsi d'odio. Si raccomanda di rilevare anche caratteristiche sociostrutturali, quali il livello d'istruzione, l'attività professionale o le reti sociali. Questo permetterebbe di identificare i gruppi a rischio e di mettere a punto contromisure specifiche. Sarebbe in particolare fondamentale raggiungere gli autori per determinare i loro moventi e le loro reti di contatti. Andrebbero privilegiati studi sul lungo periodo per monitorare gli sviluppi, le tendenze e le ripercussioni.

Se si applicano inoltre alla pratica le conoscenze scientifiche acquisite, si dovrebbe evitare un eccessivo amalgama tra discorso d'odio razzista online e altri tipi di discorso d'odio, che hanno in comune determinate caratteristiche, come la logica della semplificazione e della diffusione. L'amalgama è tuttavia problematico se le conoscenze sui discorsi d'odio in generale vengono trasposte direttamente ai discorsi d'odio di stampo razzista, nonostante queste due categorie di discorso possano differire per origine, effetti e contromisure. Per esempio, i discorsi d'odio razzisti e sessisti colpiscono gruppi diversi. Ognuno di questi gruppi ha una propria storia di discriminazione, che dovrebbe essere presa in considerazione quando si studiano le contromisure da adottare. Inoltre, il discorso d'odio può essere diretto o indiretto, palese o camuffato, isolato o ripetuto, e può essere sostenuto o no da istituzioni e personalità potenti²⁴¹. Tutto questo ha un impatto sulla portata del danno causato alle vittime.

In linea di principio, la ricerca può dare un importante contributo alla scelta informata e (finora ancora poco) basata su prove delle contromisure da adottare e alla loro attuazione. Il dialogo e lo scambio di conoscenze tra ricercatori e professionisti impegnati sul terreno possono essere utili in questo senso.

7.5 Società civile in Svizzera e all'estero

La società civile può intervenire in vari modi contro i discorsi d'odio razzisti online, in particolare laddove le misure previste dalla legge non sono efficaci o la moderazione e la rimozione operate dai gestori di social media non riescono a penetrare i contesti locali o laddove occorre mobilitare molte persone. Il presente sottocapitolo illustra alcuni approcci di progetti promossi dalla società civile che riguardano i discorsi d'odio online e, in alcuni casi, il razzismo online²⁴². Si tratta di approcci di vario genere (progetti istituzionalizzati, iniziative spontanee poco coordinate, organizzazioni specializzate o singoli esperti e strumenti). Dato il loro elevato numero, in questa sede sono menzionati soltanto quelli considerati particolarmente rilevanti. I promotori sono consultori, organizzazioni giovanili o ONG. Gli approcci sono suddivisi in quattro ambiti:

- prevenzione e sensibilizzazione;
- consulenza e segnalazione in casi concreti;
- monitoraggio;
- contronarrazione.

Spesso un approccio può coprire più di un ambito. Nella suddivisione qui proposta sono attribuiti all'ambito nel quale, secondo l'autrice del presente studio, assumono un ruolo di esempio. L'efficacia va relativizzata, dato che non esistono valutazioni sistematiche dei singoli progetti²⁴³ e le valutazioni interne hanno carattere autoreferenziale.

²⁴⁰ P. es. Catherina Blaya (LASALÉ, HEP Vaud), Dirk Baier (Istituto di criminologia e di prevenzione della criminalità, ZHAW), Lea Stahel (Istituto di sociologia, UZH), Elisabeth Stark (Seminario di filologia romanza, UZH), Katharina Lobinger e Eleonora Benecchi (Istituto di media e giornalismo e Istituto di tecnologie digitali per la comunicazione, USI), stato: maggio 2020.

²⁴¹ Delgado e Stefancic 2009: pag. 361.

²⁴² Nonostante le sovrapposizioni, in questa sede non vengono considerati gli approcci che concernono la (de)radicalizzazione digitale in generale o esplicitamente altre forme di discorso d'odio online (p. es. il discorso d'odio sessista). Alla stessa stregua, non vengono analizzati gli approcci che riguardano concernono il razzismo in generale o il razzismo offline, ma che non hanno un nesso diretto con la dimensione online.

²⁴³ Blaya 2019: pag. 163.

Oltre ai pochi studi di valutazione disponibili, sono stati considerati anche i resoconti raccolti nelle interviste. Le conoscenze qui riportate possono fornire soltanto un quadro approssimativo. Inoltre, l'efficacia degli approcci valutati non dev'essere considerata universalmente valida, in quanto dipende in ampia misura dal contesto specifico e dalle modalità concrete di realizzazione. Dato che il discorso d'odio online interessa vari settori della società (p. es. politica, diritto, industria tecnologica), molti professionisti sottolineano la necessità di promuovere lo sviluppo di reti di cooperazione tra responsabili di progetti, gestori di social media, politici, esperti di tecnica, autorità penali e altri soggetti. La piattaforma di networking tedesca [Das NETTZ](#) è un esempio di come queste reti possano essere realizzate. Anche l'[Institute for Strategic Dialogue](#) (Istituto per il dialogo strategico) e la fondazione tedesca [Amadeu Antonio](#) sono noti per le loro attività nel campo della ricerca, del lavoro sul terreno e della creazione di reti di cooperazione. Stando all'autrice, in Svizzera non vi sono approcci che mirano a contrastare in modo esplicito il razzismo online; la maggior parte di quelli menzionati in questa sede sono stati realizzati in altri Paesi.

7.5.1 Prevenzione e sensibilizzazione

Gli approcci di prevenzione e sensibilizzazione (cfr. tabella 1 per la Svizzera e tabella 2 per gli altri Paesi) hanno una prospettiva a lungo termine; mirano a promuovere la resilienza mediante un approccio educativo e a combattere sul nascere i discorsi d'odio online o, in alternativa, a contenerne gli effetti. Rappresentano soltanto uno degli elementi di un impegno molto più ampio per la costruzione di una società civile digitalmente matura. Tra le priorità vi sono:

- promuovere la competenza mediale (p. es. condizioni di comunicazione, architettura delle piattaforme);
- rafforzare una cultura del dibattito diversificata e la risoluzione costruttiva dei conflitti in Internet;
- sicurezza digitale e misure di precauzione (p. es. suggerimenti su come proteggersi, sviluppo della resilienza);
- promuovere la conoscenza delle definizioni, delle forme dei discorsi d'odio e del modo per riconoscerle, nonché delle cause, dei moventi, del retroterra, delle ripercussioni, della nocività e del modo di procedere di chi diffonde messaggi d'odio online;
- contestualizzare e/o integrare la questione dei diritti umani e della discriminazione;
- promuovere la riconoscibilità della responsabilità (la responsabilità è degli autori e non delle vittime);
- diritti e situazione giuridica;
- strategie di gestione per le vittime e gli spettatori (p. es. strategie di moderazione);
- mettere le vittime, potenziali o effettive, nelle condizioni di crearsi una rete digitale che consenta loro di trovare interlocutori, scambiare informazioni e ottenere supporto;
- supportare la pianificazione, la produzione e la promozione di progetti propri.

Gli strumenti offerti e i formati tecnici utilizzati per diffondere questi contenuti sono molteplici. Viene utilizzato materiale²⁴⁴ informativo reperibile direttamente sui siti web o scaricabile (p. es. volantini, opuscoli, relazioni e studi). L'offerta comprende anche consulenza individuale, workshop, corsi di formazione, presentazioni o webinar (online e offline), petizioni (per dare visibilità o creare pressione pubblica), campagne pubbliche, tavole rotonde e conferenze. I formati tecnici utilizzati sono testi, video (p. es. produzione di video da parte di giovani per blog), app, memi, canali YouTube, TED talks o podcast. I progetti si rivolgono a diversi gruppi target, quali il pubblico in generale, bambini e giovani, moltiplicatori (educatori, ONG ecc., ma anche utenti con una vasta rete di contatti online), genitori, giornalisti, e in parte anche aziende, mondo politico e piattaforme online. La scelta del gruppo target si riflette nel design, nel linguaggio utilizzato e nel coinvolgimento degli attori (p. es. maggiore interazione tra i giovani).

²⁴⁴ Per il materiale generale sulla competenza mediale dei bambini e dei giovani in Svizzera cfr. [Zischtig.ch](#) e per quello focalizzato sui discorsi d'odio cfr. il portale [Giovani e media](#); percorso: Giovani e media > Argomenti > Discriminazione e odio in rete.

Tabella 1 – Approcci di prevenzione e sensibilizzazione in Svizzera²⁴⁵

Progetto	(Principale) promotore del progetto / organizzazione	Lingua	Caratteristica distintiva
Stop Hate Speech	Frauentachverband; Alliance F	DE (FR)	Materiale informativo per il pubblico (in fase di pianificazione)
#NetzCourage #WirGegenHass	#NetzCourage	DE	Informazione ed educazione per contrastare i discorsi d'odio, tra cui un opuscolo informativo destinato ai politici
Hate Speech / Hassrede im Internet ²⁴⁶	Fondazione contro il razzismo e l'antisemitismo (GRA)	DE FR	Opuscolo sui discorsi d'odio razzisti e sulle misure per combatterli, soprattutto per le scuole
JASS gegen HASS	JASS	DE	Analisi della situazione (2017) e video contro i discorsi d'odio

²⁴⁵ All'autrice non risulta che vi siano approcci specifici in Ticino (eccetto la formazione impartita nelle scuole).

²⁴⁶ Percorso: www.gra.ch > Bildung > [Hate Speech](#).

Tabella 2 – Approcci di prevenzione e sensibilizzazione in altri Paesi

Progetto	(Principale) promotore del progetto / organizzazione	Paese	Caratteristica distintiva
Hate Speech: Hass Im Netz	Landesanstalt für Medien NRW; AJS NRW; Klicksafe	D	Manuale con informazioni per professionisti e genitori
Hate Speech ²⁴⁷	Deutsche Gesetzliche Unfallversicherung (DGUV)	D	Materiale didattico dettagliato da scaricare
Streitkultur 3.0	Berghof Foundation / Friedenspädagogik Tübingen	D	Rafforzare la cultura del conflitto attraverso laboratori digitali di dialogo
#NoHateNoFake	WERK 2 – Kulturfabrik Leipzig e.V.	D	Progetto di videoblog per le scuole
Werte leben - online	Juuuport e.V.	D	Mediazione di valori per la discussione digitale da parte di giovani «scout» con webinar
Hate Speech im Netz stoppen!	Campact	D	Petizione contro il discorso d'odio online con richieste alla giustizia
Tarik Tesfu	Tarik Tesfu	D	Influencer con un canale YouTube (che si posiziona contro i discorsi d'odio)
No Hate Speech Movement	Consiglio d'Europa, Neue deutsche Medienmacher	D / UE	Ampia e dettagliata offerta di informazioni elaborate per il web
ZARA	ZARA	A	Formazione pratica per adulti e giovani sul coraggio civile in generale e e su come argomentare contro il razzismo online in particolare
Rete nazionale per il contrasto ai discorsi e ai fenomeni d'odio	Vari, incl. Ufficio nazionale antidiscriminazioni razziali (UNAR)	I	Rete per promuovere e sostenere in particolare la ricerca, le campagne di sensibilizzazione e lo scambio di informazioni
CO.N.T.R.O	Ufficio nazionale antidiscriminazioni razziali (UNAR), in partenariato con l'Institute for Social Research (ISR)	I	Studio e ricerca, attività di sensibilizzazione e informazione
BRICKS – Against Hate Speech	Cospe, Zaffiria	I	Materiale didattico sui discorsi d'odio
#NOHATESPEECH Giornalisti e lettori contro i discorsi d'odio	Associazione Carta di Roma	I	Petizione contro i discorsi d'odio online rivolta a giornalisti e lettori
#SilenceHate – Giovani digitali contro il razzismo	Cospe, Zaffiria, Priscilla	I	Lotta contro la diffusione di discorsi d'odio attraverso l'educazione ai media dei giovani
No Hate Speech Movement, anche su Facebook	Dipartimento della gioventù e del servizio civile nazionale, Consiglio d'Europa	I / UE	Spot radiotelevisivi e campagne educative in particolare per i giovani, manuale didattico
Facing facts	Jewish contribution to an inclusive Europe (CEJI-A)	8 Paesi UE, incl. I	Corsi didattici online per diversi gruppi target (p. es. corpi di polizia, organizzazioni della società civile, decisori) in tedesco, francese, italiano e inglese
LICRA	Ligue Internationale contre le racisme et l'antisémitisme (LICRA)	F	Attenzione focalizzata sul razzismo in generale con tema prioritario «Combattere la haine sur Internet»
Mouvement contre le discours de haine	Non indicato, Consiglio d'Europa	F / EU	Campagne educative, in particolare per i giovani
SELMA	European Schoolnet, UE	UE	Ampio materiale pedagogico destinato ai giovani
Campaign Toolkid	Institute for Strategic Dialogue (ISD)	UK	Tool di supporto per la concezione di progetti

Sugli effetti degli approcci di prevenzione e di educazione si sa poco, anche perché gli effetti a lungo termine dell'azione educativa, in particolare sul comportamento delle persone sensibilizzate, sono difficilmente misurabili. Le valutazioni ciò nondimeno disponibili si riferiscono quasi esclusivamente alla competenza mediale dei giovani ed evidenziano effetti positivi. Oltre a trasmettere conoscenze e competenze sull'uso sicuro e consapevole di Internet, i progetti permettono per esempio ai partecipanti di acquisire maggiore sicurezza per la corretta valutazione dei contenuti di terzi pubblicati online

²⁴⁷ Percorso: www.dguv.de > Sekundarstufe I > Sucht- und Gewaltprävention > [Hate Speech](#).

(p. es. disinformazione) e per la creazione e la diffusione responsabile di contenuti propri²⁴⁸. Una meta-analisi di 51 progetti²⁴⁹ conferma l'impatto positivo sulla conoscenza dei media e (in minor misura) sugli atteggiamenti, l'autoefficacia o il comportamento. Tuttavia, si consiglia di allenare più assiduamente il comportamento, anziché limitarsi a trasmettere conoscenze. Infine, gli studi suggeriscono²⁵⁰ che i discorsi d'odio online possono essere ridotti se gli utenti sono consapevoli delle conseguenze penali e degli effetti negativi che essi possono comportare.

Per garantire l'efficacia dei progetti, diversi esperti²⁵¹ consigliano di:

- adeguare al gruppo target gli interventi in termini di abitudini mediali, realtà della vita digitale, bisogni digitali, livello di difficoltà e coinvolgimento; soprattutto per i giovani: equilibrio tra contenuti didattici ed esercitazioni ludiche;
- focalizzarsi sui gruppi vulnerabili, in particolare sulle vittime e sugli odiatori probabili e/o effettivi, anziché educare le persone già sensibilizzate;
- prevedere un procedimento didattico in tre fasi: 1. sensibilizzare e informare, 2. analizzare e riflettere, 3. applicare; quest'approccio permette l'apprendimento di nuove conoscenze, lo scambio di esperienze e l'attuazione coraggiosa di strategie;
- rafforzare le competenze dei partecipanti in modo che adottino decisioni autonome ed etico-morali; evitare un approccio moralista prescrivendo opinioni e atteggiamenti «corretti»;
- in relazione alle strategie di comportamento, focalizzarsi sui problemi specifici, incluso il training comportamentale;
- utilizzare riferimenti socioculturali popolari del mondo di Internet (per collegarsi all'universo giovanile);
- predisporre materiale di accompagnamento stampato adatto all'età del gruppo target;
- offrire una formazione continua ai moltiplicatori (p. es. insegnanti) per garantire una diffusione competente dei contenuti.

7.5.2 Segnalazione e consulenza

Se in situazioni concrete sono vittime di attacchi razzisti o se assistono ad attacchi di questo tipo, gli internauti possono segnalarlo e/o chiedere consiglio (per i progetti specifici, cfr. tabella 3 per la Svizzera e tabella 4 per gli altri Paesi). In linea di massima, i messaggi possono essere segnalati direttamente alle piattaforme social (se vi è il sospetto che violino le regole di utilizzo). In alternativa, si possono contattare i servizi che mettono a disposizione altri tool di segnalazione. Questi strumenti permettono di verificare la rilevanza penale dei post segnalati, che vengono inoltrati ai gestori della piattaforma per chiederne la rimozione e/o alle autorità penali per avviare un procedimento penale (quest'ultima possibilità vale in particolare per la Germania). In Svizzera, GRA e #Netzcourage segnalano regolarmente contenuti d'odio ai gestori di social media. #Netzcourage si avvale di una community online che, coordinando segnalazioni in massa, aumenta la probabilità che i messaggi d'odio vengano effettivamente rimossi. I tool dei servizi di segnalazione più importanti hanno impostazioni e caratteristiche diverse, per esempio:

- alcuni permettono segnalazioni soltanto tramite siti Internet, altri anche tramite app;
- il livello di dettaglio delle informazioni sulle modalità di segnalazione (p. es. cosa può/deve essere segnalato e cosa no, vantaggi della segnalazione) varia da tool a tool;
- la quantità di informazioni che l'utente deve/può fornire (p. es. screenshot, link) varia da tool a tool;
- in alcuni la segnalazione può essere anonima, in altri no;
- in alcuni chi segnala può ricevere un riscontro sulla rimozione e/o sulla procedura, in altri no;

²⁴⁸ P. es. Gatewood e Boyer 2019.

²⁴⁹ Jeong, Cho e Hwang 2012.

²⁵⁰ Aizenkot e Kashy-Rosenbaum 2018.

²⁵¹ P. es. Blaya 2018; Blaya 2015; Dittrich et al. 2020; Gatewood e Boyer 2019; Wachs et al. 2019.

- alcuni sono utilizzati per la pubblicazione di rapporti o statistiche basati sulle segnalazioni operate (numero, tipo, contenuti), altri no.

La consulenza offerta è di tipo psicologico, sociale e/o legale. Non tutte le consulenze comprendono tutti questi aspetti, anche se ciò sarebbe raccomandabile per garantire alle persone direttamente interessate un sostegno completo. Generalmente, le consulenze sono gratuite e sono offerte attraverso canali di comunicazione anonimi e non anonimi (p. es. telefono, chat, contatto personale). Alcune associazioni aiutano le vittime anche a sporgere denuncia o ad avviare un'azione presso il procuratore pubblico²⁵². Le vittime possono ricevere sostegno legale e psicosociale per l'udienza di conciliazione o l'udienza in tribunale, come pure un aiuto immediato per la gestione del profilo social preso di mira (in Svizzera principalmente #NetzCourage). In alcuni casi le vittime possono beneficiare di un aiuto finanziario per le spese processuali, anche se stando all'autrice dello studio attualmente l'unica organizzazione a farlo è HateAid in Germania. Infine, vengono messi a disposizione terapisti esterni che aiutano le vittime nel processo di elaborazione, allo scopo, tra l'altro, di evitare che si isolino e si ritirino dai social media.

Tabella 3 – Segnalazione e consulenza in Svizzera²⁵³

Progetto	(Principale) promotore del progetto / organizzazione	Lingua	Caratteristica distintiva
Tool di segnalazione ²⁵⁴	Fondazione contro il razzismo e l'antisemitismo (GRA)	DE FR	Tool di segnalazione per il razzismo online, in alcuni casi consulenza
#NetzCourage	#NetzCourage	DE	Accompagnamento (consulenza sulle possibilità di reazione, segnalazione, supporto nella segnalazione ecc.) In alcuni casi presa di contatto con gli autori

Tabella 4 – Segnalazione e consulenza in altri Paesi

Progetto	(Principale) promotore del progetto / organizzazione	Paese	Caratteristica distintiva
Hass melden	Reconquista Internet	D	Possibilità di segnalazione tramite app; valutazione della rilevanza penale, inoltro alle autorità penali
Eco Beschwerdestelle	Eco	D	Possibilità di segnalazione con inoltro alle autorità e alle piattaforme social; forte coinvolgimento di giuristi; pubblicazione di rapporti annuali
HateAid	Campact e.V., Fearless Democracy e.V. (Anna-Lena von Hodenberg)	D	Consulenza e supporto durante tutto il procedimento (in particolare consulenza e assunzione delle spese processuali grazie a un sistema di donazioni)
Ban hate	Antidiskriminierungsstelle Steiermark	A	App di segnalazione con possibilità di consultazione dello stato di avanzamento
Hass im Netz melden ²⁵⁵	ZARA	A	Consulenza e segnalazione specifiche per il razzismo online
Segnala il razzismo online	Regione Abruzzo (soltanto Italia)	I, UE	Rapporto con informazioni e proposte pratiche su come procedere
INACH cyber hate complaints base	International Network Against Cyber Hate (INACH)	22 Paesi, incl. F e I	Elaborazione delle segnalazioni in rapporti internazionali
Signaler un fait de racisme / d'antisémitisme	Ligue Internationale Contre le Racisme et l'Antisémitisme (LICRA)	F	Sito Internet per segnalare episodi di razzismo offline e online

²⁵² In Germania si è osservato che gli organi competenti non sono sufficientemente consapevoli del problema. Per ovviarvi, sono state formulate diverse richieste all'attenzione delle autorità p. es. da parte di Neue deutsche Medienmacher & No Hate Speech Movement Deutschland 2019. Tra le misure auspicate figurano agenti specializzati nella lotta ai discorsi d'odio, corsi di formazione per i corpi di polizia, sensibilizzazione di corpi di polizia, tribunali e pubblici ministeri e possibilità di segnalazione online.

²⁵³ All'autrice non risulta che vi siano progetti di questo tipo in Ticino.

²⁵⁴ Percorso: www.gra.ch > Bildung > [Hate Speech](#).

²⁵⁵ Percorso: www.zara.or.at > Beratung > [Hass im Netz anonym melden](#).

Réagis au racisme	equal.brussels	B	Informazioni sul razzismo e possibilità di segnalare episodi di razzismo offline e online (in francese, inglese e olandese)
No hate speech youth campaign ²⁵⁶	Council of Europe Youth Department	UE	Informazioni dettagliate sulle procedure di segnalazione ai principali gestori di social media

Si parte dal presupposto che la notorietà e la facilità d'accesso determinino se e con quale frequenza gli internauti si avvalgono dei tool di segnalazione e dell'offerta di consulenza. In linea di principio sembra che gli utenti segnalino raramente contenuti d'odio²⁵⁷ perché le vittime non hanno fiducia nella polizia, accettano la violenza come un dato di fatto e temono ritorsioni o la violazione della loro privacy o le barriere linguistiche²⁵⁸. Tuttavia, le persone tendono tanto più a segnalare i discorsi d'odio quanto meno discontinuità mediali devono affrontare, ossia quanto meno devono passare da un media all'altro. Una soluzione ottimale è la segnalazione diretta alle piattaforme sulle quali i discorsi d'odio sono stati pubblicati. Non si può però presumere che le persone siano in grado di valutare correttamente la rilevanza penale dei messaggi e quindi l'opportunità di procedere alla segnalazione. Nelle sue statistiche, l'organo di reclamo tedesco Eco²⁵⁹ constata che nel caso del razzismo e dei discorsi d'odio soltanto l'otto per cento delle segnalazioni erano effettivamente rilevanti dal punto di vista penale. Secondo gli autori, ciò dimostra quanto sia importante che le persone chiamate a valutare i contenuti pubblicati dispongano di competenze giuridiche per evitare che dichiarazioni consentite vengano rimosse soltanto perché indesiderate²⁶⁰. Va infine ricordato che la segnalazione può sfociare in un procedimento penale lungo, complesso e costoso e che pertanto occorre valutare se vi sono le risorse finanziarie necessarie. [HateAid](#) giustifica l'assunzione delle spese processuali adducendo che la lotta contro la violenza in rete non dovrebbe dipendere dalle disponibilità finanziarie delle vittime. Se gli organi competenti riescono a ridurre gli ostacoli a cui si è accennato, questo potrà teoricamente incoraggiare le segnalazioni. Anche sul terreno si notano grosse differenze nel numero di segnalazioni: la fondazione svizzera GRA riceve²⁶¹ in media tre segnalazioni alla settimana (soltanto dalla Svizzera), mentre in Germania [Hass melden](#) ne riceve 500 (finora sono stati denunciati complessivamente 16 000 post).

La pubblicazione di statistiche delle segnalazioni e delle condanne potrebbe incrementare la visibilità, sensibilizzare e produrre un effetto dissuasivo, in modo che Internet sia visto sempre meno come uno spazio senza legge. Anche le udienze di conciliazione, in cui vittima e autore siedono l'uno di fronte all'altro (che in Svizzera sono organizzate da #Netzcourage) potrebbe contrastare determinati meccanismi quali la distanza fisica, la percezione di anonimato e l'assenza di sanzioni (cfr. cap. 5.4.1) e ridurre la disinibizione digitale.

7.5.3 Monitoraggio

L'obiettivo fondamentale degli approcci di monitoraggio (tabella 5 per la Svizzera e tabella 6 per gli altri Paesi) è di seguire in modo sistematico i discorsi d'odio online nel loro contesto locale e di descriverli. La scienza, la società civile e i think tank hanno avviato diversi progetti volti soprattutto a rilevare l'evoluzione a lungo termine. I risultati informano gli attori della società civile sui principali gruppi che diffondono discorsi d'odio a livello locale permettendo loro di adattare la strategia delle campagne di contronarrazione o forniscono alle autorità elementi utili per identificare gruppi di autori o di vittime ed eventualmente prevenire crimini d'odio. I progetti possono essere suddivisi secondo l'approccio in progetti quantitativi (cfr. punti 1-3) e progetti qualitativi (punto 4)²⁶². I due approcci sono complementari: i progetti quantitativi possono fornire una panoramica (mappatura) della frequenza e dei tipi di discorso

²⁵⁶ Percorso: www.coe.int > Democrazia > Movimento contro il discorso dell'odio > Taking Action > [Reporting Hate Speech](#).

²⁵⁷ Un sondaggio dell'autorità di sorveglianza dei media del Land Renania Settentrionale-Vestfalia (2018) mostra che, tra gli utenti che hanno dichiarato di aver osservato discorsi d'odio in rete, soltanto una minoranza ha segnalato il post al gestore di social media (25 %) o alla polizia (1 %).

²⁵⁸ Cerase, D'Angelo e Santoro 2015: pag. 4.

²⁵⁹ Eco – Verband der Internetwirtschaft 2019.

²⁶⁰ Eco – Verband der Internetwirtschaft 2019: pag. 25.

²⁶¹ Intervista con Dominik Pugatsch (Fondazione contro il razzismo e l'antisemitismo, GRA), marzo 2020.

²⁶² Per una panoramica, cfr. Lucas 2014.

d'odio online in un contesto più ampio, a condizione che la raccolta e l'analisi dei dati siano sistematiche e che i dati siano comparabili (il che è possibile soltanto se le segnalazioni sono raccolte in modo sistematico), quelli qualitativi forniscono indicazioni più approfondite sulla natura e il funzionamento dei discorsi d'odio.

1. *Monitoraggio continuo in tempo reale*: una grande quantità di contenuti online viene analizzata con metodi automatici, quali il trattamento informatico del linguaggio naturale, l'apprendimento automatico e l'analisi dei big data²⁶³. Determinati spazi online (p. es. social media, giornali online) e piattaforme virtuali (con o senza moderazione) vengono costantemente scandagliati per individuare discorsi d'odio. Approcci di questo tipo offrono un sistema di allarme precoce e permettono di rilevare tendenze e di reagire agli eventi già nella loro fase nascente. Sono tuttavia onerosi e quindi vengono realizzati raramente (p. es. coinvolgimento di molti volontari).
2. *Monitoraggio retrospettivo*: per questo monitoraggio si utilizzano generalmente metodi automatici per analizzare grandi quantità di contenuti online. I contenuti vengono selezionati e analizzati in un lasso di tempo relativamente breve.
3. *Raccolta di contenuti segnalati*: i contenuti online inoltrati ai sistemi di segnalazione vengono raccolti in modo continuativo (in diretta) o a intervalli regolari. Le segnalazioni degli utenti sono talvolta integrate da episodi raccolti con ricerche interne. Qui il volume dei dati è limitato.
4. *Analisi qualitative*: queste analisi permettono di esaminare la qualità di un numero limitato di contenuti o anche di intere piattaforme in modo da approfondire le conoscenze sugli argomenti, i moventi o le ideologie dei diffusori d'odio.

Tabella 5 – Approcci di monitoraggio in Svizzera

Progetto	(Principale) promotore del progetto / organizzazione	Lingua	Caratteristica distintiva
Stop Hate Speech	Fraendachverband Alliance F	DE (FR)	Monitoraggio algoritmico dei discorsi d'odio sulle piattaforme online dei giornali e nei social media in tutta la Svizzera con l'ausilio di volontari (in fase di pianificazione)
Chronologie «Rassismus in der Schweiz»	Fondazione contro il razzismo e l'antisemitismo (GRA)	DE FR	Dal 1992 registra tutti gli episodi resi pubblici che poggiano su moventi razzisti o xenofobi (casi offline e online)
Antisemitismusbericht	GRA e Federazione svizzera delle comunità israelite (FSCI)	DE FR	Rapporto annuale sugli episodi di antisemitismo segnalati dal pubblico o raccolti attivamente; elenco separato degli episodi online
Rapports Antisémitisme	Coordination Intercommunautaire Contre l'Antisémitisme et la Diffamation (CICAD)	FR	Rapporto annuale sugli episodi di antisemitismo in Svizzera romanda (casi offline e online)
Episodi di razzismo trattati nell'attività di consulenza	Associazione Humanrights.ch, Commissione federale contro il razzismo (CFR)	DE FR IT	Rapporto annuale nazionale sugli episodi di razzismo trattati nell'attività di consulenza; elenco separato degli episodi online

²⁶³ Cfr. in proposito Fortuna e Nunes 2018.

Tabella 6 – Approcci di monitoraggio in altri Paesi

Progetto	(Principale) promotore del progetto / organizzazione	Paese	Caratteristica distintiva
Debate//de:hate ²⁶⁴	Amadeu Antonio Stiftung	D	Rapporti di monitoraggio principalmente qualitativi con casi concreti
Mapping Hate (strumento)	Institute for Strategic Dialogue (ISD)	UK, USA	Strumento di monitoraggio metodologicamente solido e scientificamente fondato con geolocalizzazione; identificazione dei moventi, dei temi e delle figure chiave, raggruppamento di gruppi di odiatori; utilizzabile in molte lingue ²⁶⁵
Task force contro i discorsi d'odio	Amnesty International	I	Osservazione di e reazione a discorsi d'odio online; rapporto Il barometro dell'odio sui discorsi d'odio online di politici
LIGHT ON	Regione Abruzzo, ISIG, Progetti sociali (soltanto Italia)	I, UE	Documentazione visuale di casi di discorsi d'odio offline e online
Cartographie de la Haine en Ligne: Tour d'horizon du discours haineux en France	Institut pour le dialogue stratégique (ISD) e Facebook	F	Rapporto sulle forme di discorso d'odio online in Francia
hatemeter	eCrime, Università di Trento, Facoltà di giurisprudenza	I, F, UK	Monitoraggio automatico in diretta in tutta l'UE dei discorsi d'odio anti-musulmani inteso a identificare le tendenze dei discorsi d'odio; utilizzo dell'intelligenza artificiale a supporto della contronarrazione
Research – Report – Remove: Countering Cyber Hate Phenomena	International Network Against Cyber Hate (INACH)	6 Paesi UE, incl. D e F	Rapporti semestrali sui tipi di discorso d'odio specifici a singoli Paesi, sugli eventi scatenanti, sulle tendenze e sulle piattaforme social rilevanti basati su contenuti segnalati e raccolti in modo meno sistematico
Scan	Ligue Internationale Contre le Racisme et l'Antisémitisme (LICRA)	10 Paesi UE, incl. D, F e I	Rapporti e workshop su strumenti e conoscenze per il monitoraggio

Gli approcci di monitoraggio servono a misurare efficacemente la presenza dei discorsi d'odio online e dei loro modelli a lungo termine nelle e tra le società (fungendo per così dire da barometro delle tensioni sociali). Molti rapporti di monitoraggio pubblicati finora sono ricchi di informazioni su episodi specifici e cercano di cogliere eventuali sviluppi. L'uso di algoritmi per il riconoscimento aiuta anche a categorizzare in breve tempo enormi quantità di informazioni che non possono essere trattate manualmente. Nonostante i vantaggi rispetto, per esempio, ai dati dei sondaggi, i dati «comportamentali» non consentono di quantificare in modo affidabile i discorsi d'odio online in un determinato Paese o di prevederne l'evoluzione a lungo termine. Questo è dovuto al fatto che il monitoraggio riguarda spesso soltanto singole piattaforme online e non l'intera rete. Inoltre, poiché generalmente gli algoritmi e le strategie di moderazione delle piattaforme non sono trasparenti, non si possono determinare i motivi all'origine di eventuali cambiamenti osservati in rete: si tratta di un cambiamento effettivo del comportamento degli utenti o di una modifica delle regole di rimozione dei contenuti sulle piattaforme? Pertanto, in una certa misura, i ricercatori e i servizi coinvolti devono anche interpretare e contestualizzare i fattori che stanno alla base dei discorsi d'odio²⁶⁶. Un dilemma analogo si pone per l'interpretazione dei contenuti segnalati, in quanto sono rappresentativi soltanto di ciò che è stato segnalato ed è probabile che rappresentino unicamente la punta dell'iceberg.

L'attuazione del monitoraggio è una sfida per le organizzazioni della società civile²⁶⁷: spesso l'accesso ai dati online non è garantito e i software per il riconoscimento automatico sono molto costosi²⁶⁸. La programmazione di algoritmi richiede competenze linguistiche, conoscenze informatiche e risorse umane sufficienti per garantire che gli algoritmi riconoscano in modo affidabile i discorsi d'odio in

²⁶⁴ Percorso: www.amadeu-antonio-stiftung.de > Projekt > de:hate > [Monitoring und Analyse](#).

²⁶⁵ Intervista con Jonathan Birdwell (Institute of Strategic Dialogue), aprile 2020.

²⁶⁶ Schmidt e Wiegand 2017.

²⁶⁷ Cfr. in proposito sCAN 2018b.

²⁶⁸ sCAN 2018b: pag. 18. In Svizzera esistono soluzioni assicurative contro gli attacchi online per singoli e imprese, che propongono il monitoraggio dei social media, ammoniscono gli autori e rimuovono i contenuti (p. es. [Silencio](#)). Tuttavia, visti l'impostazione e i costi, quest'approccio si presta meno per la lotta contro il razzismo.

un'ingente massa di dati. Considerate la natura complessa del discorso d'odio e la soggettività della percezione di ciò che è o non è un discorso d'odio²⁶⁹, questo processo non deve essere sottovalutato. Si dovrebbe anche tener conto del contesto locale e della lingua. Se questi fattori ambientali vengono ignorati, gli algoritmi possono facilmente indurre a errori di valutazione. Gli algoritmi più complessi non cercano soltanto singole parole, ma anche metadati (p. es. geolocalizzazione). Per scongiurare errori di classificazione è comunque necessaria un'ulteriore validazione qualitativa – manuale – dei risultati. Per poter tener conto dell'evoluzione delle espressioni d'odio nel tempo (cfr. cap. 5.4.2), gli algoritmi devono essere flessibili e modificabili sul lungo periodo. Infine, nella raccolta dei dati vanno considerate le direttive sulla protezione dei dati e sulla privacy²⁷⁰. Il procedimento per la definizione del fenomeno da analizzare, dei dati, dell'algoritmo e dell'interpretazione dei risultati deve essere comunicato in modo trasparente. Se tali requisiti sono soddisfatti, i rapporti di monitoraggio forniscono significative indicazioni specifiche a singoli Paesi sugli eventi scatenanti, le tendenze e i tipi di discorso d'odio e i gruppi che li diffondono.

7.5.4 Contronarrazione

La regola di base «Don't feed the trolls»²⁷¹ (letteralmente: non nutrire i troll, ovvero non rispondere alle provocazioni) non si applica alla contronarrazione. Per contronarrazione (*counter speech*) si intendono testi, immagini o video che mirano a contrastare il discorso d'odio apportando al dibattito un tono positivo. L'idea è di rafforzare norme quali il rispetto e l'obiettività mostrando solidarietà verso le vittime. Si tratta di affrontare chi diffonde discorsi d'odio: nel caso ideale, il confronto dovrebbe indurli a un ripensamento ma anche – cosa ancora più importante – ridurre il loro impatto sugli spettatori. La contronarrazione segnala al pubblico che il discorso d'odio non è tollerato e non rappresenta l'opinione dominante. Si tratta di confutare l'immagine di «maggioranza silenziosa» che i diffusori di discorsi d'odio online cercano generalmente di trasmettere di sé. Questi obiettivi non possono essere raggiunti semplicemente rimuovendo i contenuti.

Gli approcci rilevanti (tabella 7 per la Svizzera e tabella 8 per gli altri Paesi) si differenziano in particolare per quanto riguarda i gruppi target, le piattaforme e le modalità concrete di attuazione²⁷². I gruppi target sono selezionati in base alla distinzione tra prevenzione primaria (minori e adulti in contesti educativi), prevenzione secondaria (gruppi vulnerabili) o prevenzione terziaria (diffusori di discorsi d'odio online affrontati direttamente), ma possono coesistere all'interno delle singole iniziative. La contronarrazione è diffusa in particolare da gestori di community e internauti volontari. Questi ultimi si riuniscono in gruppi d'azione ed esprimono la loro solidarietà alle vittime per sostenerle ed evitare che si ritirino dal mondo digitale. I gruppi d'azione sono community permanenti o temporanee e focalizzano i loro interventi su uno o più tipi di discorso d'odio (razzista) online. Ci sono anche diversi stili di intervento: conflittuale o comunitario, con argomentazioni²⁷³ logiche e fattuali, con satira e umorismo, con commenti, «mi piace» e condivisioni²⁷⁴. I memi (immagini, cartoni animati, GIF, proverbi o aforismi rielaborati in chiave creativa con un contromessaggio) sono particolarmente apprezzati. Gli approcci di contronarrazione si focalizzano su vari aspetti, tra cui:

- offerte di consulenza e di formazione per l'adeguamento delle strategie di lotta ai diversi tipi di discorso d'odio, elaborazione di materiale e attuazione di contronarrazioni;
- approntamento di materiale innovativo per la contronarrazione; può essere adattato in base ai tipi di discorso d'odio o alle strategie retoriche (p. es. «benaltrismo»);
- supporto in rete per le community di contronarrazione, p. es. gruppo Facebook;
- sviluppo di una contronarrazione propria;

²⁶⁹ Salminen et al. 2018.

²⁷⁰ sCAN 2018b: pag. 21 segg.

²⁷¹ Lorentzen, TEDx Talk 2017. [Don't feed the trolls - Fight them](#), pubblicato su Youtube il 7 gennaio 2020.

²⁷² Tuck e Silverman 2016.

²⁷³ Per l'applicazione della contronarrazione esiste un data set multilingue compilato da esperti composto da 4078 coppie di parole di odio e contro-argomentazioni; cfr. Chung et al. 2019.

²⁷⁴ P. es. Ziegele et al. 2019.

- *naming and shaming* ovvero rendere di dominio pubblico le azioni di personaggi pubblici (ma non solo) identificati come diffusori di discorsi d'odio o troll (p. es. [Troll del mese](#)²⁷⁵);
- progetti artistici indiretti quali campagne innovative di raccolta fondi ([Hass hilft](#)), opere d'arte interattive ([Tools of Subversion](#)²⁷⁶) o eventi satirici ([Hate Poetry](#))²⁷⁷;
- approcci non organizzati, nell'ambito dei quali utenti di Internet si oppongono spontaneamente e in funzione della situazione a discorsi d'odio; p. es. dopo che in rete si è cominciato a incorniciare nomi ebraici o presunti tali con parentesi triple (cfr. cap. 5.4.2) molti utenti non ebrei si sono mostrati solidali e hanno aggiunto le parentesi anche al loro nome utente in modo da rendere più difficile la ricerca di contenuti antisemiti²⁷⁸; vanno inoltre menzionate le contronarrazioni degli utenti online: molti si oppongono ai discorsi d'odio non come parte di un gruppo, ma a titolo individuale²⁷⁹.

Le modalità di attuazione sono illustrate in base a due noti progetti di contronarrazione:

Il tool [Seriously](#)²⁸⁰ sviluppato dal think tank francese Renaissance Numérique combina l'approccio scientifico con quello educativo. Da un lato, i responsabili progettano e raccolgono materiale che può servire per la contronarrazione (p. es. statistiche, fatti, citazioni) e lo aggiornano continuamente per descrivere i dibattiti in corso nella società. Prima che venga messo a disposizione sulla piattaforma, una commissione di esperti indipendenti ne verifica l'idoneità e la qualità. Dall'altro, il tool ha anche carattere educativo: se necessario, gli interessati possono essere supportati nell'elaborazione dell'argomentazione critica e nello sviluppo di strategie di *de-escalation* per intervenire efficacemente nei dibattiti sui discorsi d'odio.

L'autoproclamato gruppo d'azione apartitico [#ichbinhier](#)²⁸¹ si concentra invece sulla messa in rete di contronarratori volontari, che attualmente sono 45 000 (maggio 2020). Ogni giorno gli organizzatori pubblicano nel loro gruppo Facebook link a nuovi dibattiti su canali social molto seguiti, a partire dai quali i membri organizzano spontaneamente una contronarrazione. L'obiettivo è di permettere al pubblico di formarsi un'opinione equilibrata e obiettiva. I membri del gruppo si sostengono a vicenda attribuendo «mi piace» alle rispettive contronarrazioni. Questo fa sì che l'algoritmo li sposti verso l'alto nella rubrica dei commenti, visto che nella classifica per rilevanza gli algoritmi attribuiscono un miglior posizionamento alle interazioni che generano maggior traffico. Questo dà maggiore visibilità alla contronarrazione e, nel caso ideale, mette in secondo piano il discorso d'odio.

Tabella 7 – Approcci di contronarrazione in Svizzera²⁸²

Progetto	(Principale) promotore del progetto / organizzazione	Lingua	Caratteristica distintiva
Stop Hate Speech	Frauentachverband Alliance F	DE (FR)	Contronarrazione coordinata da cittadini (in fase di pianificazione)
#NetzCourage	#NetzCourage	DE	Contronarrazione coordinata e azioni di segnalazione su Facebook
Meldezentrale für Eidgenossen	10 fondatori anonimi	DE	Contronarrazione coordinata su Facebook dal 2017 con pubblicazione dei casi più importanti sul blog

²⁷⁵ Percorso: www.getthetrollsout.org > What we do > [Troll of the month](#).

²⁷⁶ Percorso: www.goldextra.com > Projekte > [Tools of subversion](#).

²⁷⁷ Facebook: Hate Poetry.

²⁷⁸ Gunaratna, CBS News, 10 giugno 2016: [Neo-Nazis tag \(\(\(Jews\)\)\) on Twitter as hate speech, politics collide](#).

²⁷⁹ Jones e Benesch 2019.

²⁸⁰ Le informazioni sul tool Seriously provengono in particolare da un'intervista con Claire Pershan (Renaissance Numérique), aprile 2020.

²⁸¹ Facebook: #ichbinhier.

²⁸² All'autrice non risulta che vi siano progetti di questo tipo in Ticino.

Tabella 8 – Approcci di contronarrazione in altri Paesi

Progetto	(Principale) promotore del progetto / organizzazione	Paese	Caratteristica distintiva
#ichbinhier e Hatecontrol Tool	ichbinhier e.V.	D / internazionale	Organizza contronarrazioni (v. sopra); tool per la chiusura temporanea delle rubriche di commento su Facebook in caso di discorsi d'odio
No Hate Speech Movement	Neue deutsche Medienmacher, Consiglio d'Europa	D	Approntamento di materiale di contronarrazione (soprattutto in chiave umoristica) contro vari tipi di discorso d'odio razzista
Democratic Meme Factory	La Red	D	Approccio creativo alla produzione di memi
Love Storm	Bund für Soziale Verteidigung e.V.	D	Formazione pratica e azioni della community, focus sulla solidarietà alle vittime
Online Civil Courage Initiative (OCCI) Rapport d'Observation de l'OCCI	Facebook (fondatore)	D, UK, F	Sostegno a ONG e attività; sviluppo di materiale di contronarrazione
#jesuislà	Xavier Brandao (fondatore)	F	Organizza contronarrazioni; equivalente francofono di #ichbinhier
Seriously	Renaissance Numérique	F	Materiale di contronarrazione verificato su basi scientifiche e approccio pedagogico (v. sopra)
ZARA	ZARA	A	Diffusione di conoscenze e workshop sulla contronarrazione per giovani e adulti
Youth Civil Activism Network ²⁸³	Institute for Strategic Dialogue (ISD)	Globale	Rete giovanile globale per la contronarrazione

Quanto è efficace la contronarrazione? Si può dire se non altro che il suo principio di base è chiaro: soprattutto nel caso di contenuti che non sono né illegali né contrari agli standard in uso nelle community delle piattaforme, ma che vengono comunque percepiti da molti come problematici, la contronarrazione può assumere la funzione normativa della società. In particolare nel caso di attacchi non specifici riferiti a un gruppo, in cui non vi sono individui direttamente colpiti che si sentono costretti a reagire, la contronarrazione potrebbe essere un modo per contrastare il fenomeno della diffusione della responsabilità (effetto spettatore). Se sulle piattaforme online non si lascia spazio ai diffusori di discorsi d'odio, si crea la consapevolezza che è possibile inibire l'odio online («gli odiatori sono soltanto una minoranza degli utenti della rete», «i discorsi d'odio vengono sanzionati dalla rete»). Questo permette di dare un segnale al quale gli spettatori tendono a conformarsi. Questo controllo sociale informale è stato confermato nell'ambito di esperimenti²⁸⁴. L'effetto è più forte se a «sanzionare» è un membro del gruppo, a maggior ragione se gode di grande considerazione o fiducia²⁸⁵. Questo parla a favore del coinvolgimento di personalità pubbliche. L'efficacia della contronarrazione è stata dimostrata anche dall'esempio di #ichbinhier²⁸⁶: la contronarrazione ha reso la discussione meno aggressiva e ha motivato i non membri di #ichbinhier a uscire dal loro ruolo passivo e a reagire.

²⁸³ Percorso: www.isdglobal.org > Programmes > Grassroots Networks > [Youth Civil Activism Network](#) (YouthCAN).

²⁸⁴ Álvarez-Benjumea e Winter 2018.

²⁸⁵ Munger 2017.

²⁸⁶ P. es. Ziegele et al. 2019.

L'efficacia degli approcci di contronarrazione dipende in larga misura dal contesto e dalle modalità concrete in cui vengono attuati. Secondo gli esperti²⁸⁷, sono tendenzialmente più efficaci quando:

- il materiale e il tool vengono elaborati con il gruppo target e testati su di esso; il gruppo target varia a seconda dell'orientamento del progetto e comprende generalmente il pubblico (ancora) passivo e i potenziali contronarratori;
- le azioni di contronarrazione sono adattate alla community (online), alla sua cultura, ai suoi valori e alla piattaforma; nel caso ideale, i diffusori d'odio online «pentiti» vengono coinvolti visto che conoscono bene il fenomeno e sono considerati un gruppo target a sé stante;
- il pubblico online è mediamente numeroso, inattivo e indeciso;
- la contronarrazione è costruttiva e civile;
- vengono coinvolti contronarratori molto seguiti in rete;
- si tiene conto della disponibilità individuale degli utenti a fare contronarrazione attiva e regolare (questo dipende dal background sociale e dalle esperienze individuali);
- almeno uno zoccolo duro di contronarratori può impegnarsi per un certo periodo di tempo (ovviamente su base volontaria);
- il materiale di contronarrazione è costantemente adeguato ai dibattiti in corso.

Allo stesso tempo non vanno ignorati i seguenti rischi:

- rischi dovuti all'esposizione mediale in prima persona; in questo caso devono essere pianificate misure di protezione;
- effetto controproducente dell'approccio umoristico, che può inizialmente sembrare promettente, ma che può essere percepito come condiscendente o arrogante;
- rischio di infiltrazione tramite falsi profili (una soluzione può essere la conoscenza personale tra i membri);
- elevato investimento di tempo per la produzione (costante) di materiale;
- rischio di legittimare il discorso d'odio attraverso il riferimento diretto a esso;
- difficoltà di «convertire» i diffusori d'odio ideologicamente convinti.

Per quantificare il successo delle campagne di contronarrazione sono rilevanti i seguenti criteri²⁸⁸: attenzione (p. es. portata digitale, numero di tweet, visualizzazioni video ecc.), interazione (p. es. condivisione della campagna con gli utenti tramite messaggi) ed efficacia (p. es. cambiamento degli atteggiamenti e dei comportamenti degli utenti). In linea di massima si osserva tuttavia che soltanto pochissimi approcci hanno una presenza significativa sui social media o seguono una strategia ben precisa. Ciò nonostante, è possibile raggiungere un'ampia diffusione con costi modesti²⁸⁹: un'iniziativa di contronarrazione su Facebook dal Regno Unito ha raggiunto oltre 670 000 utenti con un budget di 3750 dollari²⁹⁰. Si possono inoltre contenere i costi facendo ricorso alle community online già attive.

²⁸⁷ P. es. Laubenstein e Urban 2018; Reynolds e Tuck 2016; Schieb e Preuss 2016.

²⁸⁸ Silverman et al. 2016: pag. 12.

²⁸⁹ Laubenstein e Urban 2018

²⁹⁰ Silverman et al. 2016: pag. 7.

8 SERVIZI SVIZZERI: SFIDE ED ESIGENZE

Il presente capitolo è dedicato ai servizi statali e non statali che in Svizzera si occupano di razzismo. In riferimento ai discorsi d'odio razzisti in Internet, quali sfide e interrogativi devono affrontare i servizi amministrativi, i consultori pubblici o al beneficio di un sostegno pubblico e le organizzazioni private? Di che cosa necessitano per poter (meglio) reagire a questo fenomeno? L'analisi della situazione che segue si basa su informazioni tratte dagli atti di un workshop organizzato dal SLR²⁹¹ e su interviste a persone attive sul campo. Gli aspetti qui presentati dovrebbero essere pertinenti per la maggior parte dei servizi rilevanti, il che non esclude che singoli aspetti non riguardino, del tutto o in parte, specifici organi.

Il modo in cui i singoli servizi gestiscono i discorsi d'odio razzisti in Internet dipende anche dalla dinamica dei casi che si presentano. Da un lato ci sono i casi più semplici riferiti a situazioni localmente delimitate nelle quali, per esempio, una persona non nota è vittima un'unica volta di un attacco razzista su un canale non pubblico, come WhatsApp. Dall'altro ci sono i casi più complessi, contraddistinti da attacchi di massa (*shitstorm*) o campagne d'odio con numerosi contributi online di matrice razzista che si diffondono in pochissimo tempo. Questo secondo tipo di casi è potenzialmente difficile da controllare, anche perché gli organi d'informazione tradizionali spesso ne parlano e ne amplificano così la visibilità e la dinamica. Ne sono un esempio l'accesso dibattito su un'area di transito per nomadi stranieri a Wileroltigen scatenatosi nel 2019/2020²⁹², l'appello a indossare la camicia con la stella alpina lanciato nel 2015 da allievi di una scuola secondaria su un gruppo WhatsApp²⁹³ oppure il discorso d'odio portato avanti nel 2014 da un gruppo formatosi spontaneamente su Facebook contrario a far alloggiare richiedenti l'asilo²⁹⁴ in un edificio scolastico.

8.1 Servizi amministrativi

In Svizzera, i servizi amministrativi sono strutturati secondo i principi del federalismo. A livello federale ne esiste uno specifico (SLR), mentre nei Cantoni e nei Comuni i servizi specializzati nella lotta al razzismo sono per lo più annessi a quelli di promozione dell'integrazione (in questi casi, gli interlocutori sono i delegati all'integrazione), incaricati di lottare contro il razzismo nel quadro della protezione dalla discriminazione e, in particolare, di occuparsi dell'informazione e della sensibilizzazione. Il finanziamento esclusivamente statale ne limita l'autonomia in merito al campo d'attività.

In veste di interlocutore principale per quanto attiene al razzismo in Internet, il SLR fissa quattro obiettivi di principio²⁹⁵:

1. *i consultori specializzati in razzismo devono essere dotati degli strumenti per gestire il razzismo in Internet*, in particolare per occuparsi autonomamente dei casi più semplici;
2. *l'opinione pubblica deve essere informata e sensibilizzata sui discorsi d'odio razzisti in Internet*. Il dibattito pubblico deve rendere la popolazione consapevole dei fattori d'influenza e delle ripercussioni dei discorsi d'odio in rete, consentendole così di valutare con spirito critico e senso di responsabilità i discorsi d'odio razzisti e discriminatori in Internet;

²⁹¹ Percorso: www.frb.admin.ch > Campi d'attività > Media e Internet > Prevenzione in Internet > [Workshops](#).

²⁹² Trasmissione Schweiz aktuell della televisione svizzero tedesca SRF, 9 febbraio 2020: [Ja zum Transitplatz für Fahrende in Wileroltigen](#). Un commento segnalato a titolo di esempio nel quadro di questo caso alla GRA è stato: «Für was 1000 vo franke usgä. Wesi glich überau ihre müll wärde verteile, s het platz für 36 wohnwäge, u der rest? U ner chames au jahr saniere weu ds pack eh aues kaputt macht, bi us i de läde chömesi cho chlaue bis zum geht nicht mehr, u ner schiebtmene no weiss ni was i arsch, nei sy» («Perché spendere migliaia di franchi se poi spargono lo stesso ovunque i loro rifiuti? C'è posto per 36 roulotte: e il resto? E poi bisogna rimettere tutto a posto ogni anno perché quelli rompono ogni cosa, vengono nei nostri negozi a rubare fino all'inverosimile e poi diamo loro ancora chissà che cosa. No, certamente no»).

²⁹³ Serafini, Aargauer Zeitung, 21 dicembre 2015: [Edelweiss-Streit: Auf Whatsapp blüht der jugendliche Patriotismus](#).

²⁹⁴ humanrights.ch, 23 agosto 2017: [Hassaufrufe im Internet - Schweizer Fälle und Politik der Newsportale](#).

²⁹⁵ Percorso: www.frb.admin.ch > Campi d'attività > [Media e Internet](#).

3. *le misure di prevenzione contro il razzismo in Internet devono essere rafforzate.* La prevenzione del razzismo deve oggi avvenire sia nel mondo analogico che in quello digitale, pubblicizzando in Internet le misure e i servizi preposti. Inoltre i progetti andrebbero promossi in funzione all'efficacia attesa, andrebbe valutata l'istituzione di un centro di segnalazione e andrebbe rafforzata la creazione di reti di attori rilevanti;
4. *le offerte d'intervento (p. es. a livello comunale) devono essere rafforzate.* I servizi locali vanno sostenuti affinché siano in grado di intervenire in casi più complessi.

Da approfondimenti mirati dell'attività di servizi amministrativi a parte il SLR, come quello per la promozione dell'integrazione della Città di Zurigo²⁹⁶, si evince che non è ancora attribuito un ruolo molto importante alla gestione del razzismo in Internet. Nonostante singoli dipendenti siano sensibilizzati all'argomento, nel loro complesso questi servizi paiono essere poco confrontati con casi di discorsi d'odio razzisti o quanto meno non sembrano occuparsene esplicitamente. Il motivo potrebbe risiedere nel fatto che la consapevolezza di questo tema è ancora poco diffusa al loro interno e nell'ambiente rilevante.

La tabella 9 riassume le sfide e le esigenze dei servizi amministrativi in generale.

Tabella 9 – Unità amministrative: situazione attuale

Sfide e interrogativi in riferimento ai discorsi d'odio razzisti in Internet	Di che cosa necessitano per agire con maggiore efficienza?	Rimandi (come e dove) nel presente rapporto
Lacune conoscitive	Miglioramento dello stato delle conoscenze in materia di social media e discorsi d'odio razzisti in Internet	Elaborazione di conoscenze (cap. 3-6)
Scarsa consapevolezza in materia di discorsi d'odio razzisti in Internet	Maggiore consapevolezza del fatto che il razzismo oggi si manifesta anche in rete	Elaborazione di conoscenze quale base per la sensibilizzazione (cap. 3-6)
Trasmissione ottimale delle conoscenze	Conoscenze sui gruppi di destinatari rilevanti e la loro raggiungibilità	Panoramica dei gruppi di destinatari e della loro raggiungibilità (cap. 9.3)
Carenza di conoscenze sui fattori di successo per i progetti in materia di razzismo in Internet	Criteri per la valutazione di progetti di promozione	Contromisure in atto (cap. 7.5); raccomandazione per la valutazione di progetti (cap. 9.4)
Promozione della collaborazione tra servizi rilevanti	Panoramica di persone e organizzazioni specializzate per sviluppare reti di contatti	Panoramica dei modelli di progetto disponibili in Svizzera e all'estero (cap. 7)

8.2 Consulitori pubblici (o finanziati con fondi pubblici)

I consulenti che si occupano di razzismo (ce ne sono in ogni Cantone, cfr. la Rete di consulenza per le vittime del razzismo²⁹⁷) sono interamente o in parte finanziati con fondi statali, circostanza che – a seconda della quota di fondi pubblici rispetto al budget totale – ne limita il campo d'attività.

Il tema del razzismo in Internet è «sul radar» dei consulenti, che però vedono molteplici possibilità di miglioramento (cfr. tabella 10). Il basso numero di segnalazioni registrato finora costituisce senza dubbio una base empirica troppo scarsa per poter permettere uno sviluppo sistematico delle competenze. Secondo la Rete di consulenza per le vittime del razzismo²⁹⁸, nel 2019 sono stati segnalati soltanto 23 episodi di razzismo in Internet poi confermati come atti di discriminazione. La cifra sommersa è verosimilmente parecchio superiore²⁹⁹. Questo dato rispecchia un problema fondamentale: le vittime non osano farsi avanti o non sanno a chi rivolgersi per un consiglio. Giudicano l'esperienza sconvolgente ma non di primo acchito discriminatoria oppure l'accettano come un dato di fatto e non reputano

²⁹⁶ P. es. intervista con Michael Bischof (Città di Zurigo, promozione dell'integrazione), febbraio 2020.

²⁹⁷ www.network-racism.ch.

²⁹⁸ Rete di consulenza per le vittime del razzismo 2020.

²⁹⁹ Ufficio federale di statistica 2019: dal rapporto «Convivenza in Svizzera» emerge che il 28 per cento della popolazione si considera vittima di discriminazione in generale.

necessario rivolgersi a un consultorio³⁰⁰. La dimensione digitale del razzismo, relativamente recente, potrebbe acuire questo problema.

In ogni caso, con le competenze e gli strumenti di cui dispongono, i consultori sono in grado di occuparsi almeno in parte di episodi d'odio razzista in Internet, seppure non ancora sistematicamente. Il loro dubbio è come reagire ad attacchi del genere: in base a quali criteri va classificata la rilevanza o la gravità degli attacchi? Come reagire di fronte ai diversi tipi di autore (p. es. singole persone o reti)? Come possono essere rapidamente identificati contenuti rilevanti e quale strategia va adottata per contrastarli? Da dove iniziare? Quando reagire e quando no? Quali ruoli possono assumere i consultori pubblici o finanziati con fondi pubblici, e quali sono i limiti dell'azione statale? Questi interrogativi riflettono lacune conoscitive: come si manifestano i discorsi d'odio razzisti in Internet? Come funzionano? Quali ne sono le premesse tecniche? Come sono considerati a livello giuridico? Secondo gli intervistati, urge sensibilizzare anche i livelli gerarchici più alti e i vari committenti, affinché i consulenti possano occuparsi efficientemente di questi casi.

Un altro ostacolo – ascrivibile alla lacunosa conoscenza dei discorsi d'odio razzisti in Internet – è costituito dalla mancanza praticamente totale di pubblicità digitale dell'offerta di consulenza in materia. Sui siti della maggior parte dei consultori che si occupano di razzismo, infatti, non si trovano praticamente informazioni sul tema del razzismo online o su prestazioni di consulenza riferite a questa espressione del razzismo. Al momento, le informazioni e gli esempi pubblicati sono per lo più limitati a episodi di razzismo nella ricerca di un alloggio o sul posto di lavoro. La dimensione digitale del fenomeno non viene praticamente citata. La presenza in rete dei consultori può pertanto essere migliorata, per esempio tematizzando sulle loro pagine e sui social media il razzismo online e pubblicizzando le proprie competenze in materia (non appena ne abbiano acquisite).

Intervenire nei casi più complessi si rivela rapidamente al di fuori della portata dei consultori. I Comuni o le organizzazioni coinvolti non dispongono di sufficienti competenze per accompagnare in modo tecnicamente corretto queste situazioni, soprattutto per quanto riguarda il diritto, la comunicazione e le misure di protezione psicosociali.

Tabella 10 – Consultori pubblici (o finanziati con fondi pubblici): situazione attuale

Sfide e interrogativi in riferimento ai discorsi d'odio razzisti in Internet	Di che cosa necessitano per agire con maggiore efficienza?	Rimandi (come e dove) nel presente rapporto
Consapevolezza soltanto parziale della dimensione digitale del razzismo, anche in ambito allargato	Maggiori conoscenze in seno ai consultori stessi, ma anche ai servizi sovraordinati, ai finanziatori e ai committenti su come funzionano i social media e i discorsi d'odio razzisti in Internet	Elaborazione di conoscenze quale base per la sensibilizzazione (cap. 3-6)
Strutturazione della consulenza / dell'accompagnamento in caso di episodi individuali o interpersonali	Maggiori conoscenze tecniche e giuridiche per gli interventi nei casi più semplici	Definizione delle conoscenze e delle competenze necessarie per i casi più semplici (cap. 10.1)
Carenza di conoscenze su quali misure possano essere adottate da servizi statali e non statali	Maggiori conoscenze sulla ripartizione delle competenze (e sui limiti dell'azione statale)	Panoramica delle competenze dei servizi rilevanti (cap. 9.1)
Nessuna esperienza con casi più complessi	Informazione su come gestire i casi più complessi; sostegno da parte di esperti	Definizione delle conoscenze e delle competenze necessarie per i casi più complessi (cap. 10.1.2)
Pubblicizzazione dell'offerta di consulenza	Conoscenze su come la propria offerta possa raggiungere i gruppi target	Panoramica dei gruppi target e della loro raggiungibilità (cap. 9.3); misure per rendere nota e fruibile l'offerta di consulenza (cap. 10.2)

³⁰⁰ In merito alla gestione dei casi viene inoltre segnalato che la base legale è percepita come debole (p. es. nessun diritto di azione delle associazioni), condizione che legherebbe le mani se, per esempio, si volesse rappresentare legalmente una persona privata per proteggerla dall'ulteriore esposizione a discorsi d'odio razzisti. Per informazioni di base sul quadro giuridico, cfr. humanrights.ch, 19 marzo 2020: [Rechtsentwicklung: Verstärkung des gesetzlichen Schutzes vor Diskriminierung?](#)

8.3 Organizzazioni private

Le organizzazioni finanziate esclusivamente con fondi privati godono della maggiore autonomia d'azione di tutte in virtù dell'indipendenza dal finanziamento pubblico. Il loro ruolo è tuttavia definito anche dai committenti per mezzo dei mandati. A parere dell'autrice, le poche organizzazioni private che in Svizzera si occupano di discorsi d'odio razzisti in Internet (p. es. la Fondazione contro il razzismo e l'antisemitismo GRA o #Netzcourage) ritengono il rischio in media più rilevante e acuto rispetto ai consulenti finanziati esclusivamente con fondi statali, forse perché per via della loro indipendenza si muovono in un campo in cui sono confrontate con un maggior numero di casi. Al momento, queste organizzazioni sono attive nella consulenza, nel monitoraggio e nella contronarrazione (per maggiori dettagli, cfr. cap. 7.5).

Un aspetto centrale sollevato da queste organizzazioni (per tutti gli aspetti, cfr. tabella 11) è la mancanza di sensibilità percepita in seno a organi rilevanti, come la polizia³⁰¹. Esempio in proposito è la diffusa prassi di consigliare alle vittime di cancellare il loro profilo. Questa proposta non considera però il fatto che oggi la gente costruisce i propri contatti personali e sociali (anche) tramite i social media, soprattutto i giovani, i politici e i giornalisti. La cancellazione del profilo equivale alla cancellazione della rete sociale e, autoimponendosi il silenzio, a una fuga. Oltretutto questa risposta non risolve il problema alla radice. La carenza di sensibilità è inoltre all'origine di esperienze negative con i servizi rilevanti³⁰². Le vittime, per esempio, affermano di non sentirsi sempre prese sul serio dalla polizia. Anche la procedura di denuncia è ritenuta onerosa, perché impone di passare in continuazione da strumenti digitali a situazioni analogiche e viceversa (p. es. fare uno screenshot e stamparlo, recarsi dalla polizia, la quale digita l'indirizzo del rimando ecc.).

Tabella 11 – Organizzazioni private: situazione attuale

Sfide e interrogativi in riferimento ai discorsi d'odio razzisti in Internet	Di che cosa necessitano per agire con maggiore efficienza?	Rimandi (come e dove) nel presente rapporto
Conoscenze in parte insufficienti	Aiuto per l'elaborazione di basi e strumenti	Elaborazione di conoscenze (cap. 3-6)
Percezione di una sensibilità carente in ambito allargato	Maggiore sensibilizzazione di organi rilevanti, p. es. la polizia	Elaborazione di conoscenze quale base per la sensibilizzazione (cap. 3-6); raccomandazioni (cap. 11)
Risorse umane e finanziarie insufficienti (soprattutto per i casi più complessi)	Sostegno finanziario e fattuale da parte dei servizi statali (p. es. riconoscimento pubblico, logo, percentuali d'impiego)	Consulenza e interventi (cap. 10); raccomandazioni (cap. 11)

³⁰¹ P. es. intervista con Jolanda Spiess-Hegglin (#Netzcourage), febbraio 2020.

³⁰² P. es. intervista con Jolanda Spiess-Hegglin (#Netzcourage) e Dominic Pugatsch (Fondazione contro il razzismo e l'antisemitismo, GRA), condotte rispettivamente nel febbraio e nel marzo del 2020.

9.1 Competenze e funzioni

La presente panoramica sulle attuali e possibili funzioni e competenze di organi amministrativi, servizi di consulenza pubblici o finanziati con fondi pubblici e organizzazioni private in relazione ai discorsi d'odio razzisti in Internet riprende le misure presentate nel capitolo 7.5 («Società civile in Svizzera e all'estero») e le completa presentando competenze e funzioni già esercitate da organi e servizi nell'ambito del razzismo offline³⁰³. L'accento è posto sulle misure di prevenzione, ma sono prese in considerazione anche quelle d'intervento, dato che nella prassi spesso coincidono.

Dall'elenco esposto alla tabella 12 si nota che le organizzazioni private adottano maggiormente misure rivolte specificamente al fenomeno in questione (p. es. contronarrazione, segnalazione). In teoria, questo tipo di incarico può essere assunto anche dai servizi pubblici o finanziati con fondi pubblici, a condizione che il loro mandato lo preveda e che non vi si opponga nessuna base legale. Per esempio, gli organi amministrativi potrebbero gestire un sistema di monitoraggio. Considerato però che lo Stato non può interferire nella libertà di opinione, i suoi organi non sono autorizzati a proporre controargomentazioni (a differenza degli attori politici preposti agli organi amministrativi). Il discorso è diverso quando i servizi pubblici o finanziati con fondi pubblici sono presenti nei social media con profili propri: in questo caso è importante che moderino il loro account e non tollerino i discorsi d'odio. Le organizzazioni private sono invece maggiormente autonome e possono assumere un ventaglio di incarichi più ampio, specialmente quando non sono confrontate con le attese che vanno di pari passo con il (co)finanziamento statale. In particolare nell'ambito della contronarrazione, il carattere partecipativo di un'azione nata «dal basso» o da gran parte della popolazione può generare un'accettazione più elevata³⁰⁴, il che è d'altra parte in linea con il principio secondo cui lo Stato non può interferire nella libertà di opinione. Anche per quanto concerne i casi più complessi, di natura spesso più politica, occorre chiarire quali servizi possano reagire e in che modo debbano farlo. Per esempio sarebbe ipotizzabile che un'istituzione statale prenda pubblicamente posizione lasciando a un'organizzazione privata la facoltà di intervenire con una contronarrazione. Quest'organizzazione deve a sua volta valutare nel singolo caso che cosa rientra nel suo ambito di competenza, in considerazione della quota degli aiuti statali percepiti, del contenuto del suo incarico, degli interessi dei committenti e delle proprie risorse.

Tabella 12 – Competenze dei principali servizi in Svizzera

Servizi amministrativi	<ul style="list-style-type: none"> • Prevenzione e sensibilizzazione • Funzione di coordinamento e di messa in contatto • Elaborazione e trasmissione delle conoscenze (nel quadro di strutture esistenti o nuove) • Formulazione di pareri (nel quadro del loro mandato) • Sostegno di progetti di organizzazioni private; promozione dello sviluppo di competenze in materia d'intervento
Servizi di consulenza pubblici (o finanziati con fondi pubblici)	<ul style="list-style-type: none"> • Prevenzione e sensibilizzazione • Sostegno a organizzazioni private • Trasmissione di conoscenze (p. es. mediante i social media, sul razzismo online) • Formulazione di pareri (nel quadro del loro mandato) • Consulenza e accompagnamento nel caso di episodi individuali e/o interpersonali, consultazione di esperti con esperienza pluriennale, trasferimento di casi a esperti
Organizzazioni private	<ul style="list-style-type: none"> • Prevenzione e sensibilizzazione • Consulenza e intermediazione • Protezione nel quadro di interventi nei casi più complessi • Strumenti di segnalazione e/o di monitoraggio attivo dei social media e dei media tradizionali • Messa a disposizione di materiale per la contronarrazione, coordinamento ed esecuzione di attività di contronarrazione • Segnalazione di contenuti ai gestori di social media e alle autorità

³⁰³ Il presente sottocapitolo si basa in particolare su interviste condotte con i responsabili di organi e servizi e sui risultati delle discussioni tenutesi in occasione di un workshop svolto nel 2018, cfr. Servizio per la lotta al razzismo 2018.

³⁰⁴ Intervista con Claire Pershan (Renaissance Numérique), aprile 2020. Pershan è responsabile dello strumento di contronarrazione [Seriously](#) (in francese), finanziato su base non governativa.

9.2 Strumento nazionale di segnalazione

Sul piano nazionale è raccomandabile l'istituzione di uno strumento di segnalazione centralizzato, il quale potrebbe comprendere una o più funzioni, tra cui:

- fornitura di materiale informativo e di consulenza in relazione ai discorsi d'odio razzisti in Internet;
- possibilità di segnalazione per vittime e spettatori;
- consulenza online;
- valutazione della perseguibilità di determinati contenuti;
- segnalazione (con lo status di «segnalatore attendibile») ai gestori di social media o ai media tradizionali e al mondo politico;
- assistenza nella formulazione di denunce penali;
- controllo delle segnalazioni e aggiornamento degli utenti interessati sullo loro stato di avanzamento;
- monitoraggio attivo dei social media e delle piattaforme medial tradizionali;
- pubblicazione di rapporti di monitoraggio con statistiche (su segnalazioni, eliminazioni, temi ecc.).

Quali funzioni possano essere integrate nello strumento di segnalazione dipende dal servizio a cui questo farebbe capo e dalle sue risorse. Se si intende limitarne l'utilizzazione, lo si può agganciare per esempio alla Commissione federale contro il razzismo (CFR). Se si prevede invece di istituire un servizio con svariati compiti, sarebbe opportuna una cooperazione tra organizzazioni private. Il servizio può anche essere gestito da una rete di diversi attori. Se non fosse sufficiente, si potrebbe ricorrere a un gruppo di esperti (p. es. psicologi, sociologi, giuristi), nel caso ideale in possesso di esperienza sia teorica che pratica, nonché a organizzazioni specializzate (p. es. GRA, #Netzcourage, Stop Hate Speech) e servizi di consulenza già operativi. È auspicabile che lo strumento sia disponibile in tutte le lingue nazionali (e, se del caso, anche in altre), semplice da utilizzare e facilmente accessibile (p. es. grazie alla possibilità di inviare segnalazioni in forma anonima attraverso il sito Internet e l'app). Oltre ai discorsi d'odio razzisti in Internet, si può prevedere la possibilità di segnalare anche altre dimensioni dei discorsi d'odio e forme analoghe di razzismo. La soluzione migliore per lo sviluppo e la messa in esercizio di uno strumento di respiro relativamente ampio sarebbe una vasta cooperazione tra servizi pubblici o finanziati con fondi pubblici, organizzazioni della società civile, scuole universitarie (professionali), rappresentanti del mondo dei media e dei gestori di social media, che permetterebbe di garantire la necessaria base di conoscenze e di rafforzare la visibilità. Il nome dello strumento deve infine essere tale da consentire di trovarlo rapidamente attraverso i motori di ricerca.

La creazione di uno strumento nazionale è auspicabile per diverse ragioni. Al giorno d'oggi chiunque è in grado di diffondere discorsi d'odio oltre i confini cantonali (e a livello internazionale), per cui le specificità cantonali sono rilevanti soltanto in parte. Uno strumento centralizzato premetterebbe inoltre di scongiurare la giungla di approcci, competenze e responsabilità osservabile in alcuni casi all'estero, facendo collaborare gli attori chiave fin dall'inizio (se possibile e necessario). L'autonomia dai Cantoni garantirebbe inoltre che i servizi possano reagire in modo neutrale alle segnalazioni, per esempio nel caso di discorsi d'odio pronunciati da politici locali, nei confronti dei quali i servizi cantonali potrebbero essere in un rapporto di dipendenza. Il carattere online dello strumento contribuirebbe inoltre a incoraggiare le segnalazioni poiché, rispetto alle opzioni di segnalazione fisica o telefonica, comporterebbe una minore discontinuità mediale per le persone interessate. Lo strumento non renderebbe assolutamente obsoleti i servizi di consulenza esistenti (sul loro ruolo, cfr. cap. 10).

9.3 Gruppi target e loro raggiungibilità

Per principio occorre sempre riflettere con cura sui *gruppi target* di misure e approcci. Nel caso dei discorsi d'odio razzisti in Internet, entrano in linea di conto diversi gruppi target, che possono essere raggiunti in svariati modi. Per ridurre la discontinuità mediale sono raccomandati soprattutto i canali digitali. L'offerta digitale deve inoltre adeguarsi alle abitudini medial dei gruppi che si vogliono raggiungere (vale a dire quali piattaforme vengono utilizzate, come, quando e per quanto tempo).

Poiché l'utilizzo delle piattaforme muta in modo dinamico (p. es. l'app TikTok è divenuta in poco tempo una delle piattaforme per la condivisione di video brevi più utilizzata tra i giovani) bisognerebbe rilevare le abitudini medialie dei gruppi target d'interesse poco prima dell'attuazione della relativa misura prevista. A tal fine si può ricorrere innanzitutto alle pubblicazioni disponibili sull'utilizzo dei media³⁰⁵. È inoltre possibile commissionare studi indipendenti ad hoc sulle abitudini medialie dei gruppi target di cui si vuole suscitare l'interesse. La visibilità può essere potenziata scegliendo un ampio ventaglio di canali, ovvero siti Internet, diversi social media, canali offline (workshop, relazioni) e moltiplicatori. Al contempo, per ottenere una sorta di effetto valanga, si può ricorrere alle reti digitali di diversi gruppi sociali. Alla tabella 13 sono presentati i gruppi target d'interesse e idee su come raggiungerli. Le procedure menzionate vanno però adeguate in funzione della misura scelta e sono dunque da interpretare come spunti provvisori.

³⁰⁵ P. es. [IGEM-digiMONITOR](#) del gruppo d'interesse Elektronische Medien Schweiz (IGEM).

Tabella 13 – Gruppi target e loro raggiungibilità

Gruppi target	Come raggiungerli
Ruolo in relazione ai discorsi d'odio	
Vittime	<ul style="list-style-type: none"> In particolare consulenza (cfr. cap. 10.2) o strumento nazionale di segnalazione Individuazione attraverso le caratteristiche che, in base agli studi, le vittime spesso presentano (cfr. cap. Fehler! Verweisquelle konnte nicht gefunden werden.); conseguente elaborazione di una campagna specifica in funzione del gruppo sui social media
Spettatori / pubblico	<ul style="list-style-type: none"> In funzione dello scopo delle misure previste Nell'ambito della contronarrazione, tenendo conto dei fattori di efficacia, capitolo 7.5.4
Divulgatori di discorsi d'odio online	<ul style="list-style-type: none"> Presenza di contatto diretto online (sistema noto dall'ambito dei progetti di deradicalizzazione; oneroso in termini di tempo) Contatto tramite denuncia ed eventualmente procedura penale (cfr. #Netzcourage) Individuazione attraverso le caratteristiche che, in base agli studi, i divulgatori di discorsi d'odio online presentano (cfr. cap. 5.2) Nell'ambito della contronarrazione, tenendo conto dei fattori di efficacia, capitolo 7.5.4
Gruppi vulnerabili, ovvero potenziali vittime o divulgatori di discorsi d'odio online	<ul style="list-style-type: none"> Individuazione attraverso le caratteristiche che, in base agli studi, aumentano il rischio (cfr. cap. 5.2 e Fehler! Verweisquelle konnte nicht gefunden werden.) Individuazione di sottoculture (gamification, satira, arte) e mobilitazione prima dell'intervento degli estremisti³⁰⁶
Categorie in base alla fascia d'età	
Bambini e giovani	<ul style="list-style-type: none"> A scuola; nell'ambito di moduli sulle competenze mediali (Lehrplan 21, PER, Piano di studio) Analisi del comportamento del target in materia di utilizzo dei media (p. es. attualmente di TikTok) e adeguamento del proprio
Giovani adulti	<ul style="list-style-type: none"> Analisi del comportamento del target in materia di utilizzo dei media e adeguamento del proprio
Adulti, anziani	<ul style="list-style-type: none"> Analisi del comportamento in materia di utilizzo dei media (fondamentalmente più difficile da contattare) e adeguamento del proprio
Altri gruppi sociali	
Personaggi politici attivi a livello nazionale, cantonale o comunale (in quanto vittime o autori ³⁰⁷)	<ul style="list-style-type: none"> Attraverso i partiti e la diffusione di materiale di consulenza da parte degli stessi (p. es. opuscolo informativo destinato ai politici, in tedesco) Possibilità di contattare tramite lo strumento nazionale di segnalazione i personaggi pubblici che diffondono discorsi d'odio
Migranti, svizzeri con retroterra migratorio, profughi	<ul style="list-style-type: none"> In particolare consulenza (cfr. cap. 10.2)
ONG, attivisti (in quanto vittime o moltiplicatori)	<ul style="list-style-type: none"> In particolare consulenza (cfr. cap. 10.2)

9.4 Criteri per la valutazione dei progetti presentati

Per la valutazione del diritto di un progetto a beneficiare di aiuti finanziari si rimanda ai criteri indicati sul sito Internet del SLR³⁰⁸. In aggiunta a questi, nella tabella 14 sono elencati i criteri considerati di particolare rilievo nel contesto dei discorsi d'odio razzisti in Internet. Alcuni criteri sono più importanti per determinati approcci progettuali piuttosto che per altri. Occorre dunque determinare nel singolo caso quali applicare. I criteri qui presentati in forma sintetica sono tratti da valutazioni e relative raccomandazioni su progetti passati, svolti in Svizzera e all'estero (cfr. cap. 7.5).

³⁰⁶ Ebner, citato in Heiderich, Fearless Democracy, 26 aprile 2018: [Extremismusforscherin Julia Ebner: «Hasskampagnen folgen einem klaren Muster»](#).

³⁰⁷ Per un dibattito sulla responsabilità dei politici in merito ai discorsi d'odio in Internet, cfr. sCAN 2020: [Hotspots of Hate: The online responsibility of public figures](#).

³⁰⁸ www.frb.admin.ch > [Aiuti finanziari](#).

Tabella 14 – Criteri per la valutazione dei progetti presentati

Competenze e preparazione	<ul style="list-style-type: none"> • Sono date le competenze tecniche necessarie? Sono date le competenze per la gestione di strumenti di marketing sui social media? Si può ricorrere a queste competenze in qualsiasi momento (in particolare in caso di emergenza)? • È garantita la comprensione teorica dei media digitali? • Il progetto si riallaccia a progetti in atto in Svizzera e, soprattutto, all'estero?
Gruppo target	<ul style="list-style-type: none"> • In che misura si dispone di conoscenze sul gruppo target, nonché sulle sue abitudini ed esperienze digitali? Come vengono impiegate queste informazioni nell'attuazione delle misure e per raggiungere il gruppo target? • Il gruppo target è coinvolto nello sviluppo e nel test delle misure? • Il gruppo target è delimitato in modo chiaro e opportuno? In che modo si tiene conto di gruppi vulnerabili o non ancora sensibilizzati? • In che misura gli aspetti terminologici e le prospettive sono trattati con la dovuta sensibilità? Si tiene conto dei loro possibili effetti sul gruppo target? (p. es. resistenza a causa dell'utilizzo del termine «discorso d'odio»)
Contenuti del progetto	<ul style="list-style-type: none"> • Quali tipi di discorso d'odio razzisti in Internet (p. es. antimusulmano, antisemita) sono presi in considerazione? Come viene affrontata la difficoltà dell'inclusione di più forme di discorso d'odio? (particolarmente importante per progetti di monitoraggio e di contronarrazione) • Il progetto si riallaccia al contesto locale e nazionale e a episodi d'attualità? • L'approccio scelto è impostato in modo semplice? (particolarmente importante per progetti sulle competenze medial e di contronarrazione) • L'approccio scelto è originale e creativo? (accresce tra l'altro la visibilità nei social media) • Che tipo d'impegno si richiede ai partecipanti e in che misura? Come è preparato il progetto, dal punto di vista organizzativo, a un livello di partecipazione inaspettatamente alto? • con che accuratezza sono state scelte le piattaforme di cui tenere conto (social media, siti Internet, app)? Vi è una riflessione su vantaggi, punti deboli e dinamiche dal punto di vista del raggiungimento degli obiettivi? Sono previsti servizi offline?
Protezione	<ul style="list-style-type: none"> • In che misura è prevista una protezione dei partecipanti coinvolti? (particolarmente importante per progetti di monitoraggio e di contronarrazione) • Vi è una riflessione su aspetti di tipo etico e sulla sicurezza? (p. es. sfera privata online, elaborazione di dati online) • In che misura è affrontata la questione della possibile comparsa di profili fasulli e di bot? • Sono presi in considerazione altri potenziali rischi?
Promotori	<ul style="list-style-type: none"> • In che misura il raggiungimento degli obiettivi del progetto può essere influenzato dall'orientamento ideologico o dalla reputazione dei suoi promotori? Vi è una riflessione su questo aspetto e su come contrastarlo, se del caso?
Pubbliche relazioni	<ul style="list-style-type: none"> • È previsto un investimento nella risonanza mediatica e, in caso affermativo, di che tipo? • È perseguita la creazione di una rete di contatti con esponenti della politica, dell'economia e delle organizzazioni della società civile e con altri moltiplicatori? • Vi è una preparazione a possibili critiche da parte dell'opinione pubblica o attacchi nello spazio digitale? (p. es. pianificazione delle modalità di reazione, ripartizione delle competenze per quanto concerne la moderazione del profilo sui social media e la preparazione di FAQ con risposte alle possibili domande) • Il modo di procedere è comunicato in modo trasparente (p. es. definizione di dati, algoritmi ecc.)? (particolarmente importante per progetti di monitoraggio)
Valutazione	<ul style="list-style-type: none"> • Come vengono misurati successo e/o efficacia del progetto? • Sono stati individuati parametri di misurazione che permettono di valutare il raggiungimento degli obiettivi? Vi è un chiaro metodo per l'osservazione e la rilevazione dei parametri scelti? • È garantito un certo grado d'indipendenza della valutazione?
Continuità	<ul style="list-style-type: none"> • L'orizzonte temporale del progetto è orientato sul lungo periodo a prescindere da una durata limitata del finanziamento del SLR? In caso affermativo, è prevista l'elaborazione di linee guida per la prosecuzione della sua gestione?
Finanziamento	<ul style="list-style-type: none"> • È stata formulata e debitamente motivata una stima dei costi per la produzione di materiale (in particolare nell'ambito della contronarrazione) e per la sua diffusione (p. es. campagna pubblicitaria online, sui social media)? • Nel preventivo sono previste anche le risorse per il monitoraggio, la misurazione e la valutazione (p. es. con l'aiuto di parametri)? • Nel caso in cui è prevista una vasta mobilitazione attraverso i social media: sono presi in considerazione ulteriori canali di finanziamento, come il crowdfunding (ovvero il coinvolgimento di un cospicuo numero di utenti Internet in qualità di finanziatori).

10 CONSULENZA E INTERVENTI IN SVIZZERA

10.1 Conoscenze e competenze di cui i consultori devono disporre

Per prestare consulenza nei casi di discorsi d'odio razzisti online, oltre alle conoscenze e competenze necessarie per assistere le vittime di episodi di razzismo offline occorre anche avere dimestichezza con Internet e i social media³⁰⁹. Maggiore è la complessità e la portata dei casi segnalati, maggiori sono anche le esigenze che i consultori devono soddisfare.

10.1.1 Consulenza nei casi più semplici

Nei casi più semplici e di portata limitata, in linea di massima occorre reagire rapidamente, dato che i contenuti non cancellati possono diffondersi rapidamente. Si raccomandano le seguenti conoscenze e competenze:

- Conoscenza dell'iter tipico (1: consulenza; 2: raccolta delle prove; 3: eventuale sostegno in caso di segnalazione ai gestori di social media; 4: eventuale sostegno in caso di procedimento giudiziario).
- Conoscenza del contesto tecnologico: conoscenza dei social media, del loro funzionamento, del loro raggio d'azione e delle loro dinamiche.
- Conoscenze teoriche: conoscenza della frequenza, delle manifestazioni, delle cause, dei fattori di influenza e delle ripercussioni dei discorsi d'odio online.
- Sensibilità psicosociale nei confronti delle forme «digitali» di vittimizzazione.
- Conoscenze delle strategie di autoprotezione preventive o reattive (p. es. impostazioni relative alla privacy).
- Conoscenze legali di base sui discorsi d'odio razzisti nella sfera digitale.
- Conoscenza dei possibili interventi, dei loro vantaggi e svantaggi e di come metterli in atto (p. es. blocco, segnalazione, perseguimento penale, ricorso a reti di sostegno psicologico).
- Diversificazione degli interventi a seconda dei casi.
- Conoscenza di organizzazioni specializzate o persone a cui le vittime si possono rivolgere, per esempio per «segnalazioni di massa» (numerosi utenti segnalano in modo coordinato un contenuto inappropriato, favorendone la rapida cancellazione), per controbattere o rettificare i discorsi d'odio o per sporgere denuncia.

10.1.2 Consulenza in casi più complessi e intervento

Negli ultimi anni sono regolarmente emersi casi di notevole portata, non sempre segnalati ai servizi per la lotta al razzismo. Questi episodi, verificatisi nelle aule scolastiche, in ambito politico o in altri contesti, possono creare danni notevoli, per cui richiedono un approccio rapido e accorto. Oltre alle conoscenze e competenze necessarie nei «casi più semplici», i consultori o gli esperti esterni devono avere dimestichezza con i seguenti aspetti:

- Conoscenza dei modi per affrontare i casi più complessi e delle relative dinamiche (p. es. identificazione di gruppi di utenti, tipo di comunicazione).
- Conoscenza del sistema mediatico tradizionale, dei suoi meccanismi e delle possibilità di esercitare un'influenza al suo interno (p. es. panorama mediatico, reagire mediante articoli di replica o di contestualizzazione).
- Conoscenze delle misure di protezione preventive o reattive contro attacchi di massa (p. es. identificazione di temi critici, piano di crisi per un'informazione rapida, monitoraggio di persone o istituzioni in Internet, assunzione temporanea di profili sui social media nel rispetto della protezione dei dati, contronarrazione / rettifiche / segnalazioni di massa).

³⁰⁹ Precorso: www.frb.admin.ch > Campi d'attività > Media e Internet > [Consulenza](#).

- Conoscenze dei ruoli e delle responsabilità (p. es. gestori dei social media: cancellazione di contenuti; Stato: perseguimento penale o presa di posizione; media tradizionali: community management, articoli di replica o di contestualizzazione).
- Dopo un attacco: eventualmente sostegno psicosociale per evitare un «ritiro» dal mondo digitale.
- Conoscenza di esperti con una vasta rete di contatti (p. es. ex politici, giornalisti, docenti), conoscenza della scena razzista online, personalità accattivante (*people person*: persona estroversa, comprensiva e che ispira fiducia)³¹⁰. È ipotizzabile la creazione di reti di esperti esterni specifiche in base alla regione linguistica; non tutte le regioni linguistiche ne dispongono però nella stessa misura.

10.2 Misure per far conoscere e rendere fruibili i servizi di consulenza

Vi sono disparate misure per far conoscere e rendere visibili i servizi dei consultori in materia di discorsi d'odio online alle vittime e alle persone interessate:

- «Andare verso il gruppo target», ossia adeguare la strategia di promozione dei servizi alle abitudini digitali del gruppo target (soprattutto promozione nei canali digitali, dato che le pertinenti offerte devono essere vicine al mondo di Internet e dei social media; eventualmente targeting in funzione di specifici gruppi di destinatari).
- Utilizzare non soltanto applicazioni digitali ad ampio raggio (siti Internet di consulenza, app, profili sui social media ecc.), ma anche canali offline, che consentano di instaurare un rapporto di fiducia grazie all'interazione diretta (workshop, conferenze, contatti diretti ecc.).
- Allacciare contatti con altri consultori e con moltiplicatori (p. es. esponenti del mondo politico, media tradizionali) per ampliare il raggio di diffusione delle offerte.
- Puntare sulla viralità: pubblicare contenuti virali sui social media (p. es. animazioni, video).
- Provvedere affinché l'opera di informazione e sensibilizzazione sul razzismo dei centri/servizi di integrazione e delle organizzazioni della società civile tenga conto anche della dimensione online del fenomeno.

Se l'offerta è stata trovata dalle vittime e dalle persone interessate: in che modo renderla fruibile nella maniera migliore possibile?

- Ridurre la discontinuità mediale: garantire l'accessibilità online del maggior numero possibile di prestazioni (p. es. consulenza online tramite chat, telefono o video).
- Impostare offerte accessibili e a misura di utente (p. es. garanzia dell'anonimato, sito Internet strutturato in maniera chiara e/o disponibile anche in una versione per smartphone).
- Marcare presenza online per creare fiducia e lanciare un segnale («Siamo qui e offriamo aiuto in qualsiasi momento»), tra l'altro per evitare l'«allontanamento» dei gruppi target e una loro diminuzione quali follower sui social media; a tale scopo occorre eventualmente l'aiuto di esperti nella gestione dei social media.
- Ispirare fiducia grazie alla qualità dei servizi assicurata da collaboratori con una formazione in ambito digitale (almeno una persona per consultorio).
- Ispirare fiducia grazie alla trasparenza, per esempio pubblicizzando i casi risolti e i successi ottenuti sui canali dei social media.
- Inserire un link al servizio all'interno di un eventuale futuro sistema nazionale di segnalazione.

³¹⁰ Intervista con Jonathan Birdwell (Institute of Strategic Dialogue), aprile 2020.

11 RACCOMANDAZIONI PRATICHE

Il presente rapporto raccomanda diverse misure per prevenire e limitare i discorsi d'odio razzisti online in Svizzera³¹¹. In linea di massima si può affermare che l'attuale offerta di servizi è limitata, soprattutto se paragonata all'estero, e andrebbe fortemente ampliata.

Sensibilizzazione ad ampio raggio e rafforzamento delle competenze

I social media dovrebbero essere considerati un aspetto integrante della convivenza civile e quindi presi sul serio come principale canale di diffusione dei discorsi d'odio razzisti online: eppure si constatano una mancanza di consapevolezza e insufficienti conoscenze e competenze a svariati livelli. Si raccomanda pertanto un'opera di sensibilizzazione e di rafforzamento delle competenze medialità a 360 gradi, per favorire la conoscenza del fenomeno, anche nelle sue implicazioni tecniche e giuridiche, per esempio da parte degli operatori attivi sul campo (p. es. consultori) e i moltiplicatori (p. es. docenti), nonché ai diversi livelli gerarchici di Comuni, Cantoni e Confederazione. Una sensibilità complessiva per tutti gli aspetti del razzismo in rete e le sue ripercussioni può contribuire a far sì che i discorsi d'odio razzisti online vengano considerati un problema con cui occorre fare i conti. I mandati di informazione e prevenzione dovrebbero quindi per principio tenere implicitamente ed esplicitamente conto dello spazio digitale. Per gli operatori del settore, il lavoro in questo campo può gettare le basi per l'elaborazione di una strategia efficace contro i discorsi d'odio razzisti in rete. In questo contesto, nei casi di discorsi d'odio di dominio pubblico si raccomandano anche prese di posizione visibili da parte degli organi ufficiali (p. es. la CFR).

Collaborazione tra organizzazioni pubbliche, della società civile e private

I discorsi d'odio razzisti in rete costituiscono un fenomeno dinamico e ricco di sfaccettature, che spesso si manifesta al di fuori di strutture organizzate consolidate. Per farvi fronte occorre un approccio ad ampio spettro, che tenga conto delle esigenze e delle responsabilità di svariati soggetti e sfrutti le loro capacità di intervento. Tra questi soggetti figurano le istituzioni statali, attori pubblici come i politici, le scuole e le università, la società civile, i gestori dei social media, i media tradizionali (mediante un'informazione responsabile e il community management), nonché la giustizia e la polizia (p. es. miglioramento nel primo contatto, accertamento di denunce online). La collaborazione tra diversi organi consente di cumulare le conoscenze e legittimare i provvedimenti adottati. In quest'ottica si possono per esempio ipotizzare workshop e tavole rotonde³¹² per favorire lo scambio di conoscenze ed esperienze e promuovere i contatti reciproci o incontri periodici di coordinamento e networking per rafforzare la cooperazione o il sovvenzionamento, a seconda delle esigenze, di progetti da parte dello Stato oppure la creazione di una piattaforma online centralizzata (p. es. un sistema nazionale di segnalazione) che fornisca un quadro d'insieme sugli attori rilevanti del settore e sui loro ruoli, oltre ai loro link.

Approccio olistico con diverse misure per quanto possibile fondate su prove scientifiche

Come illustrato nel presente rapporto, tutti gli approcci presentano vantaggi e svantaggi. Ognuno di essi costituisce a ogni modo un tassello integrante di un mosaico di soluzioni preventive e reattive. Pertanto si consiglia di ricorrere a una combinazione di misure complementari, che spaziano dalla semplice procedura di segnalazione mediante un apposito sistema nazionale alla consulenza competente e, se necessario, al perseguimento penale efficace e alla contronarrazione. Sono immaginabili anche la cancellazione efficiente di contenuti inappropriati da parte dei gestori dei social media e una maggiore trasparenza: in quest'ottica lo Stato, nei limiti delle possibilità date, potrebbe esercitare pressioni per migliorare i meccanismi di controllo tenendo conto delle leggi locali svizzere. In una prospettiva di lungo termine occorre anche evitare condizioni strutturali che favoriscano la frustrazione, l'odio e la violenza, quali l'emarginazione e l'umiliazione, le disparità di opportunità (p. es. la povertà di fronte al benessere),

³¹¹ Queste raccomandazioni coincidono con quelle formulate nei Paesi limitrofi dove la situazione è in parte simile a quella svizzera (p. es. Austria, Germania). Cfr. Republik Österreich: Parlamentsdirektion 2016; Baldauf et al. 2018.

³¹² Va segnalato a tale proposito l'incontro d'avvio, organizzato congiuntamente dal SLR e dalla piattaforma «Giovani e media», del programma prioritario sui discorsi d'odio, 25 agosto 2020.

un'offerta formativa insufficiente e la sensazione di iniquità, ingiustizia e sfiducia³¹³. Quanto accade nella sfera digitale riflette infatti le condizioni sociali predominanti: se si ignora questa circostanza, le misure adottate si riducono a un'effimera lotta contro i sintomi. Per garantire l'efficacia, la congruità e la proporzionalità delle misure, occorre tener conto delle conoscenze scientifiche acquisite e delle esperienze maturate nel corso degli anni: in questo modo è anche possibile garantire un equilibrio tra la protezione dei gruppi vulnerabili dalla discriminazione in rete e il diritto alla libertà d'espressione e alla tutela della sfera privata. I discorsi d'odio possono infatti essere combattuti, ma in una società aperta probabilmente non scompariranno mai. Per questo motivo, un approccio fondato su prove scientifiche offre soluzioni per contrastare efficacemente il fenomeno a lungo termine.

Investire nella ricerca e nel monitoraggio

Vista l'attuale carenza di dati sui discorsi d'odio razzisti online in Svizzera, è fondamentale acquisire maggiori conoscenze sistematiche sul fenomeno e mettere a disposizione le necessarie risorse finanziarie. Soprattutto i sondaggi tra la popolazione possono essere utili per individuare autori, vittime e spettatori e il loro background socio-strutturale e quindi elaborare contromisure ad hoc per gruppi target specifici, tenendo conto di un ampio contesto di nessi causali e ripercussioni sulla società. L'analisi qualitativa dei discorsi d'odio in rete e della loro diffusione sulle varie piattaforme può favorire la comprensione approfondita della «subcultura razzista on-line» locale. Un monitoraggio sistematico, fondato su prove scientifiche, sull'evoluzione continua dei fattori scatenanti, dei temi, delle figure chiave e dei gruppi contribuisce infine a un'attività di prevenzione e di intervento orientata strategicamente.

Rivolgersi a gruppi target selezionati, in particolare ai gruppi vulnerabili

Le misure dovrebbero essere fortemente incentrate sugli *utenti finali*. Nell'area germanofona vi sono ormai numerose campagne e materiali informativi sul tema rivolti alla collettività. Il rischio, soprattutto con le campagne online, è però di raggiungere un pubblico già «convertito alla causa». Vale dunque la pena di individuare i diffusori effettivi o potenziali di discorsi d'odio online e le rispettive vittime e di adeguare il messaggio ai loro moventi e all'ambiente digitale, ma anche analogico, in cui sono inseriti. Studi condotti sulle caratteristiche di questi gruppi target (cfr. cap. 5.2), sondaggi (ancora da effettuare) nei contesti rilevanti (p. es. nelle scuole) e interviste alle vittime e, idealmente, agli autori di discorsi d'odio forniscono un orientamento in materia. I diffusori di discorsi d'odio online (cfr. cap. 5.2) rappresentano di norma una minoranza, ma, se sono molto attivi e visibili in rete, possono ostacolare la formazione di una società civile digitalmente matura. Bambini e adolescenti in generale, ma in particolare quelli con caratteristiche specifiche (cap. 4.3 e 5.2), devono essere sensibilizzati facendo leva sul loro forte coinvolgimento e con la partecipazione dei docenti, in modo da mettere precocemente in moto importanti processi di socializzazione. Particolare sensibilità è richiesta nei confronti di chi è stato sia vittima che autore di discorsi d'odio. In linea di massima, si raccomanda di elaborare misure specifiche per lo spazio digitale invece di riproporre provvedimenti dimostratisi validi contro il razzismo offline. La creazione di reti di sostegno e solidarietà per le vittime nella vita reale assume per esempio forme diverse (soprattutto contatto personale) che nello spazio digitale (estese ma labili comunità virtuali che provvedono a una contronarrazione pubblica immediata a eventuali attacchi).

Ampliare e pubblicizzare i servizi di consulenza e gli interventi su scala locale e nazionale

Siccome al giorno d'oggi gli episodi di razzismo non si manifestano soltanto offline, ma anche – e forse già prevalentemente – online, si raccomanda ai consultori di tenere maggiormente conto della dimensione digitale del fenomeno. Così come negli ultimi 20 anni i gruppi estremisti hanno sfruttato le opportunità comunicative offerte da Internet, occorre ora recuperare terreno su questo fronte per contrastare lo squilibrio tra i discorsi d'odio e le pertinenti contromisure. La chiave in quest'ottica è il rafforzamento delle competenze digitali e della presenza digitale. Vi è infatti la possibilità che, a causa della pressoché totale assenza di servizi di consulenza esplicitamente focalizzati sul razzismo online, le vittime di simili episodi rinuncino a segnalarli, il che può indurre a sua volta gli attori del settore a ritenere non necessario il rafforzamento delle competenze e della presenza digitale. Questo ipotetico ma forse reale circolo vizioso potrebbe essere spezzato mediante un'offerta digitale che risulti ampiamente conosciuta e fruibile. In questo modo, vittime e spettatori sarebbero meglio protetti. In questo contesto i consultori locali assumono un ruolo importante poiché, grazie alla conoscenza delle

³¹³ Bauer 2011.

strutture del territorio, sono in grado di raggiungere i gruppi target rilevanti e, di fronte a casi concreti, hanno una maggiore dimestichezza con il contesto politico e culturale. Un coordinamento nazionale supplementare può tuttavia essere d'aiuto finché le competenze necessarie non sono ancora disponibili a livello cantonale. È inoltre immaginabile l'introduzione, anche a lungo termine, di un sistema nazionale centralizzato di segnalazione, dato che con un click i discorsi d'odio razzisti online possono assumere una dimensione nazionale.

12 ELENCO DELLE ABBREVIAZIONI

AfD	Alternativa per la Germania (<i>Alternative für Deutschland</i>)
AIRR	Autorità indipendente di ricorso in materia radiotelevisiva
AJS NRW	Azione per la protezione della gioventù della Renania Settentrionale-Vestfalia (<i>Aktion Jugendschutz Nordrhein-Westfalen</i>)
AWD	Divisione Armi atomiche (<i>Atomwaffen Division</i>)
CC	Codice civile svizzero (<i>RS 210</i>)
CEJI	Un contributo ebraico per un'Europa inclusiva (<i>A Jewish contribution to an inclusive Europe</i>)
CERD	Comitato per l'eliminazione della discriminazione razziale (<i>Committee on the Elimination of Racial Discrimination</i>)
CFR	Commissione federale contro il razzismo
CICAD	Coordinamento intercomunitario contro l'antisemitismo e la diffamazione (<i>Coordination Intercommunautaire Contre l'Antisémitisme et la Diffamation</i>)
Cost.	Costituzione federale (<i>RS 101</i>)
CP	Codice penale svizzero (<i>RS 311.0</i>)
DGUV	Ente tedesco per l'assicurazione contro gli infortuni (<i>Deutsche Gesetzliche Unfallversicherung</i>)
ECRI	Commissione europea contro il razzismo e l'intolleranza (<i>European Commission against Racism and Intolerance</i>)
FSCI	Federazione svizzera delle comunità israelite
GIF	Graphics Interchange Format
GRA	Fondazione contro il razzismo e l'antisemitismo (<i>Stiftung gegen Rassismus und Antisemitismus</i>)
HEP Vaud	Alta scuola pedagogica Vaud (<i>Haute École Pédagogique Vaud</i>)
IB	Movimento identitario (<i>Identitäre Bewegung</i>)
INACH	Rete internazionale contro la diffusione dell'odio in Internet (<i>International Network Against Cyber Hate</i>)
ISD	Istituto per il dialogo strategico (<i>Institute of Strategic Dialogue</i>)
ISIS	Stato islamico
LICRA	Lega internazionale contro il razzismo e l'antisemitismo (<i>Ligue Internationale Contre le Racisme et l'Antisémitisme</i>)

LRTV	Legge federale del 24 marzo 2006 sulla radiotelevisione (<i>RS 784.40</i>)
OCCI	Iniziativa per il coraggio civile online (<i>Online Civil Courage Initiative</i>)
OGN	Organizzazione non governativa
PEGIDA	Europei patriottici contro l'islamizzazione dell'Occidente (<i>Patriotische Europäer gegen die Islamisierung des Abendlandes</i>)
sCAN	Specialised Cyber-Activists Network
SLR	Servizio per la lotta al razzismo
SLR	Servizio per la lotta al razzismo
UE	Unione europea
UNICRI	Istituto interregionale delle Nazioni Unite per la ricerca sul crimine e la giustizia (<i>United Nations Interregional Crime and Justice Research Institute</i>)
USI	Università della Svizzera italiana
UZH	Università di Zurigo
ZHAW	Scuola universitaria di scienze applicate di Zurigo (<i>Zürcher Hochschule für Angewandte Wissenschaften</i>)

Articoli scientifici, rapporti e libri

- Adena M. et al. 2015. Radio and the rise of the Nazis in prewar Germany. *The Quarterly Journal of Economics* 130(4): 1885-1939.
- Aizenkot D. e G. Kashy-Rosenbaum. 2018. Cyberbullying in whatsapp classmates' groups: Evaluation of an intervention program implemented in Israeli elementary and middle schools. *New Media & Society* 20(12): 4709-4727.
- Álvarez-Benjumea A. e F. Winter. 2018. Normative change and culture of hate: An experiment in online environments. *European Sociological Review* 34(3): 223-237.
- Anderson A. A. et al. 2014. The «Nasty effect»: Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication* 19(3): 373-387.
- Awan I. e I. Zempi. 2015. We fear for our lives: Offline and online experiences of anti-Muslim hostility (ottobre).
- Baier D. 2019. Kriminalitätsoferenerfahrungen und Kriminalitätswahrnehmungen in der Schweiz: Ergebnisse einer Befragung. ZHAW (agosto).
- Baldauf J., Ebner J., e J. Guhl. 2018. *Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht Hassrede*. Institute for Strategic Dialogue, Londra.
- Barlow C. e I. Awan. 2016. «You need to be sorted out with a knife»: The attempted online silencing of women and people of Muslim faith within academia. *Social Media + Society* 2(4): 1-11.
- Barnidge M. et al. 2019. Perceived exposure to and avoidance of hate speech in various communication settings. *Telematics and Informatics* 44: 101263.
- Bauer J. 2011. *Schmerzgrenze: Vom Ursprung alltäglicher und globaler Gewalt*. Karl Blessing Verlag, Monaco di Baviera.
- Ben-David A. e A. Matamoros-Fernández. 2016. Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication* 10: 1167-1193.
- Blaya C. 2015. *Cyberviolence et école – Les Dossiers des Sciences de l'Éducation*, 33. Presses universitaires du Midi, Tolosa.
- Blaya C. 2019. Cyberhate: A review and content analysis of intervention strategies. *Aggression and Violent Behavior* 45: 163-172.
- Blaya C. e C. Audrin. 2019. Toward an understanding of the characteristics of secondary school cyberhate perpetrators. *Frontiers in Education* 4(46).
- Bliuc A.-M. et al. 2018. Online networks of racial hate: A systematic review of 10 years of research on cyber-racism. *Computers in Human Behavior* 87: 75-86.
- Boyd D. M. 2010. Social network sites as networked publics: affordances, dynamics, and implications. In: *A Networked Self: Identity, Community, and Culture on Social Network Sites*, Papacharissi Z. (ed.): 39-58. Routledge, New York.
- Brady W. J. et al. 2017. Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* 114(28): 7313-7318.
- Breuer J. 2017. Hate Speech in Online Games. In: *Online Hate Speech: Perspektiven auf eine neue Form des Hasses*. Grimme-Institut – Gesellschaft für Medien, Bildung und Kultur mbH, Marl.
- Brown A. 2018. What is so special about online (as compared to offline) hate speech? *Ethnicities* 18(3): 297-326.
- Bucchianeri M. M. et al. 2014. Multiple types of harassment: Associations with emotional well-being and unhealthy behaviors in adolescents. *Journal of Adolescent Health* 54(6): 724-729.
- Buckels E. E. et al. 2014. Trolls just want to have fun. *Personality and Individual Differences* 67: 97-102.
- Bulut E. e E. Yörük. 2017. Digital populism: Trolls and political polarization of Twitter in Turkey. *International Journal of Communication* 11(25): 4093-4117.
- Celik S. 2019. Experiences of internet users regarding cyberhate. *Information Technology & People* 32(6): 1446-1471.
- Cerese A., E. D'Angelo e C. Santoro. 2015. Monitoring racist and xenophobic extremism to counter hate speech online: Ethical dilemmas and methods of a preventive approach. *VoxPol Workshop*, Brussels (19 gennaio).
- Chan J., A. Ghose e R. Seamans. 2013. The internet and hate crime: Offline spillovers from online access, Working Papers 13-02, NET Institute.
- Chandrasekharan E. et al. 2017. You can't stay here: The efficacy of Reddit's 2015 ban examined through hate speech. *Proceedings of the ACM on Human-Computer Interaction* 1(CSCW): 1-22.
- Cho D., S. Kim e A. Acquisti. 2012. Empirical analysis of online anonymity and user behaviors: The impact of real name policy. *45th Hawaii International Conference on System Sciences* 3041-3050. IEEE.
- Christopherson K.M. 2007. The positive and negative implications of anonymity in Internet social interactions: On the internet, nobody knows you're a dog. *Computers in Human Behavior* 23(6): 3038-3056.
- Chung et al. 2019. CONAN – COunter NArratives through Nichesourcing: a multilingual dataset of responses to fight online hate speech. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 28 luglio-2 agosto 2019: 2819-2829.

- Chyzh O., M. D. Nieman e C. Webb. 2019. The effects of dog-whistle politics on political violence. *Iowa State University, Political Science Publications* 59: 1-10.
- Cleland J. 2013. Racism, football fans, and online message boards: How social media has added a new dimension to racist discourse in English football. *Journal of Sport & Social Issues* 38: 415-431.
- Costello M. e J. Hawdon. 2018. Who are the online extremists among Us? Sociodemographic characteristics, social networking, and online experiences of those who produce online hate materials. *Violence and Gender* 5(1): 55-60.
- Costello M. et al. 2019. Social group identity and perceptions of online hate. *Sociological inquiry* 89(3): 427-452.
- Craker N. e E. March. 2016. The dark side of Facebook: The dark tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences* 102: 79-84.
- Daft R. L. e R. H. Lengel. 1986. Organizational information requirements, media richness and structural design. *Management Science* 32(5): 554-571.
- Daniels J. 2013. Race and racism in Internet studies: A review and critique. *New Media & Society* 15(5): 695-719.
- Daniels J. 2018. The algorithmic rise of the Alt-Right. *Contexts* 17(1): 60-65.
- Delgado R. e J. Stefancic. 2009. Four observations about hate speech. *Wake Forest Law Review* 44: 353-370.
- Dittrich M. et al. 2020. Alternative Wirklichkeiten. Monitoring rechts-alternativer Medienstrategien. *Amadeu Antonio Stiftung*, Berlino.
- Duggan M. 2017. Online Harassment 2017. *Pew Research Center* (luglio).
- Eckes C. et al. 2018. #Hass im Netz: Der schleichende Angriff auf unsere Demokratie. *Institut für Demokratie und Zivilgesellschaft*, Jena.
- Eco – Verband der Internetwirtschaft. 2019. [Eco Beschwerdestelle: Jahresbericht 2018](#).
- Eddington S. M. 2018. The communicative constitution of hate organizations online: A semantic network analysis of «Make America Great Again». *Social Media + Society* 4: 1-12.
- Elson M. e C.J. Ferguson. 2014. Twenty-five years of research on violence in digital games and aggression. *European Psychologist* 19(1): 33-46.
- Erjavec K. e M. P. Kovačić. 2012. You don't understand, this is a new war! Analysis of hate speech in news web sites' comments. *Mass Communication and Society* 15(6): 899-920.
- Fielitz M. e H. Marcks. 2019. Digital fascism: Challenges for the open society in times of social media. *Berkeley Center for Right-Wing Studies Working Paper Series* (luglio).
- Fortuna P. e S. Nunes. 2018. A survey on automatic detection of hate speech in text. *ACM Computing Surveys* 51(4): 85:1-30.
- Fox J. e Tang W.Y. 2017. Women's experiences with general and sexual harassment in online video games: Rumination, organizational responsiveness, withdrawal, and coping strategies. *New Media & Society* 19(8): 1290-1307.
- Gagliardone I. et al. 2015. *Countering online hate speech: UNESCO Series on Internet Freedom*. Organizzazione delle Nazioni Unite per l'educazione, la scienza e la cultura, Parigi.
- Gagliardone I. et al. 2016. Mechachal: Online debates and elections in Ethiopia – from hate speech to engagement in Social Media. *SSRN Electronic Journal* 2831369.
- Gatewood C. e I. Boyer. 2019. *Building digital citizenship in France. Lessons from the Sens Critique Project building digital citizenship in France*. Institute for Strategic Dialogue, Londra.
- Gatewood C. et al. 2020. Cartographie de la haine en ligne: Tour d'horizon du discours haineux en France. Institute of Strategic Dialogue.
- George C. 2015. Hate speech law and policy. In: *The International Encyclopedia of Digital Communication and Society, First Edition*, John Wiley & Sons, Inc. (ed.): 1-10.
- Geschke D. et al. 2019. #Hass Im Netz: Der schleichende Angriff auf unsere Demokratie. *Eine Bundesweite Repräsentative Untersuchung*. Campact.
- Guhl J., J. Ebner e J. Rau. 2020. *Das Online-Ökosystem rechtsextremer Akteure*. Institute for Strategic Dialogue, Londra.
- Hawdon J., A. Oksanen e P. Räsänen. 2017. Exposure to online hate in four nations: A cross-national consideration. *Deviant behavior* 38(3): 254-266.
- Haymoz S. et al. 2019. *L'estremismo politico tra i giovani in Svizzera: entità e fattori esplicativi: Rapporto di valutazione per il Canton Ticino*. ZHAW e HETS-FR (gennaio).
- Hayne S. C. e R. E. Rice. 1997. Attribution accuracy when using anonymity in group support systems. *International Journal of Human-Computer Studies* 47(3): 429-452.
- Hine G. E. et al. 2017. Kek, cucks, and god emperor Trump: A measurement study of 4chan's politically incorrect forum and its effect on the web. *Eleventh International AAAI Conference on Web and Social Media (ICWSM, ottobre)*.
- Hsueh M. et al. 2015. «Leave your comment below»: can biased online comments influence our own prejudicial attitudes and behaviors? *Human Communication Research* 41(4): 557-576.
- Il Consiglio federale 2013. *Base legale per i media sociali. Rapporto del Consiglio federale in adempimento del postulato Amherd 11.3912 (29 settembre 2011)*. Berna: Consiglio federale.
- Il Consiglio federale 2017. *Base legale per i media sociali: nuova analisi della situazione. Rapporto complementare del Consiglio federale in adempimento del postulato Amherd 11.3912 Diamo un quadro legale ai social media (11 settembre 2017)*. Berna: Consiglio federale.

- Isbister T. et al. 2018. Monitoring targeted hate in online environments. arXiv preprint arXiv: 1803.04757.
- Jakubowicz A. et al. 2017. How cyber users experience and respond to racism: Evidence from an online survey. In: *Cyber Racism and Community Resilience*, Palgrave Hate Studies, Springer: 65-94.
- Jeong S.-H., H. Cho e Y. Hwang. 2012. Media literacy interventions: A meta-analytic review. *Journal of Communication* 62(3): 454-472.
- Jones D. e Benesch S. 2019. *Combating hate speech through counterspeech*. Berkman Klein Center (9 agosto).
- Jourová V. 2019a. *Code of conduct on countering illegal hate speech online. Fourth evaluation confirms self-regulation works*. European Commission (12 febbraio).
- Jourová V. 2019b. *How the Code of Conduct helped countering illegal hate speech online*. Commissione europea (febbraio)
- Kaspar K., 2017. Hassreden im Internet – Ein besonderes Phänomen computervermittelter Kommunikation? In: K. Kaspar, L. Grässer und A. Riffi. *Online Hate Speech: Perspektiven auf eine neue Form des Hasses*. kopaed verlagsgmbh, Düsseldorf/München.
- Keen A. 2007. *The Cult of the Amateur*. Currency, New York.
- Kenski K., K. Coe e S. A. Rains. 2017. Perceptions of uncivil discourse online: An examination of types and predictors. *Communication Research* 1-20.
- Klein A. 2017. Hate speech in the information age. In: *Fanaticism, Racism, and Rage Online*. Klein A. (ed.): 25-39. Palgrave Macmillan, Cham.
- Kreissel P. et al. 2019. *Hass auf Knopfdruck: Rechtsextreme Trollfabriken und das Ökosystem koordinierter Hasskampagnen im Netz*. Institute for Strategic Dialogue, Londra.
- Landesanstalt für Medien NRW. 2018. *Ergebnisbericht Hassrede*. Forsa.
- Lanzke A. et al. 2013. *Viraler Hass: Rechtsextreme Wortergreifungsstrategien im Web 2.0*. Amadeu Antonio Stiftung, Heidelberg.
- Lapidot-Lefler N. e A. Barak. 2012. Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior* 28(2): 434-443.
- Latzer M., Büchi M. e N. Festic. 2019a. Internetverbreitung und digitale Bruchlinien in der Schweiz 2019: *Themenbericht aus dem World Internet Project – Switzerland 2019*. Zurigo: Università di Zurigo (ottobre).
- Latzer M., Büchi M. e N. Festic. 2019b. Internetanwendungen und deren Nutzung in der Schweiz 2019. *Themenbericht aus dem World Internet Project – Switzerland 2019*. Zurigo: Università di Zurigo (ottobre).
- Laubenstein S. e A. Urban. 2018. Fallbeispiele: Welche Arten von Kampagnen gegen Hass und Extremismus im Internet funktionieren, welche funktionieren nicht und warum? In: *Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht*. Baldauf J. et al. (ed.): 55-63. Institute for Strategic Dialogue, Londra.
- Ledwich M. e A. Zaitsev. 2019. Algorithmic extremism: Examining YouTube's rabbit hole of radicalization. arXiv preprint arXiv: 1912.11211.
- Lim M. 2017. Freedom to hate: Social Media, algorithmic enclaves, and the rise of tribal nationalism in Indonesia. *Critical Asian Studies* 49(3): 411-427.
- Lingiardi V. et al. 2019. Mapping Twitter hate speech towards social and sexual minorities: A lexicon-based approach to semantic content analysis. *Behaviour and Information Technology*. DOI: 10.1080/0144929X.2019.1607903.
- Lowry P. B. et al. 2016. Why do adults engage in cyberbullying on social media? An integration of online disinhibition and deindividuation effects with the social structure and social learning model. *Information Systems Research* 27(4): 962-986.
- Lucas B. 2014. *Methods for monitoring and mapping online hate speech*. GSDRC Applied Knowledge Services.
- Marwick A. E. 2015. Instafame: Luxury selfies in the attention economy. *Public Culture* 27(1/75): 137-160.
- Marwick, A. e R. Lewis. 2017. *Media manipulation and disinformation online*. Data & Society Research Institute, New York.
- Matamoros-Fernández A. 2017. Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication & Society* 20(6): 930-946.
- Mathew B. et al. 2019. Spread of hate speech in online social media. In: *11th ACM Conference on Web Science (WebSci '19), 30 giugno-3 luglio 2019*. Boston, MA, USA.
- Mittos A. et al. 2019. And we will fight for our race! A measurement study of genetic testing conversations on Reddit and 4chan. arXiv preprint arXiv:1901.09735.
- Mondal M. et al. 2018. Characterizing usage of explicit hate expressions in social media. *New Review of Hypermedia and Multimedia* 24(2): 110-130.
- Müller K. e C. Schwarz. 2019. Fanning the flames of hate: Social media and hate crime. *SSRN Scholarly Paper*.
- Munger K. 2017. Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Political Behavior* 39: 629-649.
- Murthy D. e S. Sharma. 2019. Visualizing YouTube's comment space: Online hostility as a networked phenomena. *New Media & Society* 21(1): 191-213.
- Musial J. 2017. «My Muslim sister, indeed you are a mujahidah» - Narratives in the propaganda of the Islamic State to address and radicalize Western Women. An Exemplary analysis of the online magazine Dabiq. *Journal for Deradicalization* (9): 39-100.
- Naguib T. 2014. *Begrifflichkeiten zum Thema Rassismus im nationalen und internationalen Verständnis: Eine Auslegeordnung unter Berücksichtigung des Völker- und Verfassungsrechts*. Expertise im Auftrag der

- Fachstelle für Rassismusbekämpfung FRB, Eidgenössisches Departement des Innern EDI. Zürcher Hochschule für Angewandte Wissenschaften (ZHAW), Winterthur / Berna (27 agosto).
- Neue deutsche Medienmacher & No Hate Speech Movement Deutschland. 2019. *Wetterfest durch den Shitstorm. Leitfaden für Medienschaffende zum Umgang mit Hass im Netz*. Berlino.
- O'Callaghan D. et al. 2012. An analysis of interactions within and between extreme right communities in social media. In: *Ubiquitous social media analysis*, Atzmueller M. (ed.): 88-107. Springer, Berlino/Heidelberg.
- Olteanu A. et al. 2018. The effect of extremist violence on hateful speech online. In: *Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM)*: 221-230. AAAI Press, Palo Alto, California.
- Ortiz S. M. 2019. The meanings of racist and sexist trash talk for men of color: A cultural sociological approach to studying gaming culture. *New Media & Society* 21(4): 879-894.
- Pariser E. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin, UK.
- Perry B. e P. Olsson. 2009. Cyberhate: The globalization of hate. *Information & Communications Technology Law* 18(2): 185-199.
- Quent M. 2018. Zivilgesellschaft: Das globale Dorf verteidigen: Strategien gegen den kulturellen Backlash in sozialen Medien. In: *Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht*. Baldauf J. et al. (ed.): 48-54. Institute for Strategic Dialogue, Londra.
- Rafael S. e A. Ritzmann. 2018. Hintergrund: Das ABC des Problemkomplexes Hassrede, Extremismus und NetzDG. In: *Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht*. Baldauf J. et al. (ed.): 11-19. Institute for Strategic Dialogue, Londra.
- Refaeil N. e A. Wiecken. 2018. *Rassismus Im Netz – Schweizer Recht*. Fachstelle für Rassismusbekämpfung. Präsentation, Berna (dicembre).
- Reichelmann A. et al. 2020. Hate knows no boundaries: Online hate in six nations. *Deviant Behavior*: 1-12.
- Republik Österreich: Parlamentsdirektion. 2016. [Digitale Courage](#) (novembre).
- Rete di consulenza per le vittime del razzismo. 2020. *Episodi di razzismo trattati nell'attività di consulenza: Gennaio – Dicembre 2019*. associazione humanrights.ch, Commissione federale contro il razzismo, Berna (aprile).
- Reynolds L. e H. Tuck. 2016. *The counter-narrative monitoring & evaluation handbook*. Institute for Strategic Dialogue, Londra.
- Ribeiro M. H. et al. 2018. Characterizing and detecting hateful users on Twitter. In: *Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM)*: 676-679. AAAI Press, Palo Alto, California.
- Rost K., L. Stahel e B. S. Frey. 2016. Digital social norm enforcement: Online firestorms in social media. *PloS One* 11(6): e0155923.
- Salminen J. et al. 2018. Online hate interpretation varies by country, but more by individual: A statistical analysis using crowdsourced ratings. In: *Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 15-18 ottobre 2018*: 88-94.
- Salminen J. et al. 2019. Online hate ratings vary by extremes: A statistical analysis. *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval (CHIIR), 10-14 marzo 2019*: 213-217. Glasgow.
- Schabas, W. A. 2001. Hate-Speech in Ruanda: The Road to Genocide. *McGill Law Journal*, vol. 46, 2001, pag. 301-315.
- Schieb C. e M. Preuss. 2016. Governing hate speech by means of counterspeech on Facebook. In: *66th Annual Conference of the International Communication Association (ICA), 9-13 giugno 2016*. Fukuoka: 1-23.
- Schmidt A. e M. Wiegand. 2017. A survey on hate speech detection using natural language processing. In: *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*. Valencia: 1-10.
- Sellars A. 2016. Defining hate speech. *Berkman Klein Center Research Publication No. 2016(20)*; *Boston Univ. School of Law, Public Law Research Paper No. 16(48)*
- Servizio per la lotta al razzismo 2018. Workshop «Rassismus im Netz»: Blitzlichter auf die Inhalte der Referate und Diskussionen. Berna (11 dicembre).
- Servizio per la lotta al razzismo. 2019. Discriminazione razziale in Svizzera. Berna (settembre).
- Shapiro A. 2019. Predictive policing for reform? Indeterminacy and intervention in big data policing. *Surveillance & Society* 17(3/4): 456-472.
- Siegel A. 2020. Online hate speech. In: *Social Media and Democracy: The State of the Field*. Tucker J. e Persily N. (ed.), Cambridge University Press.
- Siegel A. et al. 2019. *Trumping hate on Twitter? Online hate in the 2016 US election and its aftermath*. Working Paper.
- Silverman T. et al. 2016. *The impact of counter-narratives*. Institute for Strategic Dialogue, Londra.
- Smahel D. et al. 2020. *EU kids online 2020: Survey results from 19 countries*. London School of Economics and Political Science, Londra.
- Soral W., B. Michal e M. Winiewski. 2018. Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior* 44(2): 136-146.
- Stahel L. e C. Schoen. 2019. Female journalists under attack? Explaining gender differences in reactions to audiences' attacks. *New Media & Society*: 1461444819885333.

- Sticca F. e S. Perren. 2013. Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying. *Journal of Youth and Adolescence* 42: 739-750.
- Stiftung gegen Rassismus und Antisemitismus (GRA) und Schweizerischer Israelitischer Gemeindeverband (SIG). 2019. *Antisemitismusbericht*. Zurigo.
- Suler J. 2004. The online disinhibition effect. *Cyberpsychology & Behavior* 7(3): 321-326.
- Sunstein C. R. 2000. Deliberative trouble? Why groups go to extremes. *The Yale Law Journal* 110(1): 71-119.
- Tuck H. e T. Silverman. 2016. *The counter-narrative handbook*. Institute for Strategic Dialogue, Londra.
- Tynes B. M. e S. L. Markoe. 2010. The role of color-blind racial attitudes in reactions to racial discrimination on social network sites. *Journal of Diversity in Higher Education* 3(1): 1-13.
- Tynes B. M. et al. 2008. Online racial discrimination and psychological adjustment among adolescents. *Journal of Adolescent Health* 43(6): 565-569.
- Tynes B., J. Torro e F. Lozada. 2019. An unwelcomed digital visitor in the classroom: The longitudinal impact of online racial discrimination on school achievement motivation. *School Psychology Review* 44(4): 407-424.
- Ufficio federale di statistica. 2019. *Indagine sulla convivenza in Svizzera: risultati 2018*. UST, Neuchâtel.
- Vosoughi S., D. Roy e S. Aral. 2018. The spread of true and false news online. *Science* 359(6380): 1146-1151.
- Wachs S. et al. 2019. Associations between witnessing and perpetrating online hate in eight countries: The buffering effects of problem-focused coping. *International Journal of Environmental Research and Public Health* 16(20): 3992.
- Waqas A. et al. 2019. Mapping online hate: A scientometric analysis on research trends and hotspots in research on online hate. *PloS one* 14(9).
- Winter A. 2019. Online hate: From the far-right to the 'Alt-Right' and from the margins to the mainstream. In: *Online Othering, Palgrave Studies in Cybercrime and Cybersecurity*, Maras M. e T.J. Holt (rd.): 39-63. Palgrave Macmillan.
- Wolfe C. 2019. Online trolls, journalism and the freedom of speech: Are the bullies taking over? *Ethical Space: The International Journal of Communication Ethics* 16(1): 11-21.
- Yanagizawa-Drott D. 2014. Propaganda and conflict: Evidence from the Rwandan genocide. *The Quarterly Journal of Economics* 129(4): 1947-1994.
- Ziegele M. et al. 2019. *Aufräumen im Trollhaus: Zum Einfluss von Community-Managern und Aktionsgruppen in Kommentarspalten*. Institute for Internet and Democracy, Düsseldorf.
- Ziegele M., C. Koehler e M. Weber. 2018. Socially destructive? Effects of negative and hateful user comments on readers' donation behavior toward refugees and homeless persons. *Journal of Broadcasting & Electronic Media* 62(4): 636-653.

Articoli di giornale

- Berners-Lee T. 2017. [I invented the web. Here are three things we need to change to save it](#). *The Guardian* (12 marzo 2017).
- Fisher M. 2018. [Inside Facebook's secret rules for global political speech](#). *New York Times* (27 dicembre 2018).
- Germann M. 2018. [7000 gesperrte Kommentare](#). *WOZ* (20 settembre 2018).
- Gunaratna S. 2016. [Neo-Nazis tag \(\(\(Jews\)\)\) on Twitter as hate speech, politics collide](#). *CBS News* (10 giugno 2016).
- Lizza R. 2016. [Twitter's anti-semitism problem](#). *New Yorker* (19 ottobre 2016).
- Newton C. 2019. [The Trauma Floor](#). *The Verge* (25 febbraio 2019).
- Priebe M. 2020. [Wie Rassisten das Coronavirus für sich nutzen](#). *FAZ* (3 febbraio 2020).
- Serafini S. 2015. [Edelweiss-Streit: Auf Whatsapp blüht der jugendliche Patriotismus](#). *Aargauer Zeitung* (21 dicembre 2015).
- Spiegel.de. 2016. [Freundlicher Frosch wird Hasssymbol](#) (28 settembre 2016).
- Schweiz aktuell. 2020. [Ja zum Transitplatz für Fahrende in Wileroltigen](#). *SRF* (9 febbraio 2020).
- Stahel L. 2019. [Gehässige Leserreaktionen können die Qualität der Berichterstattung auch erhöhen](#). *NZZ* (7 maggio 2019).
- 20min. 2019. [Primarschüler machen sich über Dunkelhäutige lustig](#) (30 dicembre 2019).
- 20min. 2019. [«Rechtsextreme Chats gibt es an jeder Schule»](#) (19 marzo 2019).

Altre fonti in Internet

(i siti Internet di cui nel testo è stato menzionato soltanto il nome non sono più elencati qui di seguito)

- Alexander N. 2019. [Reconquista Germanica meldet sich ab](#). *Youtube* (30 novembre 2019).
- Anglin A. 2016. [A Normie's Guide to the Alt-Right](#). *Daily Stormer* (31 agosto 2016).
- Banaszczuk Y. 2019. [Toxic Gaming: Rassismus, Sexismus und Hate Speech in der Spieleszene](#). *Bundeszentrale für politische Bildung* (26 luglio 2019).
- Committee on the Elimination of Racial Discrimination CERD. 2013. [Recommandation générale no 35 sur la lutte contre les discours de haine raciale](#) (26 settembre 2013).
- Council of Europe. 1997. [Council of Europe's Committee of Ministers' Recommendation 97\(20\)](#) (30 ottobre 1997).

European Commission against Racism and Intolerance ECRI. 2015. [General Policy Recommendation N°15 on Combating Hate Speech](#) (8 dicembre 2015).

Garland C. 2008. [Klan's new message of cyber-hate](#). *Cult Education Institute* (27 marzo 2008).

Generation D. 2017. [Das Shitposting 1x1](#) (18 maggio 2017).

Gettethrollout!. 2018. «[Troll of the month](#)» (5 dicembre 2018).

Gold extra. [Tools of Subversion](#).

Heiderich K. 2018. [Extremismusforscherin Julia Ebner: «Hasskampagnen folgen einem klaren Muster»](#). *Fearless Democracy* (26 aprile 2018).

Humanrights.ch. 2017. [Hassaufrufe im Internet - Schweizer Fälle und Politik der Newsportale](#) (23 agosto 2017).

Humanrights.ch. 2017. [Hassreden: Die Grenzen der Meinungsäusserungsfreiheit](#) (6 febbraio 2017).

Humanrights.ch. 2020. [Rechtsentwicklung: Verstärkung des gesetzlichen Schutzes vor Diskriminierung?](#) (19 marzo 2020).

INACH. 2019. [The state of policy on cyber hate in the EU](#) (19 novembre 2019).

ISD Global. [Youth Civil Activism Network \(YouthCAN\)](#)

Jetzt. 2018. [Rechtsextreme «Feministinnen» stören die Berlinale](#) (20 febbraio 2018).

Lorentzen M. K. 2017. [Don't feed the trolls - Fight them](#). TEDx Talk. *Youtube* (7 gennaio 2020).

Pfeiffer J. 2020. [Hate Speech in der Incel-Szene](#). *Belltower News* (25 marzo 2020).

Reitman J. 2018. [All-American Nazis](#). *Rolling Stones* (2 maggio 2018).

sCAN. 2018a. [Hate Ontology](#).

sCAN. 2018b. [Mapping study: Countering online hate speech with automated monitoring tools](#). Ligue Internationale Contre le Racisme et l'Antisémitisme (LICRA).

sCAN. 2020. [Hotspots of Hate: The online responsibility of public figures](#).

sCAN. Ohne Jahr. [Monitoring Report 2018-2019](#).

Servizio per la lotta al razzismo SLR. [Guida giuridica on-line](#).

Servizio per la lotta al razzismo SLR. [Media e Internet](#).

Southern Poverty Law Center. 2018. [McInnes, Molyneux, and 4chan: Investigating pathways to the alt-right](#). (19 aprile 2018).

Southern Poverty Law Center. 2020. [Alt-Right](#).

Southern Poverty Law Center. 2015. [The Council of Conservative Citizens: Dylann Roof's gateway into the world of white nationalism](#) (21 giugno 2015).

Stiftung gegen Rassismus und Antisemitismus (GRA). [Hate Speech / Hassrede im Internet](#).

Stiftung gegen Rassismus und Antisemitismus (GRA). [Meldetool](#).

Wikipedia. [Metapedia](#).

Zivilcourage und Anti-Rassismus-Arbeit ZARA. [Hass im Netz melden](#)

Fonte del materiale grafico

Grafico 7: <https://derpicdn.net/img/2013/4/6/289767/full.png>

Grafico 8: <https://me.me/i/nazi-pony-tumblr-912b75e26c2345f8aea216a60a0723c9>

Indice delle interviste

Birdwell, Jonathan (Institute of Strategic Dialogue, London)

Bischof, Michael (Integrationsförderung, Stadt Zürich)

Pershan, Claire (Renaissance Numérique, Paris)

Pugatsch, Dominic (GRA, Zürich)

Spiess-Hegglin, Jolanda (#Netzcourage, Zug)

Voigt, Hansi (dasnetz.ch; bajour)