Selezione algoritmica dei contenuti in Svizzera - uno studio su Google e YouTube

Sintesi

Mykola Makhortykh*, Maryna Sydorova*, Aleksandra Urman**, Franziska Keller*, Silke Adam*

* Istituto per la comunicazione e gli studi sui media, Università di Berna
** Social Computing Group, Università di Zurigo

Marzo 2023

I sistemi algoritmici di selezione dei contenuti, come la ricerca di Google o le raccomandazioni di YouTube, danno forma all'ecosistema mediatico odierno. Filtrando automaticamente e dando priorità alle informazioni in risposta alle richieste degli utenti, questi sistemi aiutano le persone a navigare nella miriade di contenuti digitali disponibili. Tuttavia, gli algoritmi che alimentano questi sistemi possono portare a prestazioni improprie che rendono eventualmente più facile la diffusione di informazioni di fatto errate o fuorvianti, compromettendo la fiducia nelle istituzioni democratiche o convincendo le persone a prendere decisioni nocive per il proprio benessere personale o quello collettivo.

Per esaminare l'impatto dei sistemi algoritmici sull'esposizione alla disinformazione in Svizzera, il progetto ha preso in esame Google e YouTube, due piattaforme monopoliste nei mercati della ricerca e dell'hosting di video. Il progetto ha analizzato quali fonti hanno la priorità presso Google e YouTube in relazione al COVID e all'Olocausto, due argomenti spesso oggetto di disinformazione. Il progetto ha analizzato come le prestazioni degli algoritmi sono influenzate 1) dalla formulazione delle richieste di ricerca (ovvero se le query che esprimono interesse per la disinformazione comportano una maggiore esposizione ad essa), 2) dai profili comportamentali degli algoritmi (ossia se il fatto di consultare siti web con specifiche inclinazioni politiche prima di effettuare la ricerca comporti una maggiore esposizione alla disinformazione); e 3) dal tempo (ossia se l'esposizione alla disinformazione cambia nel tempo).

Il progetto ha utilizzato un audit algoritmico basato su agenti virtuali, una tecnica di ricerca che simula il comportamento umano per generare input per il sistema algoritmico e registrare i risultati. Sono stati realizzati tre cicli di raccolta dati nei mesi di maggio e giugno 2022. Nel caso della ricerca di Google e di YouTube, l'algoritmo ha raccolto i primi 10 risultati di ricerca in risposta a 28 query in tedesco: 14 sull'Olocausto e 14 sul COVID, dove metà delle query esprimevano interesse per la disinformazione (ad es. "corona-impfung mikrochip" – microchip vaccinazione coronavirus) e metà erano prive di un tale interesse (ad es. "corona-impfung" – vaccinazione coronavirus). Nel caso del sistema di raccomandazione di YouTube, nel progetto è stata esaminata una serie di raccomandazioni per i video con un interesse esplicito per la disinformazione sull'Olocausto e sul COVID e video senza un tale interesse. Per analizzare i risultati, il team qualificato di assistenti ha esaminato manualmente l'intero contenuto dei risultati della ricerca di Google, i primi tre minuti dei video della ricerca di YouTube e gli output del sistema di raccomandazione.

I risultati del progetto suggeriscono che, per rispondere alle query relative al COVID e all'Olocausto, sia Google che YouTube tendono a dare priorità ai contenuti provenienti da siti e canali giornalistici. Questi contenuti, inoltre, provengono solitamente da media esteri (in genere tedeschi), mentre i media svizzeri compaiono meno frequentemente tra i risultati in cima alla classifica (ad eccezione degli output raccomandati da YouTube). Allo stesso modo, le fonti educative a cui gli algoritmi danno la priorità provengono di solito dall'estero, diversamente delle fonti governative/amministrative che tendono ad essere svizzere.

Sono relativamente pochi i risultati di Google e YouTube che sostengono la disinformazione, in particolare in relazione all'Olocausto; tuttavia, per il COVID tali risultati appaiono più spesso (fino al 40 % dei risultati per le query "corona-diktatur" (dittatura del coronavirus) e "corona-lüge" (menzogna del coronavirus) per la ricerca in Google). Questa osservazione indica che è più difficile per gli algoritmi far fronte alla disinformazione in fase di sviluppo su certi argomenti. La disinformazione era più presente su Google Search rispetto a YouTube, nonostante YouTube abbia la reputazione di essere una piattaforma che facilita l'esposizione a informazioni false attraverso i suoi algoritmi. È tuttavia importante notare che la nostra analisi ha maggiori probabilità di identificare la disinformazione su Google piuttosto che su YouTube, poiché l'analisi dei contenuti di YouTube si concentra solo sui primi tre minuti di ogni video.

In termini di fattori specifici che influenzano l'esposizione alla disinformazione, il progetto ha riscontrato una scarsa influenza dei profili comportamentali algoritmici, indicando quindi che i siti web visitati prima della ricerca non sono necessariamente presi in considerazione dagli algoritmi in modo coerente. I fattori che contano di più sono la selezione delle query di ricerca (le query con un interesse esplicito per la disinformazione comportano una maggiore esposizione ad essa) e il momento in cui l'input è stato generato, come mostrato dalla variazione della visibilità dei contenuti legati alla disinformazione.

L'insieme di questi risultati dimostra che in determinate circostanze (in particolare quando gli utenti segnalano il loro interesse), i sistemi algoritmici possono facilitare l'esposizione dei cittadini svizzeri alla disinformazione. Sottolinea inoltre quanto sia importante aumentare la consapevolezza della disinformazione algoritmica nella popolazione svizzera, creare un'infrastruttura di monitoraggio per tracciare il modo in cui gli algoritmi affrontano la disinformazione nel tempo e promuovere una discussione per definire sul piano normativo il ruolo degli algoritmi nel contesto svizzero.