



Centro di ricerca sulle amministrazioni pubbliche "Vittorio Bachelet"

**LABORATORIO DI MONITORAGGIO E ANALISI
SULLE DISCRIMINAZIONI *ON LINE***

SOFT LAW, BUONE PRATICHE E POLICIES DEI SOCIAL NETWORKS

ottobre 2014

Premessa: delimitazione dell'ambito di ricerca

La scelta di indagare il settore delle cosiddette “buone prassi” in materia di discriminazione on line deriva essenzialmente dalla necessità di ricomprendere in questa macro categoria tutta una serie di strumenti - che vanno dalle linee guida, ai codici di autoregolamentazione, alle dichiarazioni di intenti, alle prassi amministrative, quali protocolli d'intesa, piani d'azione ed altro ancora - che hanno in diverso modo ed in varie occasioni affrontato il tema che, al contrario, risulta spesso difficoltoso sistematizzare attraverso apposite previsioni di carattere propriamente normativo.

La *self-regulation*, dunque, è stata interpretata e utilizzata in questo caso come strumento in grado di migliorare la qualità della regolazione - caratterizzata tra l'altro da una maggiore specificità ed uniformità rispetto alla legge - nella tutela, nella salvaguardia dei diritti di una determinate categorie di individui.

L'assunto – riconosciuto in dottrina – da cui si parte è, dunque, che l'ordinamento giuridico non si esaurisce nella produzione di atti normativi voluti dall'apparato statale, bensì prevede l'opportunità di essere alimentato, arricchito di atti dal carattere complementare e, per certi versi, innovativo, come appunto le fonti di autoregolamentazione. Naturalmente, ci si riferisce a tutti quegli strumenti che nascono con l'obiettivo di produrre effetti applicabili ad un determinato settore, in virtù della primaria legittimazione proveniente dalle leggi nazionali e dal necessario recepimento delle vigenti prescrizioni di carattere internazionale e dell'Unione europea.

Sulla base di studi e ricerche di carattere giuridico è facile sostenere come, negli ultimi decenni, l'attenzione manifestata verso gli strumenti di *soft regulation* sia stata senz'altro crescente.

Si è assistito ad una chiara tendenza allo sviluppo di queste nuove forme di regolazione, ancor più forte oggi, che nel dibattito dottrinale si parla di “ritorno alla regolazione”, di “rafforzamento dei poteri in mano alle autorità di regolazione” (per tutti, M. Clarich), nonostante tali strumenti presentino, accanto a profili d'indubbio interesse, anche aspetti problematici sul piano dell'efficacia.

Come già accennato, il ricorso a tali strumenti gode innanzitutto di una legittimazione a livello di Unione europea: a partire dall'accordo interistituzionale del 2003 “Legiferare meglio”, tra Parlamento, Consiglio e Commissione, l'Unione europea ha definito il quadro e le condizioni generali a base della politica legislativa degli Stati Membri, prevedendo e regolamentando anche il ricorso a metodi di regolamentazione “alternativi”, a meccanismi di co-regolamentazione e di autoregolamentazione che la Commissione stessa ha definito “espressione di una nuova cultura legislativa”.

A proposito di Internet, in vista dell'approvazione della Direttiva 2007/65/CE in materia di servizi di media audiovisivi, si considera rilevante l'intervento della Commissione europea sul tema. La Commissione, infatti, con la Comunicazione “Una migliore regolamentazione per la crescita e l'occupazione nell'UE”, ha evidenziato la necessità di valutare con attenzione l'approccio normativo più appropriato, “al fine di determinare se per ciascun settore e problema sia preferibile un atto legislativo oppure se debbano essere prese in considerazione soluzioni alternative, come la coregolamentazione o l'autoregolamentazione”. Su argomentazioni analoghe si fonda il contenuto della recente pubblicazione dell'OSCE, l'*On line Media Self-Regulation Guidebook 2013*, che avvalorata l'importanza della *self-regulation* per l'industria mediatica nell'era digitale.

Con riguardo all'Italia, negli ultimi anni, parte della dottrina ha sostenuto come, in taluni casi, l'autoregolamentazione, oltre ad essere vista come una forma di *de-regulation*, possa effettivamente costituire uno di quegli strumenti di semplificazione normativa in grado di rispondere all'esigenza di implementare la qualità della regolazione nel nostro Paese, integrando l'ordinamento giuridico attraverso la formazione di nuove regole "auto prodotte"; esigenza, questa, che si avverte fortemente in un settore come quello delle discriminazioni, caratterizzato da un evidente vuoto normativo nella legislazione nazionale.

"Autoregolazione" – dunque – come "capacità di produrre regole di diritto da parte degli stessi destinatari". Forme di regolazione *soft*, quali i codici di condotta e di comportamento promossi dai "colossi" del *social networking* e delle *web company* in generale, da intendersi, da una parte, come meccanismo di regolazione del mercato che consente di "superare" – nel senso di anticipare – il ricorso a norme "esterne"; dall'altra, come capacità, per molti versi innovativa, di incidere anche rispetto alla tutela di posizioni soggettive di rilevanza costituzionale nel sistema italiano.

Tra i fattori che incoraggiano il ricorso a tecniche di autoregolazione, tra le potenzialità di questi strumenti, vi è, senza dubbio, la maggiore flessibilità e rapidità che ne caratterizza l'adozione rispetto alla legge statale (fattori, questi, che giocano un ruolo determinante in un settore in continua evoluzione come quello delle nuove tecnologie e dei servizi di media, spesso caratterizzato, inoltre, dalla variabile dell'a-territorialità.). Da questo punto di vista, la *self-regulation* costituisce un indubbio vantaggio per l'impresa, in termini di: compressione del fattore temporale, ovvero di riduzione della complessità di formazione della regola (dunque semplificazione) e conseguente diminuzione degli oneri burocratici: obiettivo, questo, che risulta essere tra le priorità d'azione del programma della Commissione europea in materia di *Smart regulation*.

Tra i punti di forza dell'autoregolamentazione è riconoscibile la capacità effettiva, ove prevista, di produrre non solo regole, ma anche sanzioni – perlomeno economiche se non giuridiche – specie nel caso in cui venga riconosciuto ad appositi organismi competenti un potere di verifica e di controllo rispetto all'applicazione delle norme.

Al contrario, la debolezza di tali strumenti può risiedere nell'assenza di meccanismi sanzionatori – ove non previsti –, nella mancanza di appositi organismi competenti cui riconoscere un potere di verifica e di controllo. Da ciò deriva essenzialmente la discutibile efficacia della *self-regulation* da parte della dottrina, per quanto sia nota l'impossibilità di generalizzare riguardo la loro inefficacia *tout court*.

Nell'orientare la ricerca di questa tipologia di strumenti tesi a prevenire e contrastare le discriminazioni on line, nei diversi fattori in cui esse si articolano (genere, etnia, età, ecc.), in questa prima fase del lavoro, si è scelto di focalizzare l'attenzione in particolare su due "filoni" specifici di discriminazione: "genere" ed "età". Per questo, nel report si darà conto, principalmente, di strumenti e misure – adottate tra i mesi di maggio-settembre 2014, a contrasto delle discriminazioni online – che hanno per oggetto la prevenzione ed il contrasto del fenomeno dell'omofobia e della transfobia, da una parte, e del bullismo in rete, dall'altra (si vedano le segnalazioni riportate in fondo al documento).

Prevenzione contrasto omofobia

Con riferimento al contrasto e alla prevenzione dell'omofobia, in questo quadro di ricerca, vale ad esempio la pena richiamare,, nel periodo di riferimento preso in considerazione, la *Dichiarazione europea contro l'omofobia e la transfobia*, firmata anche dal Governo italiano - assieme ad altri 16 governi, tra cui quello maltese, governo ospitante - in occasione della Giornata internazionale contro l'omofobia e la transfobia, che si celebra il 17 maggio di ogni anno. Si tratta di una Dichiarazione d'Intenti di matrice europea, per promuovere l'uguaglianza, la tutela dei diritti umani delle persone LGBTI e la garanzia delle loro libertà fondamentali; temi, quelli della discriminazione omofobica e transfobica, che presentano non poche implicazioni anche rispetto all'utilizzo del web.

Da questo punto di vista, infatti, per quanto la Dichiarazione in sé non presenti un esplicito riferimento alle discriminazioni on line, essa costituisce oggi, per il nostro Paese, lo strumento di riferimento per l'attuazione della Strategia nazionale LGBT che, al contrario, richiama in diverse occasioni l'utilizzo di Internet. La Rete appare infatti, all'interno della Strategia nazionale *"to prevent and contrast discrimination on grounds of sexual orientation and gender identity"*, come uno strumento dalla duplice veste: da un lato, Internet è individuato come "mezzo" funzionale al proliferare di fenomeni discriminatori da prevenire ed arginare il più possibile; dall'altro, nell'ottica di quel necessario bilanciamento tra diritti che ben conosciamo, l'on line è riconosciuto come un territorio "sconfinato" sul quale poter diffondere buone prassi e politiche contro le discriminazioni.

Come accennato, la Dichiarazione di Malta parla, in generale, di garanzie di uguaglianza e di dignità umana a prescindere dall'orientamento sessuale e dall'identità di genere; si riferisce all'urgenza di evitare che persone LGBTI continuino ad essere vittime dell'odio, della violenza, della discriminazione e del bullismo, attraverso l'adozione di misure appropriate, legislative o di altro tipo, funzionali a combattere la discriminazione e garantire, poi, la libertà di espressione.

Nella Dichiarazione si parla anche di "rilancio delle azioni di sensibilizzazione del pubblico sulla diversità e sull'inclusione a tutti i livelli della società"; ma, come accennato in precedenza, si tratta di un documento di intenti che, in realtà, è da considerarsi come fonte di legittimazione per attuare, sul territorio nazionale dei singoli Paesi aderenti, la strategia nazionale a favore delle LGBT.

Per quanto riguarda l'Italia, la strategia nazionale LGBT per gli anni 2013 – 2015 è chiara, esplicita e prescrittiva nei confronti del Web e delle discriminazioni perpetrate attraverso Internet.

La Strategia è suddivisa in diverse aree strategiche e, nella specifica area dedicata a "COMUNICAZIONE E MEDIA", contiene un riferimento specifico alla Raccomandazione del Comitato dei Ministri del Consiglio d'Europa agli Stati membri sulle misure per combattere la discriminazione fondata sull'orientamento sessuale o l'identità di genere (CM / Rec (2010) 5), laddove si parla di "hate speech", o "Incitamento all'odio", invitando gli Stati membri ad adottare *"misure appropriate per combattere tutte le forme di espressione, comprese quelle diffuse attraverso i media e su Internet, che possano essere ragionevolmente intese come probabili di produrre l'effetto di incitare, diffondere o promuovere l'odio, o altre forme di discriminazione contro le persone lesbiche, gay, bisessuali e transgender."*

Il cosiddetto "hate speech" nella strategia nazionale, dovrebbe essere una forma di espressione da vietare e sconsigliare pubblicamente allorché si verifica, pur nel rispetto del diritto fondamentale alla libertà di espressione, riconosciuto ai sensi dell'articolo 10 della CEDU e della giurisprudenza della Corte europea dei diritti dell'uomo.

E ancora, nell'area strategica dedicata all'"OCCUPAZIONE", la strategia nazionale ha poi riconosciuto tra i propri obiettivi quello di prevenire e combattere la diffusione di stereotipi che possano in qualche modo incoraggiare, anche attraverso Internet, il cosiddetto "discorso di odio" nei confronti delle persone LGBT. Infine, la strategia riconosce l'importanza del monitoraggio del fenomeno e prescrive la necessità di raccogliere in modo sistematico i dati, in particolare quelli relativi ad Internet ed ai social network, sul linguaggio omofobico e transfobico a livello nazionale, europeo ed internazionale, diffuso attraverso i mass media, anche al fine di realizzare e promuovere un'adeguata attività di informazione sul fenomeno.

Giovani e web: l'importanza di tutelare e di "istruire"

L'importanza di intervenire a supporto di strategie anti-discriminatorie è essenziale, anche alla luce delle dimensioni che il fenomeno ha assunto, con riferimento a fenomeni di bullismo, e più in generale riguardo all'utilizzo consapevole di internet (e dei social network) da parte dei più giovani.

Internet ha un impatto sempre crescente quale veicolo di diffusione di informazioni e contenuti per i ragazzi e i siti più visitati risultano essere i social network ed i motori di ricerca (alcuni dati sono reperibili sul Libro bianco "Minori e media", pubblicato dall'AGCOM a gennaio 2014). Il grande ampliamento delle possibilità di circolazione dei contenuti comporta, peraltro, un'enorme facilità per giovani (spesso adolescenti e, dunque, minori) di accedere anche a contenuti nocivi per il loro sviluppo psico-fisico e morale. Per questo si avverte una sempre maggiore esigenza di tutela specifica di tali utenti ipersensibili dal momento che la rete internet, per le sue caratteristiche intrinseche potrebbe non garantire un adeguato livello di sicurezza per i minori.

Anche per questo – se si tiene conto che la legislazione è prevalentemente rivolta ai mezzi di comunicazione classici (tv e radio) – gli strumenti di "formazione negoziata del diritto" nel campo della rete (peraltro incoraggiati dalle fonti comunitarie e internazionali) risultano i più adatti a garantire un livello di protezione adeguato agli utenti, soprattutto se appartenenti a categorie meritevoli di tutela (quali sono i minori).

Non è un caso che lo stesso MIUR ha avviato negli anni una serie di iniziative, di contrasto al bullismo (con particolare attenzione a quello online), in molti casi iniziative intraprese o coordinate dal Ministero, ma portate avanti con il coinvolgimento di vari attori, e attive in diverse aree territoriali. A titolo esemplificativo, si possono segnalare:

- gli Osservatori regionali permanenti sul bullismo, istituiti con la direttiva ministeriale n. 16 del 5 febbraio 2007 e attivi presso gli Uffici scolastici regionali;
- il portale URP social, primo social tematico realizzato da una pubblica amministrazione;
- il progetto «Generazioni connesse», sull'utilizzo sicuro dei nuovi media da parte dei più giovani, cofinanziato dalla Commissione europea e coordinato dal MIUR, con la collaborazione di Save the Children Italia, di Telefono Azzurro, del Ministero dell'interno - Ufficio di Polizia postale e delle comunicazioni, dell'Autorità garante per l'infanzia e l'adolescenza, della Cooperativa E.D.I. e del Movimento difesa del cittadino;
- il progetto Tabby (Threat Assessment of Bullying Behavior: valutazione della minaccia di cyberbullismo nei giovani), approvato nel quadro del programma Daphne III nell'ambito della programmazione europea 2007-2013.

Si tratta di iniziative che, come è evidente, necessitano di competenze istituzionali integrate e che evidenziano l'esigenza di un coordinamento interministeriale che metta a fattor comune le energie pubbliche e del privato sociale. Il tutto con un obiettivo: quello di definire un piano d'azione al quale concorrano, insieme alle istituzioni, associazioni e organizzazioni di carattere nazionale e internazionale che si stanno confrontando con il problema, e aperto ai principali operatori del web.

Del resto, gli episodi di cyberbullismo che si verificano in misura sempre crescente su scala nazionale (stando alle segnalazioni delle autorità competenti) impongono la necessità di individuare “strumenti di primo soccorso”, a disposizione della cittadinanza, che agiscano con i medesimi processi e canali multimediali su cui fa leva il cyberbulling, al fine di riunire in un solo momento condiviso i servizi di sportello telematico messi a disposizione dai soggetti impegnati nel contrasto al fenomeno, con particolare riferimento ai minori.

Le iniziative avviate nel corso degli ultimi anni sono tantissime, e in molti casi si tratta di eventi o documenti che hanno come principali soggetti di riferimento, da un lato, i giovani (nativi digitali e, talvolta, poco accorti nell'utilizzo specie dei social network), dall'altro, gli adulti (genitori e/o insegnanti, per metterli al corrente, anzitutto, delle “insidie” del web e degli strumenti per contarstarle). Attraverso documenti snelli (e talvolta esemplificativi) si cerca, allora, di “far passare” alcuni messaggi essenziali:

- la possibilità di utilizzare l'**anonimato** come forma di protezione personale, ma non come strumento di offesa o di violazione dei diritti degli altri e delle persone più deboli (tenendo, peraltro, conto che le autorità competenti hanno molti strumenti a disposizione per scoprire l'“autore anonimo”);
- la **permanenza delle informazioni** su internet (e la difficoltà di rivendicare il diritto all'oblio). Come è stato evidenziato dagli studiosi (per tutti, di recente, A. Papa, RivistaAIC2014), la rete Internet ha aperto la strada ad una diffusione del pensiero che non si esaurisce in un momento – più o meno lungo, più o meno ripetuto – ma acquista carattere di permanenza. Nella maggior parte dei casi, infatti, ciò che viene inserito in Rete si colloca in una dimensione spaziotemporale indefinita, nella quale le informazioni, fatta eccezione per alcune ipotesi specificamente previste, rimangono fino ad una eventuale (ed in realtà poco frequente) rimozione.
- il rispetto della **privacy** (altrui, ma anche propria), e a questo proposito l'importanza di leggere attentamente le condizioni d'uso e le garanzie offerte nel contratto che si sottoscrive quando si decide di iscriversi a un social network.

Le policies dei social network con riferimento alla lotta alla discriminazione online

Di seguito vengono brevemente richiamate le policies di tre tra i principali social networks, con particolare attenzione agli strumenti e alle strategie rivolte a evitare o a contrastare la discriminazione online.

Facebook: una cultura basata sulla responsabilità

“La comunità di Facebook rappresenta una delle difese migliori contro comportamenti offensivi, spam e altre forme di contatti indesiderati. Le persone usano i propri nomi reali su Facebook, e in questo modo le azioni eseguite online vengono connesse con delle identità reali. Siamo fermamente convinti che le persone dovrebbero trattare gli altri esattamente come vorrebbero che gli altri trattassero loro, sia online che offline. Se vedi qualcosa di inappropriato o che ti fa sentire a disagio, faccelo sapere. Diamo molta importanza alle segnalazioni della nostra comunità e ci impegniamo al massimo per rispondere il più velocemente possibile.”

Questa breve introduzione è disponibile sul sito ufficiale e consultabile alla voce <<Centro per la sicurezza>>. E' stato inoltre creato un Comitato di consulenza per la sicurezza, col fine di fornire suggerimenti in materia di sicurezza online; vi prendono parte importanti organizzazioni internazionali come Childnet International, ConnectSafely, Family Safety Online Institute, The National Network to End Domestic Violence e WiredSafety.

La sicurezza è, dunque, il perno centrale intorno al quale ruotano le policies di Facebook, il quale invita in modo esplicito gli utenti a collaborare col team, attraverso strumenti appositi che vengono loro messi a disposizione, tra cui, in prima istanza, le segnalazioni. L'orientamento generale verso cui propendono i gestori della società è quello di rimuovere tutto ciò che viola le Condizioni di Facebook o gli Standard della comunità. Agli iscritti al social network viene, inoltre, ricordato che le segnalazioni sono confidenziali e che la piattaforma fornisce loro degli strumenti per proteggere la loro privacy, per rimuovere delle persone dal gruppo di amici o bloccarle, e per aiutarli a proteggere le informazioni personali su Facebook e altrove. Tale linea di condotta ben si concilia con quelli che sono i Principi su cui si basa la piattaforma (richiamati di seguito):

“Abbiamo ideato Facebook per rendere il mondo più aperto e trasparente, nella convinzione che ciò possa creare una maggiore comprensione e migliorare le comunicazioni. Facebook promuove l'apertura e la trasparenza offrendo agli utenti un'ampia libertà di condivisione e di contatto e si propone di raggiungere tali scopi sulla base di alcuni principi. Tali principi saranno limitati solo da limitazioni legali, tecnologiche e da nuove norme sociali”.

Tornando a quelli che sono gli Standard della comunità, diversi sono gli obiettivi che il team di Facebook si prefigge di perseguire. Tra questi, vi è ad esempio l'obiettivo di combattere il bullismo e le intimidazioni, limitare le autolesioni, ostacolare la pornografia, tutelare la proprietà intellettuale, proteggere i dati personali. Vi rientrano in particolare anche la lotta alla violenza e alle minacce e, soprattutto, la battaglia contro l'HATE SPEECH. Alla voce “contenuti che incitano all'odio” si legge, infatti: “Facebook non consente i contenuti che incitano all'odio, ma attua una distinzione tra contenuti seri e meno seri. Se da un lato incoraggiamo gli utenti a mettere in discussione idee, eventi e linee di condotta, non consentiamo la discriminazione di persone in base a razza, etnia, nazionalità, religione, sesso, orientamento sessuale, disabilità o malattia”.

Facebook invita gli utenti a segnalare eventuali contenuti offensivi, dal momento che, appunto, rientra nei principi della società eliminarli. Allo stesso tempo, tuttavia, i potenziali segnalatori vengono avvertiti del fatto che una loro segnalazione non garantisce una sicura rimozione del contenuto in oggetto. In particolar modo, il s.n. rispetta e promuove la diversità della sua comunità, per questo motivo è possibile che ciò che agli occhi di una persona potrebbe apparire inaccettabile o fastidioso non debba necessariamente essere rimosso o bloccato in base ai criteri della società in materia. In questa ottica, Facebook offre anche delle impostazioni personali che consentono agli utenti di scegliere cosa vedere: ad esempio, ognuno ha la possibilità di nascondere o tagliare i rapporti con persone, Pagine o applicazioni che ritiene offensive in modo semplice e discreto.

Le “ferree” regole di Twitter: <<sei quello che twitti>>.

All'interno della società è stato appositamente creato un team, Trust & Safety, con il compito di condurre indagini e rispondere alle segnalazioni di violazioni delle Regole di Twitter e dei Termini di Servizio.

“Il nostro obiettivo consiste nel fornire un servizio che ti permetta di scoprire e di ricevere contenuti provenienti da fonti che ti interessano e di condividere i tuoi contenuti con gli altri. Rispettiamo la proprietà dei contenuti che gli utenti condividono e ogni utente è responsabile per il contenuto pubblicato. In virtù di questi principi, non controlliamo attivamente e non censuriamo i contenuti creati dagli utenti, salvo in circostanze limitate descritte di seguito(…)”

In questo modo il team di Twitter intende dare una chiara indicazione di come la società si pone nei confronti degli iscritti e dei loro “tweet”, ma soprattutto di come intende risolvere eventuali controversie da questi generati, auto sollevandosi da ogni responsabilità, dal momento che non rientra nella loro politica il controllo preventivo e la censura dei post. Tuttavia, la società non ha mancato di redigere linee guida sufficientemente dettagliate in merito al corretto utilizzo della piattaforma, inserendo soprattutto quelle relative alla “limitazione dei contenuti”: impersonificazione, marchio, informazioni private, violenza e minacce, copyright, uso illecito e abuso. Alla voce “violenza e minacce”, in particolare, si legge: <<l'utente non può pubblicare o postare minacce dirette e specifiche di violenza contro gli altri>>.

Twitter si riserva il diritto (ma non l'obbligo!) di rimuovere o rifiutare, in ogni momento, la distribuzione di Contenuti sui Servizi, di sospendere o chiudere utenze e di richiedere la restituzione di alcuni nomi utente senza alcuna responsabilità nei confronti dell'utente, qualora quest'ultimo risulti palesemente in violazione delle regole relative a Contenuti sui Servizi. Twitter si riserva altresì il diritto di accedere, leggere, conservare e divulgare le informazioni che ritenga necessarie per:

- conformarsi a ogni legge, regolamento, procedimento legale o richiesta governativa applicabile;
- imporre l'osservanza delle Condizioni, anche mediante l'accertamento di potenziali violazioni delle stesse;
- individuare, impedire o in altro modo affrontare frodi o problematiche inerenti la sicurezza o di natura tecnica;
- rispondere alle richieste di assistenza da parte degli utenti o proteggere i diritti, la proprietà o la sicurezza di Twitter, dei propri utenti e del pubblico.

Conformemente alla sua politica, il team non media i contenuti né interviene nelle controversie tra gli utenti, ad eccezione di molestie o abusi mirati, inequivocabilmente in violazione delle regole e dei Termini di Servizio.

Peraltro, a partire dal 2012 (come richiamato da uno studio pubblicato su www.ilpost.it), Twitter ha inasprito le proprie regole di utilizzo del social network, introducendo per la prima volta un criterio geografico di censura selettiva. In seguito a una richiesta formale da parte di un'autorità verificata, oggi la società può decidere di oscurare, soltanto in una determinata nazione, i tweet o gli account che violino le leggi di quel paese (Country Withheld Content), mentre il messaggio continua a essere visibile per gli utenti di altre nazionalità. Nel frattempo Twitter ha deciso altresì di raccogliere le richieste di rimozione da parte degli utenti in un sito

ufficiale in cui fornisce statistiche semestrali sul numero complessivo di richieste pervenute.

Youtube: “patti chiari, amicizia lunga”.

“Non pretendiamo il tipo di rispetto riservato a suore, anziani e neurochirurghi. Desideriamo soltanto che non venga fatto cattivo uso del sito. Ogni nuova fantastica funzionalità della community su YouTube prevede un determinato livello di responsabilità. Confidiamo che tu sia responsabile: milioni di utenti rispettano la nostra fiducia, quindi cerca di essere uno di loro.”

Il team di Youtube (Google) stabilisce sin da subito un contatto diretto e informale con gli utenti, ai quali spiega in modo chiaro quali sono i termini di utilizzo del canale. Stessi strumenti a disposizione di tutti, le segnalazioni appunto, per fare notare agli “addetti ai lavori” eventuali video considerati non appropriati. Prima che gli stessi possano essere rimossi devono, tuttavia, essere verificati per stabilire se violino Termini e condizioni d'utilizzo. I video segnalati, quindi, non vengono automaticamente rimossi dal sistema.

YouTube è contro la pornografia o i contenuti sessualmente espliciti, ragion per cui non sono ammessi filmati appartenenti a una di queste categorie, anche se a pubblicare è il protagonista del filmato stesso. Il team lavora a stretto contatto con le forze dell'ordine e denuncia lo sfruttamento minorile. Non è permessa alcuna forma di violenza evidente o gratuita, né tantomeno video disgustosi di incidenti, persone morte e altri soggetti simili.

Youtube promuove e sostiene la libertà di espressione di tutti, ognuno è perciò libero di esporre la propria opinione e il proprio punto di vista, anche se non condivisibili. E' tuttavia vietata ogni forma di incitamento all'odio (linguaggio che attacchi o umili un gruppo in base a razza o origine etnica, religione, disabilità/invalidità, sesso, età, condizione sociale o orientamento sessuale/identità di genere). Non sono tollerati comportamenti sleali o che infastidiscono altri utenti, minacce, molestie, violazioni della privacy o la rivelazione di informazioni personali di altri membri. Pena il divieto assoluto e permanente di accesso al canale per chi non rispetti tali condizioni. Le violazioni dei Termini e condizioni d'uso potrebbero causare, quindi, l'invio di un avvertimento o la sospensione dell'account, avvenuta la quale non è più consentito crearne uno nuovo.

Sono assolutamente vietati e suscettibili di essere rimossi i <<contenuti che incitano alla violenza o all'odio nei confronti di individui o gruppi sulla base di determinati attributi, come ad esempio: razza o etnia, religione, disabilità, sesso, età, stato di veterano di guerra, orientamento/identità sessuale>> .

Il team ci tiene, inoltre, a precisare la sottile ma fondamentale differenza, a tutela della libertà di espressione, tra ciò che è considerato incitamento all'odio e ciò che non lo è. In linea generale è consentito criticare uno stato o una nazione; non lo è invece inviare commenti rancorosi o cattivi nei confronti di un gruppo di persone solo sulla base della loro etnia. In quel caso si è in presenza di una discriminazione grave, suscettibile di essere segnalata e quindi rimossa. Non è, tuttavia, così scontato che tutto ciò che potrebbe risultare offensivo o cattivo debba considerarsi incitamento all'odio. Se i contenuti pubblicati da una particolare persona creano in qualche modo fastidio, si può sempre prendere in considerazione la possibilità di bloccare l'utente, a meno che i contenuti in oggetto non violino le norme relative all' hatespeech. In quest'ultimo caso l'utente può inviare una segnalazione a YouTube affinché esamini i contenuti, o un intero account di un utente.

Le questioni più frequenti che Google si trova a dover correggere manualmente sono legate alla funzione di auto-completamento (i suggerimenti proposti da Google in base alle ricerche più

frequenti effettuate dagli utenti). Ogni giorno vengono regolarmente esclusi dai risultati di ricerca i contenuti che violano le norme. La correzione dei suggerimenti richiede un lavoro di aggiornamento continuo da parte degli operatori umani i quali devono prontamente rimuovere ogni termine appartenente al lessico dell'hate speech (come nigger) digitato dagli utenti. Può succedere infatti che mentre l'utente effettua la sua ricerca digitando una determinata parola, o il nome di una persona, contemporaneamente, nel box di ricerca, possano comparire autonomamente termini o aggettivi offensivi. Nel primo caso Google disattiva la funzione di auto-completamento, in altri casi occorrono delle segnalazioni per far rimuovere dei suggerimenti inopportuni. Come è accaduto nel caso delle persone affette da autismo, nei confronti delle quali l'utente, cominciando la sua ricerca con la stringa di testo autistic people should ("gli autistici dovrebbero"), l'auto-completamento suggeriva predicati come die ("morire") o be exterminated ("essere sterminati"). In questi casi Google non può disattivare l'auto-completamento per qualsiasi occorrenza della parola autistic, ma deve eliminare manualmente ciascun suggerimento scorretto o nocivo.