

## **"Libertà vigilata. La lotta per il controllo di Internet" di David Kaye**

*di Silvia Cegalin*

Recensione a: David Kaye, *Libertà vigilata. La lotta per il controllo di Internet*, Prefazione di Enrico Pedemonte, Treccani Libri, Roma 2021, pp. 192, 18 euro (scheda libro)

In un periodo storico in cui Internet è divenuto sempre più determinante all'interno della definizione dei fatti politici, si pensi ad esempio alla vittoria di Donald Trump alle elezioni presidenziali del 2016 e al ruolo giocato dai social media, è necessario interrogarsi su quali siano le strategie di controllo in rete e su come cercare di arginare contenuti altamente manipolatori tesi a trasformare le dinamiche del mondo reale.

All'interno di questo contesto tematico si inserisce il primo libro tradotto in italiano di David Kaye, *Libertà vigilata* edito da Treccani Libri, che si posiziona tra le analisi più attuali e complete per affrontare il tema del controllo di Internet. Se infatti la questione della sorveglianza non appare nuova nel dibattito pubblico, e a testimoniare è il successo sempre maggiore dei surveillance studies che specialmente nei lavori di David Lyon e di Shoshana Zuboff trovano le disquisizioni più importanti; la materia della sorveglianza in rete attualmente manca ancora di una struttura coerente, e proprio per questo il testo di Kaye rappresenta un'apripista essenziale per incominciare a discutere in maniera profonda il rapporto che lega censura e libertà di espressione. In *Libertà vigilata* David Kaye - giurista, docente di Legge alla University of California e Relatore speciale delle Nazioni Unite sulla promozione e la protezione del diritto alla libertà di opinione e di espressione - mette in discussione i metodi di controllo adottati dalle piattaforme che, come vedremo, definisce inadeguati e incapaci di tutelare o aumentare il grado di libertà di espressione presente nel web.

Kaye nel suo testo asserisce inoltre che una censura avente parametri di giudizio non ben regolamentati comporta ripercussioni anche verso utenti o comunità che vogliono usare Internet a scopo informativo e/o ricreativo, perché le loro pubblicazioni rischiano di essere oscurate per un semplice errore di valutazione. Non è un caso infatti che l'inizio del libro sia dedicato proprio alle piattaforme e ai loro metodi censori. Kaye mostra che se da un lato i dirigenti delle maggiori piattaforme (Facebook, YouTube, Twitter) hanno nel corso degli anni attivato misure di controllo per prevenire la pubblicazione di materiale pericoloso, dall'altro lato è anche vero che queste misure non hanno evitato che proprio tali piattaforme fossero usate dagli estremisti per divulgare le loro ideologie.

Al riguardo Kaye espone il caso dei rohingya del Myanmar - uno tra i popoli più discriminati al mondo - e di come Facebook ha permesso che i militanti buddisti antislamici e gli ufficiali del Tatmadaw (l'Esercito della Birmania) potessero pubblicare incitamenti all'odio contro i rohingya, elemento che ha avuto un ruolo centrale nell'incentivare i pregiudizi di cui i rohingya erano vittime, aumentando il sentimento di intolleranza nei loro riguardi. «Facebook ha messo a disposizione una piattaforma importante per la diffusione di questo incitamento all'odio, benché le norme di Facebook che lo vietano, avrebbero potuto essere applicate, in teoria, per scopi positivi. [...] Facebook non è stato capace di moderare quei contenuti. Ha invece rimosso, inspiegabilmente, i post che descrivevano aggressioni reali ai danni dei rohingya e ha sospeso un gruppo di utenti creato per segnalare le località sotto attacco e permettere ai rohingya di evitarle» (p. 47).

Il problema che evidenzia Kaye appare di fondamentale importanza perché fa emergere la chiara contraddizione di metodo presente all'interno delle piattaforme: pur essendoci delle regole, queste non si rivelano essere sempre in favore della tutela dei diritti fondamentali dell'individuo, talvolta

infatti concedono visibilità a chi questi diritti li viola. Le piattaforme, scrive Kaye, in molti casi non riescono a combattere in maniera decisiva le diverse forme di violenza espresse in rete, perché le regole vigenti sono elementari o utili soltanto per intervenire in maniera emergenziale, quindi quando contenuti potenzialmente dannosi sono già stati esposti; e se a ciò si aggiunge che spesso i dirigenti delle piattaforme non riescono a far rispettare queste regole, il panorama della sorveglianza in Internet si rivela più complicato e critico di quello che superficialmente può apparire.

A dispetto di ciò, tuttavia, è importante precisare che il ruolo delle piattaforme non risulta essere tra i più semplici, perché se da un lato esse sono obbligate a tutelare la sicurezza, dall'altro si trovano a dover mantenere l'immagine del web come luogo flessibile e privo di confini in cui vigono la libertà di espressione e lo scambio aperto di idee. Kaye, tuttavia, interviene anche su questo ultimo punto sfatando il mito di Internet come spazio sregolato e anarchico in cui ognuno può fare e divulgare quello che vuole. Le forti trasformazioni [1] che specialmente negli ultimi tempi hanno interessato Internet (come ad esempio la nascita dei social network, la formazione estemporanea di notizie e la viralità) ha modificato in profondità la natura originaria del web, costringendolo ad adottare misure sempre più restrittive che però, come sottolinea Kaye, non funzionano o funzionano male.

Tra le metodologie per contrastare la diffusione di materiale inappropriato le principali piattaforme si sono affidate ai moderatori e agli algoritmi, argomento a cui Kaye dedica il terzo capitolo Umani e Macchine, in quanto all'interno del binomio controllo/libertà entrambi giocano una parte essenziale. I moderatori e/o gli algoritmi infatti hanno il compito di rintracciare, segnalare ed eliminare tutti quei contenuti che sono ritenuti di natura pericolosa. Anche in questo caso Kaye individua delle contraddizioni.

Grazie alla testimonianza di Burcu Gültekin Punsmann, ex moderatrice assunta dall'Arvato - divisione della multinazionale Bertelsmann alla quale Facebook appalta la moderazione dei contenuti per la Germania - emerge uno spaccato dell'azione delle piattaforme piuttosto compromesso. La Punsmann in un'intervista rilasciata nel 2018 a Kaye dichiara che l'Arvato, oltre esercitare una considerevole pressione psicologica sui propri dipendenti (ad esempio se non usava il mouse per 5 minuti lo schermo le inviava una notifica), chiedeva ai moderatori di evadere all'incirca 1.300 segnalazioni al giorno, di non pensare troppo e di agire esclusivamente in base alle linee guida stabilite dalla compagnia senza soffermarsi sul contesto o sulle varianti. Una linea di azione che ha prodotto, e non poche volte, la censura di contenuti di natura informativa, trasformando il web in un luogo sempre meno ospitale per la condivisione aperta di idee.

Dall'altro la spinta accelerata dell'intelligenza artificiale ha fatto sì che anche le macchine ricoprissero questo ruolo, con risultati che ad oggi si possono definire ancora molto incerti. Il dubbio che espone Kaye riguarda non tanto la capacità dell'AI di identificare materiale pericoloso, quanto se affidare la libertà di espressione ad un sistema algoritmico sia una scelta saggia, perché facendo riferimento al pensiero di Safiya Noble [2], scrive: «I programmatori hanno dei pregiudizi, come li hanno quanti creano le regole per loro, pregiudizi che a volte sono integrati nelle prassi, e che riguardano il genere, la razza, la politica e molti altri temi carichi di conseguenze. Esasperando questi problemi sostanziali, le operazioni dell'AI sono nebulose per la maggior parte degli utenti, e rappresentano una grossa sfida alla trasparenza nella regolamentazione e moderazione di quanto viene detto» (p. 86). Lasciare all'AI la completa autonomia di scelta, di conseguenza, non significa aumentare il livello di oggettività nella decisione di cosa vada censurato e cosa no. Per garantire la reale certezza che il web non diventi spazio per il radicamento di odio sarebbe perciò opportuno valutare i casi singolarmente, soluzione che per la sua complessità e dispendiosità le compagnie non sembrano voler prendere in considerazione.

Se le piattaforme nella lotta ai contenuti illeciti sembrano assumere l'intelligenza artificiale come metodologia preferita, nella seconda parte del libro e in particolare nei capitoli intitolati rispettivamente *Wir Schaffen Das* (possiamo farcela) e *Scegliete da che parte stare*, Kaye sembra

chiedersi: «Nella gestione del controllo qual è la posizione degli organi governativi occidentali e quali sono i provvedimenti governativi atti a prevenire la violenza in rete?». Purtroppo anche su questo aspetto Kaye non è portatore di buone notizie.

Il quadro generale che emerge in *Libertà vigilata* è che i governi, in particolar modo quelli europei, non stanno facendo molto per contribuire alla vigilanza in rete, lasciando così l'intero controllo alle imprese private, eclatante a proposito è il caso della Germania [3]. «In tutto il mondo i governi, sia democratici che autoritari, stanno chiedendo che i social media e le compagnie dei motori di ricerca risolvano il problema dei contenuti di stampo terroristico in rete. Governate lo spazio che gestite, altrimenti concluderemo che siete dalla parte dei terroristi» (p. 105). Soluzione discutibile perché, sottolinea Kaye, delegare parte della sicurezza pubblica e della lotta al terrorismo ai privati significa che la tutela e la protezione dei cittadini viene amministrata da soggetti che agiscono secondo logiche dettate dal profitto, e quindi molte decisioni non saranno prese seguendo principi "umanistici", di interesse generale o facendo riferimento ai diritti umani, ma seguendo interessi prevalentemente economici. Inoltre, continua l'autore, il fatto che ogni piattaforma stabilisca autonomamente le proprie norme e condizioni di servizio, differenziate anche in base alle località geografiche in cui agiscono, comporta un diverso trattamento dei contenuti di stampo terroristico, e ciò significa avere aree più esposte al rischio rispetto ad altre.

Ma la riflessione di Kaye non si ferma qui e si spinge, nell'ultima parte di *Libertà vigilata*, ad inquadrare un altro problema tipico degli spazi condivisi nel web: la disinformazione. La mancata presa di posizione da parte dei governi riguarda infatti anche la tematica della disinformazione, principalmente conosciuta con il termine inglese *fake news*. È emerso che il maggior potere conferito alle piattaforme si traduce anche in una gestione dei contenuti che, dove non classificati come pericolosi o inneggianti all'odio, lasciano ai dirigenti delle compagnie la piena libertà d'azione. Caso, questo, che riguarda le notizie false create per indirizzare l'opinione pubblica in una determinata direzione specialmente in periodi di crisi o durante le elezioni (si pensi all'esempio fatto all'inizio sulla vittoria di Donald Trump o al caso della Brexit).

Kaye, supportando la sua analisi con le teorie di Samantha Bradshaw [4] - esperta di new media e propaganda computazionale - dichiara che se per i contenuti considerati potenzialmente terroristici le compagnie adottavano delle misure di controllo, percependosi responsabili, nel caso della disinformazione, al contrario, essi non si sentono coinvolti nella questione, anche a fronte del fatto che le leggi sui diritti umani non prevedono restrizioni per quanto concerne la falsità delle informazioni. A conferma di questo atteggiamento è il fatto che la lotta contro la disinformazione è per i maggiori social network quasi unicamente relegata alla segnalazione da parte della community, sistema di dubbia efficacia deresponsabilizzante per le piattaforme che da sempre minimizzano tale problematica e che, invece, risulta essere di primaria importanza per la preservazione della democrazia.

*Libertà vigilata* invita perciò a ridiscutere i meccanismi che regolano oggi il web, perché una scarsa regolamentazione può comportare scelte e decisioni che possono influenzare irreversibilmente la realtà, mettendo in seria discussione i diritti e la democrazia. Il libro di Kaye funziona come un ottimo punto di partenza per poter valutare un approccio inerente la sicurezza in Internet in cui piattaforme e governi cooperino per raggiungere soluzioni più soddisfacenti, e dove la libertà di espressione non giochi un ruolo secondario.

È bene tuttavia precisare che Kaye nel suo testo non mostra soluzioni, questo libro non può infatti da solo svolgere una funzione risolutiva per un tema che da più parti interessa il dibattito pubblico, di certo però fornisce strumenti e argomentazioni utili per rendere il lettore consapevole su come si sia dispiegato all'interno del web il rapporto tra le nozioni di controllo e libertà, aiutandolo ad interpretare alcuni parametri che spesso sono taciuti. In attesa di una migliore stagione di Internet auspichiamo che queste riflessioni di Kaye siano da stimolo per considerare possibili ristrutturazioni

circa il controllo e l'informazione nella rete.

[1] Per comprendere al meglio le trasformazioni che hanno interessato Internet si rimanda alla lettura del saggio *The failure of Internet Freedom* di Jack Goldsmith, contenuto in *The Perilous Public Square: Structural Threats to Free Expression Today* di David E. Pozen, Columbia University Press, New York 2020.

[2] Per una maggiore disamina delle teorie di Safiya Noble si invita alla lettura di *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York University Press, New York 2018.

[3] Nel 2017 il Ministro tedesco Heiko Maas, all'epoca Ministro federale della giustizia e della tutela dei diritti del consumatore, per combattere il terrorismo e gli attacchi di odio nei social ha proposto la legge per la tutela dei diritti sui social network. La legge *Netzwerkdurchsetzungsgesetz (NetzDG)* infatti obbligava i social network con più di due milioni di utenti registrati in Germania a rimuovere a livello locale i contenuti illegali entro 24 ore dalla ricezione della notifica, e se il social non provvedeva all'oscuramento del contenuto si procedeva con una sanzione economica tutt'altro che irrisoria. La responsabilità della presenza di materiale illegale ricadeva perciò totalmente sulle compagnie private.

[4] Per un approfondimento si consiglia: Samantha Bradshaw e Philip N. Howard, *Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation*, Oxford Internet Institute, luglio 2018.