



CODICE DI CONDOTTA EUROPEO

Nel 2016, la Commissione Europea ha siglato con alcune piattaforme digitali (nello specifico: Facebook, Microsoft, Twitter e YouTube) un **“Codice di condotta per lottare contro le forme illegali di incitamento all’odio online”**.

Il Codice di condotta rappresenta uno strumento elaborato al fine di contrastare l’odio. Pertanto, il Codice prevede una serie di impegni a carico delle aziende informatiche, quali la predisposizione di adeguate procedure di verifica delle segnalazioni di contenuti illeciti, la regolare pubblicazione di dati relativi al numero di segnalazioni ricevute da ciascuna piattaforma e al numero di contenuti effettivamente rimossi dalle stesse piattaforme in quanto incitanti all’odio.



Tutte le piattaforme digitali e i Social Network prevedono dei divieti rispetto alla pubblicazione e condivisione di contenuti che rientrano nei casi di hate speech. Il mancato rispetto di tale divieto comporta delle sanzioni: la rimozione del contenuto, la sospensione o il blocco del profilo di chi abbia prodotto o condiviso il contenuto in questione, la cancellazione del canale ecc.

Il Codice di condotta è quindi uno strumento molto importante, perché obbliga le aziende dei social a seguire i regolamenti a cui hanno dato vita, monitorando i tempi e i modi entro cui ogni piattaforma garantisce il proprio sostegno agli utenti vittime di manifestazioni dell’odio. Conoscere il Codice di condotta ci permette di sapere quali sono i nostri diritti sui social e rivendicarli quando le piattaforme non svolgono il loro dovere.



Proponiamo adesso una rassegna sintetica delle linee guida predisposte nel merito dalle più diffuse “aziende informatiche”, in parte simili tra loro. Rimandiamo alle rispettive piattaforme per un approfondimento sulle rispettive linee guida.

Gli standard della community di **FACEBOOK** prevedono il divieto dei discorsi di incitazione all’odio, perché ritenuti capaci di creare un «ambiente di intimidazione ed esclusione» e perché, «in alcuni casi, possono promuovere violenza reale».

Il discorso di incitazione all’odio, secondo dunque i parametri di Facebook, è un attacco diretto alle persone sulla base di loro caratteristiche tutelate a norma di legge, ove per “attacco” si intende «un discorso violento o disumanizzante, dichiarazioni di inferiorità o incitazioni all’esclusione o alla segregazione». Il sito menziona espressamente le seguenti categorie protette: razza; etnia; nazionalità di origine; religione; orientamento sessuale; casta; sesso; genere; identità di genere; disabilità o malattie gravi. Sono inoltre fornite alcune misure di protezione per lo status di immigrato.

Facebook ammette che, in taluni casi, determinati contenuti pubblicati sul proprio profilo possano avere scopi perfettamente legittimi, come condividere contenuti altrui che incitano all’odio con lo scopo di sensibilizzare o informare le altre persone. In questi casi, quindi, parole o termini che potrebbero in altro modo violare i nostri standard vengono usati in modo autoreferenziale o per rafforzare una causa, oppure ancora a scopi umoristici o ironici, su taluni argomenti.

Pertanto, Facebook consente il contenuto, ma si aspetta che le persone indichino chiaramente la loro intenzione, aiutando la piattaforma a capire meglio la motivazione alla base. Quando l’intenzione non è chiara, il contenuto viene rimosso.



CODICE DI CONDOTTA EUROPEO

TWITTER vieta ai suoi utenti di «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». Il nome utente, il nome visualizzato, l'immagine e la bio del profilo non possono essere utilizzati per commettere abusi.

Queste scelte sono da Twitter giustificate nei seguenti termini: lo scopo di Twitter è dare a tutti la possibilità di creare e condividere idee e informazioni, di esprimere opinioni e convinzioni senza barriere. La libera espressione è un diritto dell'essere umano. Il ruolo della piattaforma è quello di agevolare la conversazione pubblica, dunque la rappresentazione di una vasta gamma di punti di vista.

Oltre alle categorie e i gruppi che si impegna a tutelare di fronte a contenuti violenti e discriminatori, Twitter fornisce, tra l'altro, una lista piuttosto esaustiva anche di tipologie di contenuti vietati. Alcuni esempi: minacce di violenza, augurare il peggio ad una persona o gruppo di persone, riferimenti a genocidi, diffondere stereotipi paurosi su un gruppo protetto, usare metafore sessiste e/o razziali ed immagini che incitano all'odio.

YOUTUBE vieta contenuti che incitano alla violenza o all'odio nei confronti di individui o gruppi sulla base delle seguenti caratteristiche protette: età; casta; disabilità; etnia; identità ed espressione sessuale; nazionalità; razza; condizione di immigrato; religione; sesso/genere; orientamento sessuale; condizione di vittima di un grave evento violento e di familiare di una vittima; condizione di veterano.

La piattaforma invia al titolare del canale un avviso dell'illiceità del contenuto postato, senza l'applicazione di alcuna sanzione laddove si tratti della prima segnalazione ricevuta rispetto ai contenuti di quel canale. In caso contrario, YouTube invia al titolare un avvertimento; al terzo avvertimento, il canale viene chiuso. Le linee guida indicano una lunga lista di tipologie di contenuti che violano le linee guida della piattaforma. Considerano, d'altro canto, l'ipotesi di contenuti che includano forme di incitamento all'odio laddove lo scopo sia di tipo «didattico, documentaristico, scientifico o artistico». Si pensi, per esempio, al caso di un documentario riguardante un gruppo che incita all'odio senza sostenerne però le idee; o ancora, a filmati di valenza storica (si pensi ad un video contenente un discorso pronunciato da Adolf Hitler). Il contesto "didattico" in senso lato deve, comunque, apparire nelle immagini o nell'audio del video stesso: non è sufficiente che esso sia semplicemente indicato nel titolo del contenuto.

Secondo le linee guida, **INSTAGRAM** condanna un utilizzo della piattaforma contrario alle norme di legge, ivi incluso il sostegno e appoggio a "gruppi inneggianti all'odio". Si propone, inoltre, di «creare una community positiva ed eterogenea»: a tal fine, Instagram rimuove «contenuti relativi a minacce verosimili o incitanti all'odio, contenuti atti a umiliare o a mettere in imbarazzo singoli individui, informazioni personali usate per ricattare o intimidire qualcuno e l'invio ripetuto di messaggi indesiderati».

In aggiunta, Instagram pone come divieto assoluto di pubblicare contenuti che contengano minacce alla sicurezza pubblica e personale, incluse le minacce fisiche, quelle relative a furto, atti di vandalismo e altri danni di natura economica.

TIKTOK vieta l'incitamento all'odio: gli account che in più occasioni violino questo divieto sono soggetti alle sanzioni della sospensione o del blocco del profilo. La piattaforma, in particolare, distingue tra "attacchi a gruppi protetti", "ingiurie" e contenuti che promuovano "ideologie basate sull'odio".

Gli attacchi a gruppi protetti sono quelle fattispecie di incitamento all'odio che attacchino, minaccino, disumanizzino o incitino alla violenza contro un individuo o un gruppo di individui sulla base di caratteristiche protette (razza; etnia; nazionalità; religione; casta; orientamento sessuale; sesso; genere; identità di genere; grave disabilità o malattia; condizione di immigrato). A tal fine, TikTok ritiene che costituiscano attacchi a gruppi protetti quei contenuti che incitino alla violenza o all'odio o disumanizzino determinati individui o gruppi di individui:

- a. Sostenendo che sono fisicamente o moralmente inferiori;
- b. Promuovendo o giustificando la violenza contro di loro;
- c. Descrivendoli come criminali;
- d. Riferendosi negativamente a loro come animali, oggetti inanimati o altre entità non umane; e. Promuovendo o giustificando l'esclusione, la segregazione o la discriminazione nei loro confronti.



CODICE DI CONDOTTA EUROPEO

CASO STUDIO

Monitoraggio annuale a cura della commissione europea per verificare applicazione "codice di condotta per lottare contro le forme illegali di incitamento all'odio" da parte delle piattaforme social.

Come abbiamo visto nella sezione teorica, le piattaforme IT si sono impegnate a seguire il codice di condotta per lottare contro le forme illegali di incitamento all'odio promosso dalla commissione europea.

Il loro impegno però come si traduce nella pratica della vita digitale degli utenti? È un impegno che rimane teorico o che si realizza in strategie efficaci di contrasto all'odio online?

Nel tentativo di trovare risposta a queste domande, annualmente, la commissione europea effettua un monitoraggio di verifica dell'effettiva applicazione del codice di condotta da parte delle piattaforme IT.



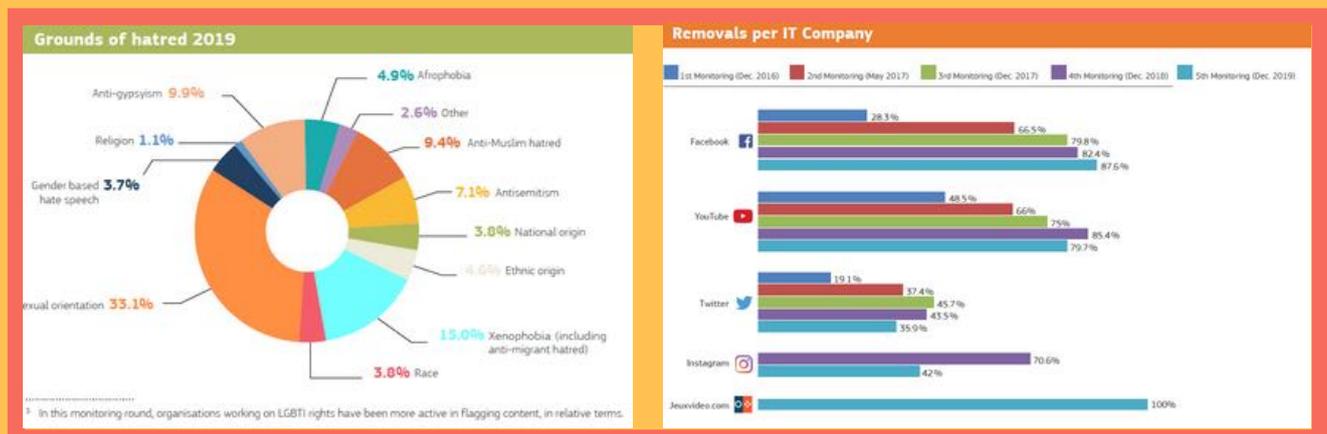
La Commissione europea organizza il monitoraggio coinvolgendo associazioni di tutta Europa che si impegnano a monitorare con i propri profili i diversi social network. Il monitoraggio consiste nell'individuare i commenti d'odio sui diversi social per poi segnalarli ai social seguendo la procedura da loro indicata nelle loro impostazioni di sicurezza. Dal momento in cui viene inviata la segnalazione, le associazioni coinvolte analizzano i tempi e reazioni che le piattaforme IT hanno alle segnalazioni. Vengono analizzati e monitorati i tempi entro cui le piattaforme rispondono alle segnalazioni, quali commenti effettivamente rimuovono e quali invece sembrano non violare le norme delle piattaforme nonostante la presenza di linguaggi d'odio.

Alla fine del periodo di monitoraggio, una volta raccolti tutti i dati da parte delle associazioni coinvolte, la Commissione europea redige un report con i risultati emersi, quali piattaforme rispettano il Codice di Condotta e quali invece devono migliorare le loro policy per essere a norma con gli standard europei. Questo esercizio di monitoraggio rappresenta un importante strumento di verifica volto a tutelare le persone nella loro esperienza online.

I report prodotti dal monitoraggio europeo sono disponibili al seguente link:

https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en

Parte dei risultati del monitoraggio 2020 condotto dalla commissione europea:





CODICE DI CONDOTTA EUROPEO

Provate a sfidare la classe nel simulare, per una settimana, un monitoraggio online rispetto al Codice di condotta dell'Unione Europea, e vedete quali risultati si ottengono.

Dopo aver analizzato nelle pagine teoriche le linee guida che il Codice di condotta propone alle piattaforme osservate se, nelle esperienze quotidiane e personali, queste sembrano rispettarlo ed applicarlo o meno.

ESERCITAZIONE

1. Lavorate individualmente da casa o in aula computer, tramite il vostro profilo su una piattaforma social a scelta. **Provate a monitorare diverse pagine pubbliche, gruppi, profili privati.**

2. Quando vi capita di trovare **un commento d'odio**, secondo le definizioni fornite nelle pagine teoriche, **segnalate** seguendo le procedure della piattaforma e **osservate** quanto tempo ci mettono effettivamente le piattaforme a **rispondere** alla vostra segnalazione e, soprattutto, che tipo di risposta vi arriva. Per esempio: il contenuto viene cancellato sì o no? Viene identificato come discorso d'odio sì o no? In caso negativo, che tipo di risposta vi viene fornita?

3. Per tenere traccia del monitoraggio, chiedete di creare una **tabella dove riportare i dati** come nell'esempio.

4. Dopo una settimana **riportate in classe** in plenaria il vostro lavoro individuale, e **commentate** insieme com'è andato il monitoraggio, cosa è successo, cosa è stato rilevato, se gli strumenti di segnalazione sembrano efficaci o meno.



Riportate: il tipo di commento segnalato, la piattaforma su cui è stato rilevato, data e ora della segnalazione e data e ora della risposta ottenuta, così come il contenuto della risposta.

Segnalazione	Commento 1	Commento 2	Commento 3
Rilevatore: Nome, Cognome			
Piattaforma			
Data e ora della segnalazione			
Data e ora della risposta			
Tipo di risposta			

CODICE DI CONDOTTA EUROPEO

The EU Code of conduct on countering illegal hate speech online

The robust response provided by the European Union

Nelle pagine precedenti abbiamo illustrato il funzionamento del codice di condotta per contrastare le forme illegali di incitamento all'odio redatto dall'Unione Europea per contrastare i discorsi d'odio online. **Ma com'è fatto questo documento e dove è possibile trovarlo?**

Nell'attività riportata di sotto vi proponiamo di scaricare il documento ufficiale del codice di condotta, leggerlo, analizzarlo e usarlo per dar vita ad una riflessione partecipata tra i componenti della classe o del gruppo di giovani coinvolti.



1. Scaricate il Codice di condotta europeo in lingua italiana al seguente link:
https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en#relatedlinks

2. Dividete la classe in gruppi, assegnando ad ognuno di essi una parte del Codice di condotta da analizzare. Riportare l'analisi in un cartellone riassuntivo da condividere con la classe.

Avviate una riflessione:

- In che modo il codice di condotta europeo per il contrasto ai discorsi d'odio online ci tutela? Quali sono gli aspetti che cerca di monitorare?
- Quali aspetti pensi non siano tutelati da questo documento e che sarebbe invece importante monitorare?
- Vi sentite tutelati quando siete in rete?
- Credete che i social media proteggano i nostri diritti o incontrate molte manifestazioni d'odio quando navigate in rete?
- In che modo reagite quando vi trovate di fronte ad un commento d'odio online?
- Conoscere qualche strumento da usare per denunciare e segnalare i discorsi d'odio?
- Vi è mai capitato di segnalare qualcuno? Se sì, come ha risposto la piattaforma che stavate usando?