

I Nostri Progetti di Sicurezza: Progress Report

Da Gennaio ci stiamo impegnando per rendere Twitter più sicuro. Abbiamo condotto una ricerca per capire meglio i problemi rilevati nella piattaforma, e poi abbiamo creato una serie di aggiornamenti per migliorarla. Successivamente abbiamo riunito il nostro Trust and Safety Council per analizzare il feedback dell'aggiornamento and capire come procedere.

Nonostante ci sia ancora molto lavoro da fare, i nostri utenti si aspettano meno abusi su Twitter oggi rispetto a 6 mesi fa. Di seguito vi presenteremo dei dati sul nostro progresso, quello che abbiamo imparato, e come intendiamo migliorare.

Maggiori Controlli su Account Abusivi

Abbiamo promesso di fare di più con la nostra tecnologia. Ora abbiamo aumentato il controllo sugli account abusivi di 10 volte paragonato a esattamente un anno fa. Stiamo anche limitando le funzioni o addirittura eliminando migliaia di account abusivi al giorno.

Molte trasgressioni si ripetono anche dopo la sospensione tramite la creazione di nuovi account. Il nostro nuovo sistema, solamente negli ultimi 4 mesi, ha rimosso il doppio di questo tipo di account. Al di là della tecnologia, il nostro team continua a monitorare i contenuti della piattaforma e a migliorare l'implementazione delle nostre policy.

Cambio di Comportamento

La comunicazione riguardo ai tweet problematici, specialmente a chi li scrive, è fondamentale. Gli account che mostrano comportamenti abusivi vengono posti in una funzione limitata e, inoltre, viene anche data una motivazione all'utente. Gli account che mettiamo in funzione limitata generano il 25% in meno di tweet abusivi e il 65 % di questi account rimangono in questa funzione solo una volta.

Controlli Personalizzati

Abbiamo introdotto un nuovo modo per personalizzare l'esperienza dei nostri user. Il nostro Quality Filter ci ha aiutato ad avere meno interazioni indesiderate: i blocchi posti dopo @menzioni da persone non seguite sono scese del 40%. Questi numeri varieranno, ma il nostro approccio sta avendo risultati positivi.

Abbiamo definizioni e criteri di molestie che valgono per tutti, nonostante persone possano pensare all'abuso in modo diverso. Dunque, usando queste nuove funzioni, ogni persona ha il controllo su quello che vede sulla sua pagina Twitter.

Idee per il futuro

Più miglioramenti, trasparenza e rapidità. La sicurezza online è un problema che può essere meglio risolto in collaborazione con il nostro Trust and Safety Council, esperti del settore, e intellettuali, sollecitando un feedback continuo. Vogliamo rendere Twitter un luogo sicuro dove i nostri users possono esprimersi senza problemi.

###