

**GRUPPO DI LAVORO
ODIO ONLINE
RAPPORTO FINALE**

5 febbraio 2021

**POST
NO
HATE**

SACSIX

ys.

one

BTM

Este di Delle...

Premessa	3
Definizioni	5
Fenomenologia	7
Esempi internazionali	9
Raccomandazioni	11
a. Attività di prevenzione	11
b. Interventi normativi.....	14
c. Sostegno all'infodiversità	18
Ipotesi Organizzative	19
Ringraziamenti	21

P R E M E S S A

Gli esseri umani sono liberi di provare sentimenti. L'odio è uno di questi. È una libertà inalienabile.

L'odio è il “sentimento di forte e persistente avversione, per cui si desidera il male o la rovina altrui”¹. Quando dal desiderio del male altrui si passa all'azione, per favorire o realizzare tale male, subentrano le responsabilità.

Internet ha reso possibile un dialogo globale ininterrotto nel quale miliardi di persone fanno esperienza della libertà di espressione e dell'amplificata diffusione dei loro messaggi, compresi quelli che manifestano odio. Gli strumenti digitali offrono nuove modalità per passare dall'espressione di odio all'azione tangibile, come si vedrà nelle prossime pagine². Le conseguenze di azioni di comunicazione online motivate dall'odio possono essere terribili per le vittime. Possono essere organizzate da chi voglia sfruttare le debolezze altrui per ottenere vantaggi economici o per aggregare consenso in chiave politica. Gli algoritmi dei media sociali possono diventare strumenti di manipolazione.

Le piattaforme compiono scelte importanti per quanto riguarda l'accesso all'informazione, la sua selezione e diffusione, e sempre più esplicitamente si prendono la responsabilità di interpretare a modo loro il bene della comunità degli utenti. Come dimostra il Digital Services Act recentemente proposto dalla Commissione Europea, i legislatori a loro volta si impegnano a definire una strategia sull'estremismo, la disinformazione, la manipolazione dei sentimenti, l'amplificazione di messaggi di odio che hanno ristretto lo spazio del discorso pubblico e aumentato la polarizzazione nella mediasfera. Il cambio di policy sui doveri delle grandi piattaforme che gestiscono la comunicazione digitale sembra orientato a richiamarle alle loro responsabilità attraverso un percorso di co-regolamentazione. I rischi degli interventi normativi in questa materia però sono importanti: regole pensate per proteggere la democrazia dagli estremismi favorendo un dibattito informato e rispettoso possono essere a loro volta distorte per inibire la libertà di espressione.

I legislatori devono dimostrare grande saggezza. Perché su questa materia si osservano diritti in tensione tra loro, come la libertà di espressione, appunto, il diritto alla privacy, il diritto al rispetto della libertà di pensiero, coscienza e religione, il diritto di proprietà e la

¹ Vocabolario Treccani

² È successo anche in passato con l'avvento di ogni nuovo strumento di comunicazione

libertà di mercato, il diritto a essere difesi contro le violenze, come ha osservato la Corte Europea dei Diritti Umani. Il problema fondamentale è quello di bilanciare i diversi diritti.

D'altra parte, qualunque policy in materia di odio online non può non considerare che ci sono diverse gradazioni della gravità dei comportamenti dettati dall'odio e diverse manifestazioni di odio: quelle che si concentrano contro specifiche persone e quelle che si focalizzano contro certe categorie sociali. Senza trascurare ogni altro problema - dalle sofferenze individuali, appunto, fino alle forme di discriminazione contro precise categorie di persone - l'argomento prioritario di una policy sull'odio online riguarda i casi in cui si arriva a generare una lesione dell'ordine pubblico.

Chiunque si presenti con la soluzione in tasca e una ricetta magica per contrastare le conseguenze dell'odio online, probabilmente, non tiene adeguatamente conto della complessità e della delicatezza del fenomeno³.

³ Ed è con questo spirito sensibile, prudente, aperto che il gruppo di lavoro sul fenomeno dell'odio online ha operato. Gli esperti riuniti nel gruppo hanno operato nella consapevolezza della complessità del fenomeno e dell'enorme insieme di sfaccettature che ogni aspetto del problema presenta. Le analisi e le raccomandazioni finali sono basate sulla vastissima letteratura prodotta da chi ha studiato il fenomeno, sull'esperienza di ricerca e di azione dei partecipanti al gruppo, sui consigli ricevuti nel corso di una chiamata per contributi aperta a tutti, e sulle discussioni tenute prevalentemente online nel corso dei mesi di lavoro del gruppo, concentrati essenzialmente tra ottobre 2020 e gennaio 2021. Per tutte queste considerazioni, speriamo che questo rapporto finale possa servire da stimolo allo sviluppo di un approccio consapevole anche in chi dovrà assumersi la responsabilità delle policy in Italia e in Europa sul tema dell'odio online. Con una sola certezza: nel quadro di uno stato di diritto che faccia valere con fermezza ed efficacia la legge nei confronti di chi compie reati approfittando delle debolezze altrui, va sempre tenuto presente che in democrazia, l'odio si combatte con l'amore per le molte, rispettabili facce della vita umana.

DEFINIZIONI

Il concetto di odio riguarda esperienze soggettive e dunque non standardizzabili, come suggerisce “Article 19”, l’organizzazione britannica per i diritti umani che difende e promuove la libertà di espressione nel mondo. Esprimere l’odio può rendere manifeste forme di profonda intolleranza e pregiudizio, fino alla negazione dell’umanità altrui. E tali espressioni possono costituire incitamento all’organizzazione di atti di violenza contro singole persone o intere comunità.

L’espressione di odio attraverso internet non riduce la difficoltà di definire il sentimento in modo standardizzato ma aggiunge un ulteriore elemento di complessità, per la capacità delle tecnologie digitali di influenzare i comportamenti delle persone e amplificarne la portata, creando nel contempo le condizioni per implementare peculiari modalità di contenimento.

Per l’Unesco l’odio online è un esempio di come Internet offra opportunità ma anche sfide difficili per quanto riguarda la libertà di espressione e la dignità umana. In “Countering Online Hate Speech” (2015)⁴, l’Unesco individua quattro differenze sostanziali tra l’odio offline e l’odio online. La prima è la permanenza dell’odio online, ossia la possibilità che rimanga “attivo” per lunghi periodi di tempo e in diversi formati. La seconda differenza è il ritorno imprevedibile delle espressioni di odio che, anche se sono rimosse da un luogo possono riapparire altrove. La terza differenza è l’importanza che assume online l’anonimato: di per sé un diritto che consente in certe condizioni di fare emergere verità che chi si trova in condizioni di svantaggio potrebbe aver timore di comunicare⁵, l’uso di pseudonimi e nomi falsi può rendere le persone meno consapevoli del valore delle proprie parole e ingenerare più o meno fondate aspettative di irresponsabilità e impunità. La quarta differenza è la transnazionalità, che aumenta l’effetto dell’hate speech e rende più complicato individuare i meccanismi legali per combatterlo.

⁴ Unesco, “Countering Online Hatespeech” (2015*), ([*https://unesdoc.unesco.org/ark:/48223/pf0000233231](https://unesdoc.unesco.org/ark:/48223/pf0000233231))

⁵ La protezione delle fonti è una delle pratiche tradizionali del giornalismo. La protezione del whistleblowing è entrata a far parte delle tecniche di lotta alla corruzione nella pubblica amministrazione. Le persone che si trovano in territori nei quali il crimine organizzato è particolarmente forte possono trovarsi a collaborare con la giustizia solo se la loro identità viene tenuta segreta. Le donne perseguitate da mariti violenti possono trovare il coraggio di denunciarli se il loro anonimato è garantito. E nei sistemi politici autoritari, l’anonimato è una forma di lotta indispensabile. Ma l’anonimità è importante non solo come strumento di difesa, ma anche per creare spazi liberi dalla pressione sociale per esprimere idee controverse, innovative, inattese. Cfr: Robert Bodle, “The ethics of online anonymity. Or Zuckerberg vs. Moot”, College of Mount St. Joseph, USA.

Qual è dunque l'approccio giusto al problema delle definizioni? Cercare comunque una definizione standard dell'odio online, oppure descrivere pragmaticamente il fenomeno per facilitarne l'individuazione e attivare le azioni di contenimento? Di fatto, il sistema giuridico nel suo complesso è attrezzato, in Europa, per riconoscere il fenomeno dell'odio e contrastare se è necessario le sue conseguenze e poiché, come osserva il Council of Europe⁶ «non esiste una definizione universalmente accettata del termine “hate speech”», il suo trattamento concettuale ricade nell'attività delle magistrature.

⁶ Council of Europe (<https://rm.coe.int/16800c170f>)

FENOMENOLOGIA

Da una prospettiva sociologica, per analizzare l'odio occorre tener conto di una molteplicità di fattori che vanno dalle condizioni politico-sociali delle specifiche realtà studiate, alle generali trasformazioni culturali e dei costumi, alle forme con le quali si aggregano i gruppi e si costruiscono le loro identità, il loro potere, la loro visione del mondo. A seconda delle diverse situazioni storiche, l'odio può diventare socialmente rilevante come strumento di consenso a favore di particolari gruppi di potere e può invece essere considerato come la conseguenza di un conflitto sociale teso alla lotta contro le ingiustizie, oppure può semplicemente essere un fatto individuale collegato a particolari condizioni educative, psicologiche, antropologiche, familiari.

I media digitali sono strumenti potenti e possono essere trasformativi anche per i fenomeni di odio. Non solo perché i media che tutti possono usare per esprimersi liberano e talvolta amplificano la voce di ciascuno senza le intermediazioni tradizionali, ma anche perché gli algoritmi e le interfacce influenzano l'esposizione, la selezione e la diffusione dell'informazione, diventando veri e propri filtri cognitivi alla percezione della realtà.

Gli intermediari della comunicazione, peraltro, mettono in campo consapevoli attività di moderazione dei contenuti, anche per mitigare la diffusione dei discorsi di odio in rete. Questi interventi avvengono con metodi automatici o ricorrendo a valutazioni di persone fisiche ma in ogni caso con criteri standardizzati e determinano la segnalazione, la rimozione o la disabilitazione all'accesso di milioni di contenuti e account degli utenti. Questa autoregolamentazione è pensata per ottenere il massimo di libertà di espressione per gli utenti e, insieme, il minimo di critica nei confronti delle piattaforme da parte di coloro che ritengono di voler contribuire alla limitazione della "tossicità" dei media sociali come gli inserzionisti pubblicitari, i governi o le associazioni a tutela dei diritti umani.

La conoscenza delle logiche di progettazione e manutenzione della tecnologia è ovviamente indispensabile per la comprensione del fenomeno dell'odio online. Sebbene nella sostanza non si discosti dall'odio offline, l'odio online si esprime in un contesto nel quale la comunicazione è veloce, si muove in spazi potenzialmente molto ampi, sviluppa linguaggi e tecniche retoriche peculiari, con forme di ostilità indirette o implicite, spesso difficili da intendere per chi non è direttamente coinvolto. Oltre alla mutevolezza delle forme di hate speech, occorre tenere conto dell'ampio numero di attori, dell'incertezza e talvolta indistinzione dei ruoli (produttori, diffusori, vittime) e della velocità delle dinamiche di propagazione. Tutti elementi che contribuiscono alla complessità del fenomeno.

Del resto, non è solo il contenuto dei messaggi d'odio postati o diffusi online a fare la differenza, ma anche il modo in cui essi transitano tra i diversi media, il modo in cui vengono trasformati e diffusi da innumerevoli soggetti tra loro diversi (cittadini, politici e così via) e gli effetti a catena che possono innescare, difficili da delimitare nel tempo, nello spazio e nel numero dei soggetti sui quali vi potrebbero essere ricadute.

In primo luogo, si possono osservare pratiche di produzione, circolazione e consumo di discorsi d'odio che hanno l'obiettivo di sfidare i tabù e le convenzioni della correttezza politica, in contrasto con le narrazioni dei media generalisti. In questi casi, c'è il rischio di normalizzare l'incitamento all'odio con la scusa che si tratti di ironia, una forma espressiva, apparentemente innocua, che può rafforzare gli stereotipi contro persone o gruppi deboli.

In secondo luogo, si possono manifestare pratiche comunicative nelle quali il vortice dell'incitamento all'odio può iniziare online e, talvolta, trasferirsi offline trasformandosi in un generale atteggiamento di intolleranza. È il caso per esempio dell'uso di tecniche tipiche dei social media, come il trolling e i meme virali, che degenerano in molestie mirate, *shit storm*, e così via.

In terzo luogo, ci possono essere casi in cui certi media tradizionali si prestano a funzionare da cassa di risonanza dei contenuti di odio online (e viceversa, ovviamente). Avviene allora che i meccanismi di amplificazione finiscano per “disegnare un bersaglio” su persone o gruppi etnici, rendendo più probabile che gli stessi discorsi d'odio siano usati in futuro, rendendo più visibili gli attori negativi, sovrastimando la consistenza delle loro opinioni, rendendo più efficaci gli intenti manipolatori e propagandistici. L'amplificazione dell'odio nelle versioni online di certi media tradizionali trova ragioni strutturali nel mercato della visibilità delle notizie e della raccolta pubblicitaria che induce a pubblicare qualsiasi contenuto possa attirare utenti.

Il tema dell'odio (online e non) è un terreno scivoloso: nel tener conto del difficile equilibrio tra diritto all'informazione, libertà d'espressione, rischio di censura, non si può non pensare a come l'amplificazione tecnica fornisca ossigeno alla generazione di un ambiente comunicativo tossico nel quale i discorsi d'odio diventano una merce comune.

ESEMPI INTERNAZIONALI

È finora mancata un'azione comune e concordata tra i Paesi europei. Non c'è un accordo sulla definizione di “hate speech” e sul concetto di istigazione all'odio online. Non c'è stato un coordinamento sulle modalità di contenimento del fenomeno, nella previsione di tempi certi di cancellazione dei messaggi o sul ruolo di garanzia della magistratura nella gestione della rimozione dei contenuti quando impatta sulla libertà di espressione.

La Germania è stato il primo Paese a intervenire, nel 2018, con una normativa specifica molto aggressiva nei confronti delle piattaforme e mirata soprattutto ad imporre maggiore collaborazione da parte di grandi social network. L'efficacia di quelle misure sarà sottoposta a verifica dopo tre anni dall'entrata in vigore per valutare eventuali criticità. Le domande sono molte: da chi provengono le segnalazioni, se utenti o autorità; quali sono i tempi di reazione; se si manifestano rimozioni di contenuti legittimi operati dalle piattaforme solo per evitare problemi. Tale verifica potrebbe fornire elementi di valutazione per comprendere se sia corretto affidare ai provider il ruolo di “giudice di primo grado” sullo hate speech.

La Francia, seguendo l'esempio della Germania, ha elaborato una normativa simile che, però, è finita sotto la scure del Consiglio Costituzionale. In altri Paesi, come Polonia, Spagna, Austria e Paesi del Nord Europa, sono in corso studi e iniziative politiche finalizzate non tanto a pratiche repressive ma a cercare di capire come “governare” e “mitigare” l'odio, magari creando consultori e luoghi di ascolto, fisici o virtuali, per le vittime e i responsabili di aggressioni, soprattutto se minori o adolescenti, mitigando così il ricorso alle sanzioni penali. Il tutto in un'ottica di collaborazione con le piattaforme.

Ma queste policy in ordine sparso sono destinate a trovare un coordinamento. Il 15 dicembre 2020 è stata resa pubblica dalla Commissione la Proposta di Regolamento del Parlamento Europeo e del Consiglio sul mercato unico dei servizi digitali⁷.

La proposta di Regolamento ha come obiettivo dichiarato quello di “stabilire regole uniformi per un ambiente online sicuro, certo e affidabile, dove i diritti fondamentali sanciti dalla Carta siano effettivamente protetti”⁸.

⁷ DigitalServiceAct -DSA -Proposta COM (2020) 825 2020/0361 (COD); nel gennaio 2021 la Germania ha varato una legge che salvaguarda la concorrenza nei mercati digitali e in un certo senso anticipa la nuova normativa europea <https://www.politico.eu/article/germany-shows-eu-the-way-in-curbing-big-tech/amp/>

⁸ Art.2 comma 2 lett.b

Per raggiungere questo obiettivo, che riguarda la diffusione in rete di tutti i “contenuti illegali” e dunque anche le manifestazioni di odio che per la loro lesività integrano fatti illeciti secondo le diverse leggi nazionali, la proposta di Regolamento stabilisce:

- a) un nuovo impianto normativo per l'esenzione di responsabilità dei prestatori di servizi di intermediazione, modificando la vecchia direttiva 31/2000/CE e-commerce e ridisegnando le condizioni di tale esenzione che nel Regolamento è definita “condizionata”.
- b) nuovi e specifici obblighi di necessaria diligenza ritagliati e adattati su determinate e specifiche categorie di fornitori di servizi di intermediazione;
- (c) nuovi meccanismi di cooperazione e coordinamento tra le autorità competenti per l'attuazione e il rispetto degli obblighi del regolamento.

Il nuovo Regolamento, se approvato dal Parlamento, costituirà una base normativa comune per tutti i paesi dell'Unione e troverà applicazione per tutti i servizi accessibili ai cittadini europei, indipendentemente dal paese di stabilimento del provider.

È indubbio che tale Regolamento, unitamente alla seconda proposta avanzata dalla commissione sempre il 15 dicembre scorso e volta a regolamentare il mercato digitale sotto i profili della concorrenza e dell'equità⁹ modificherà profondamente il governo dei contenuti illeciti online e influirà sulle future modalità di “autoregolamentazione” delle piattaforme.

L'art. 6 della proposta di Regolamento introduce infatti in Europa la clausola del Buon Samaritano che incentiva la moderazione dei contenuti da parte dei fornitori dei servizi, senza che tale attività “editoriale” possa minare l'applicabilità delle esenzioni di responsabilità.

⁹ Digital Market Act-DMA - COM(2020) 842 final 2020/0374 (COD)

RACCOMANDAZIONI

Per trovare soluzioni capaci di rispettare tutti i diritti umani, in un equilibrio che sui social network è apparso spesso messo in discussione, occorrono interventi che coinvolgano le persone, le famiglie, le scuole, le associazioni, le università e i centri di ricerca, le aziende, i media, i partiti, le istituzioni.

Si può immaginare una strategia articolata su tre livelli:

A. Azioni di prevenzione, con obiettivi di lungo termine, centrate sull'educazione civica e digitale, la cultura giuridica, la ricerca, l'informazione, la comunicazione;

B. Innovazione normativa capace di costruire un quadro giuridico adeguato all'ambiente digitale sia per le aziende private che per i cittadini che fruiscono dei servizi digitali. Un quadro nel quale, con metodo aperto, inclusivo e basato sull'evidenza siano chiaramente definiti ruoli e responsabilità nel pieno rispetto dei diritti fondamentali della persona;

C. Sostegno fattivo alle iniziative orientate a progettare, sperimentare e costruire nuove piattaforme e ambienti mediatici, in modo da favorire l'infodiversità nell'ecosistema digitale e aumentare le probabilità che emergano soluzioni adatte a favorire condizioni di vita online più rispettose dei diritti umani e del valore della conoscenza di qualità.

Se valutate positivamente, queste ipotesi strategiche richiederanno l'adozione di misure organizzative destinate alla loro implementazione.

La Commissione europea, in ogni caso, sembra muoversi in questa prospettiva strategica.

A. ATTIVITÀ DI PREVENZIONE

L'attività più profonda che si possa progettare per ottenere risultati nel lungo periodo è quella che incide sulla cultura, a livello di formazione, informazione e comunicazione. Il progetto di un grande programma educativo per la consapevolezza digitale in Italia può rivelarsi strategico, ovviamente non solo per il tema dell'odio online. Come mostrano i dati registrati dal Digital Economy and Society Index¹⁰, l'Italia è nelle ultime posizioni su tutte le variabili che misurano l'avanzamento digitale dell'economia e della società nei Paesi europei,

¹⁰ <https://ec.europa.eu/digital-single-market/en/digital-economy-and-society-index-desi>

ma è ultima dal punto di vista del “capitale umano”: è ultima in quanto a capacità basilari ed è ultima in quanto a presenza di specialisti. Questi dati possono essere letti come una proxy dell’arretratezza culturale generale sul tema digitale, che rende anche meno facile per la società italiana il raggiungimento di una conoscenza dei pregi e dei limiti delle tecnologie. In questo modo, le probabilità di comportamenti poco consapevoli, nell’uso quotidiano degli strumenti digitali, aumentano. Un programma educativo per l’epoca digitale, di pari ampiezza di quello che ha reso possibile accompagnare l’industrializzazione nel corso del “miracolo economico” è in gran parte necessario. E non si rivolgerà soltanto ai giovani, ma all’insieme della società. Così, troveranno spazio, profondità e motivazione anche le iniziative educative più specificatamente orientate a prevenire le condizioni di debolezza cognitiva che possono costituire contesti poco difesi dalle espressioni di odio online. Le strutture prioritariamente coinvolte saranno: la scuola, la famiglia, le aziende, i media.

Il Ministero dell’Istruzione può lanciare un grande programma di modernizzazione culturale di base rafforzando il percorso di educazione all’uso consapevole del digitale previsto nel quadro dell’insegnamento dell’educazione civica introdotto nel 2019¹¹ (per esempio, con l’introduzione nel curriculum scolastico di un’ora settimanale di cittadinanza digitale per le scuole di ogni ordine e grado a partire dall’infanzia) e investendo molto più fortemente nella formazione dei formatori. Nel corso dell’ultimo anno, 70mila insegnanti hanno ricevuto una formazione digitale, come si ricorda al ministero per l’Innovazione tecnologica e la digitalizzazione. Ma il programma va generalizzato: sia per alimentare un uso creativo delle soluzioni digitali - anche in vista di una modernizzazione della didattica che le tecnologie possono favorire - e sia per accrescere la consapevolezza degli insegnanti sulle opportunità e i rischi dell’uso che si può fare del digitale nella vita quotidiana.

Naturalmente, la scuola può servire anche per favorire il coinvolgimento delle famiglie. Ma un programma culturale orientato alla consapevolezza digitale deve coinvolgere gli italiani tutte le età. E quindi va favorita la collaborazione al progetto delle aziende e dei media. Riportiamo qui sotto una raccolta di proposte, tutte da approfondire.

Scuola

- Coinvolgimento di animatori digitali, team innovazione e responsabili cyberbullismo nella individuazione di buone pratiche da selezionare e rendere disponibili, con licenze aperte o creative commons, al fine di replicarle in altri contesti.
- Dopo attenta valutazione dei risultati della ricerca nelle materie di riferimento, progettazione di percorsi educativi “riabilitativi” che il MI potrebbe proporre ai

¹¹ Articolo 5, Legge 20 agosto 2019, n. 92: Educazione alla cittadinanza digitale. A quanto risulta, la previsione di 33 ore per tutta l’educazione civica lascia uno spazio di sole 11 ore per la cittadinanza digitale: questo spazio può essere allargato.

dirigenti scolastici come alternativa alle tradizionali sanzioni (note, espulsioni), rivedendo i Regolamenti di Disciplina.

- Realizzazione di una campagna nazionale di prevenzione realizzata dai giovani per i giovani, coinvolgendo il maggior numero di contesti educativi locali (Scuola, Associazioni sportive, Centri di aggregazione, Oratori)
- Immaginare le coperture per l'attuazione delle azioni di cui sopra attraverso il ricorso alla Cassa delle Ammende¹²

Famiglie

- Proposta per le scuole da parte del MI del riconoscimento di crediti formativi agli studenti delle famiglie coinvolte nelle attività di formazione su questi temi.

Aziende

- Premiare le aziende che attivano percorsi di formazione rivolti alle famiglie sui temi della consapevolezza digitale e in particolare sul contrasto ai discorsi di odio e a tutti i reati legati all'espressione (diffamazione, incitazione alla violenza e così via).
- Sostenere la formazione sul tema con l'integrazione di fondi professionali o di formazione del personale aziendale

Media

- Proporre all'Ordine Nazionale dei Giornalisti l'introduzione di una quota obbligatoria di crediti per la formazione continua dei giornalisti sul tema del contrasto all'hate speech
- Proporre all'Ordine Nazionale dei Giornalisti l'integrazione del codice deontologico dei giornalisti con uno specifico articolo che possa vietare azioni mirate alla diffusione di discorsi di odio
- Introduzione di una strategia di contrasto alla violenza sui media, con esemplari disposizioni da inserire nel contratto di servizio che regola il servizio pubblico, della Rai o delle aziende incaricate di svolgerlo.
- Condizionamento del finanziamento pubblico ai giornali al rispetto dell'obbligo - da introdurre - di astenersi dal ricorso a campagne di incitamento all'odio e dall'uso di un linguaggio discriminatorio.
- Pubblicità progresso e campagna nazionale di comunicazione per sensibilizzare alla responsabilità di ciascuno.

¹² https://www.giustizia.it/giustizia/it/mg_1_29_29_0_1_3.page ; https://it.wikipedia.org/wiki/Cassa_delle_ammende

B. INTERVENTI NORMATIVI

Strategia repressiva

Un controllo massivo dello Stato e della pubblica autorità che tenda a governare e filtrare il discorso pubblico in rete che travalichi la legittima e doverosa attività di prevenzione e repressione delle specifiche fattispecie di reato previste per legge sarebbe ovviamente inammissibile: la libertà d'espressione è principalmente, come recita l'art.11 della Carta dei Diritti Fondamentali dell'Unione Europea, "la libertà di ricevere o di comunicare informazioni o idee senza che vi possa essere ingerenza da parte delle autorità pubbliche".

Intervenire con politiche repressive su di un fenomeno quale l'odio on line, che come visto è magmatico ed incerto, ipotizzando nuovi illeciti è assai complesso e rischioso. L'illegalità di una comunicazione, indipendentemente dalla sua forma espressiva e dalle modalità di diffusione, risiede sempre, a fronte del diritto fondamentale alla libera manifestazione del pensiero (e dei sentimenti), nella concreta lesione di specifici beni giuridici che debbono esser determinati e ben profilati nella realtà. Tali beni giuridici sono usualmente già tutelati e presidiati da adeguate fattispecie di reato. I reati d'odio razziale etnico e religioso ed in futuro quelli legati all'identità di genere e alle possibili differenze identitarie atte a generare discriminazioni di gruppi o comunità tutelano l'ordine pubblico, e dunque la dimensione collettiva, al pari dei reati di apologia e di istigazione previsti dall'ordinamento. I diritti della persona e le libertà individuali trovano, anche in rete, piena tutela nei reati previsti dal codice penale che hanno subito adeguamenti, usualmente con specifiche aggravanti nella loro dimensione online (si pensi alle molestie, alle minacce, ai reati di atti persecutori, alla diffamazione ed ai delitti legati alle nuove tecnologie quali il recente reato di diffusione illecita di immagini o video sessualmente espliciti). Le dinamiche proprie dell'ambiente on line difficilmente determinano nuove lesioni non presidiate e tutelate da norme repressive. Ciò che è illegale online è illegale offline e viceversa. Certamente l'uso delle piattaforme di social network ed in generale il potenziamento della libertà di espressione dovuto a internet e il carattere pervasivo e permanente della comunicazione digitale può generare nuove forme e nuove manifestazioni di aggressione a beni giuridici tradizionali già presidiati o può aggravarne la lesione e le conseguenze, ma tali caratteristiche della dimensione online sono già previste o possono esser correttamente interpretate dalla magistratura sulla base dei principi generali propri del diritto penale.

Sul punto:

- l'ordinamento già determina quali contenuti sono illeciti e quali comunicazioni in rete consentono da parte dello Stato interventi di rimozione e sanzione; ipotizzare una strategia repressiva con nuove forme di illeciti legati all'odio online appare arduo e inutile¹³.

Strategie rieducative / sanzionatorie

Per efficaci politiche di contenimento potrebbe esser utile creare centri di ascolto e percorsi assistiti per generare consapevolezza negli utenti resisi responsabili di condotte qualificabili come "hate speech", soprattutto se minori o adolescenti. Anche per le vittime potrebbe esser utile prevedere percorsi di supporto per cercare di attenuare l'impatto subito da condotte di odio online.

- In caso di procedimenti per fattispecie penali legate all'odio on line, potrebbero esser previste forme di definizione del processo con percorsi di recupero ed educativi, o forme di messa alla prova con condotte riparatorie in grado di estinguere il reato.

Strategia preventiva, di contenimento e moderazione: il ruolo delle piattaforme

Una strategia volta a governare i discorsi d'odio on line non può non coinvolgere appieno gli intermediari della comunicazione, ovvero tutti quei soggetti che, tramite i loro servizi di condivisione, diffondono contenuti generati dagli utenti del web.

Una prima considerazione, forse banale, crediamo sia necessaria: il mezzo, internet, e dunque i servizi della società dell'informazione che veicolano i contenuti degli utenti, non sono i responsabili del fenomeno dell'odio online e la criminalizzazione tout court delle piattaforme di social network è un errore prospettico che può condurre a soluzioni errate e controproducenti. La proposta della Commissione Europea contenuta nel Digital Service Act che modifica il regime di responsabilità degli intermediari della comunicazione all'art. 7 (*No general monitoring or active fact-finding obligations*) ribadisce correttamente l'assenza di obblighi di monitoraggio e controllo preventivi pur rimodulando il sistema della (ir)responsabilità per i contenuti di terze parti.

¹³ L'unico punto che potrebbe forse esser oggetto di approfondimento sotto il profilo repressivo riguarda una lacuna di tutela in relazione alla diffusione illecita di dati personali: il c.d. *doxing* è costituito dalla diffusione di singoli dati personali per fini di nocimento del soggetto interessato ed è spesso connesso alle c.d. "tempeste mediatiche" (più efficacemente definite in inglese *shit-storm*) che possono travolgere singoli utenti rei di aver suscitato una qualche forma di facile indignazione. In Italia tale condotta, la diffusione non autorizzata di singoli dati personali spesso trattati da privati e dunque esclusi dall'applicazione del GDPR, non è più presidiata da sanzione dopo l'adeguamento del Codice Privacy ad opera del D.L.vo 101/2018. Sul punto si suggerisce un approfondimento tecnico.

Ciò precisato, è indubbio che gli intermediari della comunicazione giochino un ruolo essenziale nella diffusione (anche) dei contenuti potenzialmente dannosi: la struttura, il modello di business, i termini del servizio e le stesse interfacce utente dei diversi servizi indubbiamente modificano le forme di manifestazione e diffusione dell'odio amplificandone e talvolta falsandone la percezione da parte degli utenti, tanto per le vittime quanto per i responsabili. Il ruolo delle piattaforme, in particolare là dove queste si trovino in posizione dominante rispetto al dibattito pubblico, risulta centrale. Sotto questo profilo il tema della collaborazione delle piattaforme e della responsabilità, intesa come dovere di collaborazione e di rispetto delle regole di sistema stabilite per gli intermediari della comunicazione, è essenziale nelle politiche volte a mitigare e prevenire le degenerazioni dell'odio.

Se è errato pensare di delegare alle piattaforme ed a provider privati il ruolo di “gatekeeper” che decidono ciò che è consentito e ciò che non è consentito comunicare pubblicamente, è però necessario incentivare la moderazione dei contenuti caricati dagli utenti a opera degli stessi provider, per permettere alle piattaforme la creazione di ambienti digitali più sicuri.

E se è vero che solo lo Stato, la magistratura, può qualificare come illecito un contenuto d'odio foriero di danno - e solo nel caso sia illecito lo Stato può imporre la rimozione - è altresì vero che

- la piattaforma deve essere libera, sulla base di regole chiare e condivise, e con procedure trasparenti e controllabili, di moderare ed eliminare i contenuti ritenuti dannosi, indipendentemente dal fatto che questi siano illeciti o che vi sia un ordine della magistratura. Tale attività deve poter essere svolta senza che ciò possa minare la natura di mera intermediazione del servizio e il conseguente regime di irresponsabilità¹⁴;
- l'introduzione nell'ordinamento della c.d. “clausola del buon samaritano”, proposta anche dal Digital Service Act della Commissione Europea, che esclude la responsabilità del provider per le scelte editoriali in buona fede legate alla moderazione sui contenuti di terze parti - norma da tempo presente negli Stati Uniti - è essenziale per richiedere la collaborazione delle piattaforme e dovrebbe essere un obiettivo del Governo Italiano.

In vero, la più volte citata proposta di Regolamento Digital Service Act presentata in data 15 dicembre dalla Commissione Europea contiene alcune proposte che costituiscono valida base normativa per l'attuazione delle raccomandazioni che questo gruppo di lavoro ha individuato quali punti strategici nel contenimento dei contenuti dannosi e di odio sul web.

¹⁴ cfr. <https://santaclaraprinciples.org>

Ferma l'irresponsabilità delle piattaforme per i contenuti veicolati – elemento essenziale per garantire la libera circolazione di idee e informazioni - deve essere creato un quadro normativo articolato che imponga ai fornitori di servizi intermediari specifici e modulati modelli organizzativi e concreti doveri di diligenza, trasparenza e collaborazione.

In particolare sarebbe opportuno prevedere:

- Regole certe e standardizzate per la segnalazione e la rimozione dei contenuti ritenuti dannosi, eventualmente agevolando accordi con “segnalatori privilegiati” individuati in relazione al ruolo o all'attività svolta e obblighi di fornire informazioni all'autorità giudiziaria;
- Chiare tutele, con possibilità di celere contraddittorio, per gli utenti il cui contenuto sia stato cancellato dalle piattaforme a seguito di notifica o per violazione dei termini di servizio contrattualmente stabiliti, con la predisposizione di chiari punti di contatto per gli utenti e l'indicazione di organismi (umani!) indipendenti per la risoluzione extragiudiziale delle controversie;
- Misure di trasparenza ad ampio raggio, diversificate in relazione al tipo di servizio offerto: nella predisposizione dei termini contrattuali e delle condizioni previste per la rimozione o la restrizione di determinati contenuti; nella realizzazione di report periodici circa i contenuti e le informazioni rimosse o disabilitate; sugli algoritmi utilizzati per proporre e consigliare i contenuti agli utenti e sugli strumenti legati alla profilazione e alla pubblicità personalizzata, nonché sull'uso di strumenti automatizzati per la rimozione di contenuti;
- Nuovi obblighi per piattaforme in posizione dominante di intraprendere valutazioni basate sul rischio di abuso dei loro servizi e di predisporre strumenti di mitigazione degli abusi;
- Creare nuovi poteri in capo ad Authority indipendenti (AGCOM) per esaminare il funzionamento delle piattaforme, anche facilitando l'accesso dei ricercatori ai dati chiave della piattaforma;

Tali principi sembrano trovare oggi riconoscimento nella Proposta di Regolamento¹⁵ Digital Services Act recentemente presentata dalla Commissione Europea e che impegnerà le istituzioni dell'Unione nei prossimi mesi.

¹⁵ COM (2020) 825 2020/0361 (COD)

C. SOSTEGNO ALL'INFODIVERSITÀ

È possibile che la struttura di una piattaforma sia più favorevole di un'altra alla diffusione di comportamenti e manifestazioni di odio? Sicuramente la preconditione principale è che ci siano altri odiatori. Ma come si affermano le abitudini di chi si lascia andare a esprimere odio? Secondo alcune autorevoli ipotesi, un contesto nel quale la struttura della piattaforma, i suoi algoritmi, la sua interfaccia, sia definita essenzialmente intorno al modello di business della raccolta pubblicitaria, nel quale dunque qualunque tipo di attenzione è un valore che si può vendere sul mercato della pubblicità, può tendere ad accettare qualche messaggio tossico in più. Anche per questo, a quanto pare, certi comportamenti sono stati consentiti sulle grandi piattaforme sociali attuali. Per l'enorme successo delle piattaforme sociali basate esclusivamente sulla pubblicità, una sorta di monocultura pubblicitaria ha ridotto la biodiversità in questo settore. Una sorta di esternalità negativa del modello di business pubblicitario ha aumentato le probabilità, se così si può dire, di un "inquinamento" dell'ecosistema mediatico.

- Una strategia di lungo termine può essere orientata a favorire la diversità di modelli di business nei media sociali. Nuove piattaforme, progettate in base a logiche più consapevoli dei diritti umani, possono essere favorite con forme di sostegno finanziario, ma anche con la creazione di un contesto normativo più fortemente guidato dall'antitrust, con deconcentrazione del possesso dei dati e sostegno all'interoperabilità a tutti i livelli, nel quadro di una forte e netta neutralità della rete. In questo modo, le nuove piattaforme non sarebbero troppo penalizzate dall'effetto rete che favorisce le grandi imprese esistenti e potrebbero anzi trarre vantaggio dall'esistenza di altri servizi che hanno accumulato informazioni riutilizzabili da parte dei nuovi entranti sul mercato.
- Antitrust, privacy, fisco, potrebbero essere altrettanti quadri normativi all'europea che se ben interpretati potrebbero costituire un nuovo terreno di gioco per nuove piattaforme strutturate per essere interoperabili e non giardini chiusi.

Si può pensare che tra i tanti progetti così emergenti, alcuni sarebbero particolarmente adatti a contenere la diffusione dell'odio, a ridurre l'effetto amplificatore dei rimbalzi da tra un account e l'altro, oltre che tra un medium e l'altro. Non solo per via algoritmica, ma anche per via di aggregazione di persone che abbiano come loro atteggiamento mentale e sociale quello di evitare, rifiutare, disprezzare, le manifestazioni di odio. L'indefinibilità dell'odio e dei contesti semantici dei messaggi rende difficile trattare algoritmicamente la questione. Ma l'influenza sul comportamento garantito da ambienti che attraggono

soprattutto persone civili, rispettose degli altri, orientate all'ascolto e al dialogo sereno anche quando conflittuale o critico, sarebbe invece un forte incentivo ad adottare comportamenti civicamente avvertiti.

L'ipotesi sottostante è che una correzione dei media esistenti e comunque un maggiore equilibrio tra le possibilità offerte ai cittadini possa emergere con la creazione di alternative che siano disegnate per contrastare l'inquinamento mediatico.

- Progetti per questo possono essere sviluppati in chiave europea, a partire dalle nuove norme orientate all'antitrust dei dati e all'interoperabilità delle piattaforme, anche grazie alle conseguenze del GDPR. Un'iniziativa che lanci la creatività, l'investimento, la credibilità dei tentativi di far nascere nuovi media sociali caratterizzati da concetti fondativi diversi da quelli esistenti, potrebbe riservare le migliori sorprese. Far partire dal mondo della scuola, dei musei e delle biblioteche, le piattaforme della conoscenza di qualità, tali da far sentire inadeguati e fuori posto i comportamenti tossici e volgari, sarebbe per esempio un tentativo da sostenere. Rilanciare una qualità nelle relazioni sociali in nuove piattaforme puntando non tanto su forme di protezione algoritmica ma piuttosto sulla qualità dei progetti che attraggono le persone a frequentarle sarebbe forse la soluzione con le maggiori probabilità di successo. In fondo, Wikipedia dimostra che intorno a un progetto comune di valore si possono sviluppare comportamenti virtuosi.

IPOTESI ORGANIZZATIVE

Una volta valutate sul piano politico e tecnico queste proposte, si ipotizza l'istituzione di gruppi di lavoro operativi per passare all'azione. In qualche caso si potrà ricorrere a centri di ricerca e a organizzazioni non profit per la manutenzione delle misure proposte.

- Probabilmente saranno necessari gruppi operativi di lavoro per progettare le attività di formazione, comunicazione, partecipazione alle attività di co-regolamentazione che saranno avviate a livello europeo.
- Uno schema di collaborazione tra le autorità preposte alle garanzie nelle comunicazioni, alla protezione dei dati personali, all'antitrust, sarà con ogni probabilità necessario per la tenuta dei risultati della co-regolamentazione. Il tema dell'odio online farà probabilmente dei temi che un approccio olistico all'ecosistema digitale dovrà tener presente.

- Probabilmente sarà necessario stabilire un sistema di linee guida e di forme di valutazione per l'accREDITAMENTO dei centri di ricerca e delle organizzazioni non profit che potranno candidarsi a monitorare le attività delle piattaforme, la qualità dei risultati delle norme previste, le novità emergenti nel settore dell'informazione tossica, l'innovazione in materia che si può osservare a livello internazionale, la valutazione delle nuove piattaforme in termini di potenziale risposta all'esigenza di miglioramento della qualità dell'ecologia dei media in termini di riduzione dell'incidenza e della viralità dell'informazione tossica¹⁶.

¹⁶ Sarà probabilmente importante un coinvolgimento dell'ISTAT e, si spera, dell'Eurobarometro per definire rilevazioni periodiche e standard di misurazione sullo specifico della percezione e dell'esperienza del fenomeno.

RINGRAZIAMENTI

Il gruppo di lavoro sull'odio online ha lavorato per raggiungere gli obiettivi indicati nel decreto istitutivo: analisi e studio del fenomeno dell'odio online (monitoraggio, rilevazione e confronto su casi concreti); analisi e valutazione delle misure correttive attualmente previste; analisi e individuazione delle possibili misure da adottare per limitare e contrastare il fenomeno dell'odio online; redazione di una relazione conclusiva contenente i risultati dell'analisi dei precedenti punti e una o più proposte concrete per contrastare il fenomeno dell'odio online, in vista dell'adozione di atti di indirizzo da parte del Ministro per l'innovazione tecnologica e la digitalizzazione o della formulazione di proposte normative.

Hanno partecipato alla realizzazione di questo rapporto i membri del gruppo di lavoro sull'odio online: Giovanni Boccia Artieri, Sara Bentivegna, Carlo Blengino, Gioele Brandi, Valter Campanile, Agnese Nadia Canevari, Giovanna Cosenza, Luca De Biase, Alessandra De Marco, Juan Carlos De Martin, Lorenzo Del Giudice, Stefano Epifani, Simona Genovese, Paolo Iabichino, Triantafillos Loukarelis, Anna Maserà, Sonia Montegiove, Martina Pennisi, Laura Pertici, Walter Quattrociochi, Emma Rizzato, Rosy Russo, Ferruccio Sepe, Ilaria Sotis, Giovanni Ziccardi.

Hanno collaborato inviando proposte e documenti: AlmavivA; Roberto Bortone dell'UNAR; Cinzia Boschiero di ECPartners; Silvia Brena di Vox - Osservatorio italiano sui Diritti; Eva Campi di Newton Consulting; Maria Lucia Caspani di SocialExperience; Elena Castellini di SEC Newgate; Giuliano Castigliego; Martina Chichi di Amnesty International Sezione Italia; Diego Ciulli di Google, Julia Creet, York University, Mario De Ghetto, Francesco Di Costanzo di PA Social; Forma.Azione srl; Federico Faloppa della Rete nazionale per il contrasto ai discorsi e ai fenomeni d'odio, Miriam Frigerio di Sorgenia SpA; Giovanni Grandi dell'Università di Trieste; Bianca Teresa Iaccarino del Liceo G.B. Vico di Napoli; Michele Kettmajer di Mediavox; Elena Lamberti e Anna Maria Lorusso dell'Università di Bologna; Stefano Nicoletti di #Playthenow; Oryza Srl, Viviana Patti dell'Università di Torino; Paola Picierno, Parlamentare Europea; Nicoletta Polliotto di Muse Comunicazione; Cristina Raso di La Piazza Comune; Maria Luisa Stasi di ARTICLE 19; Cristina Tagliabue; Ivano Zoppi della Fondazione Carolina.

Foto di copertina: [Jon Tyson](#) on [Unsplash](#)