# S&V FOCUS | La disinformazione corre sull'AI. Spunti di riflessione




In precedenti occasioni editoriali, più volte abbiamo avuto modo di sottolineare l'efficacia e utilità di alcune applicazioni dell'IA in diversi campi, soprattutto in area medica. Ma questi potenti e innovativi strumenti, purtroppo, potrebbero essere impiegati anche per finalità e interessi del tutto estranei al bene comune. Ad esempio – in base al campanello d'allarme di molti esperti di sicurezza informatica – per condizionare le elezioni democratiche in varie parti del mondo. Prassi riprovevoli e illegali, quali la sistematica disinformazione tramite i social media, ad opera di un paese contro un altro, avevano già qualche anno fa fatto capolino sulla scena politica internazionale in tempo di elezioni, per provare a condizionare le opinioni degli elettori. Ad esempio nel 2016, quando i russi hanno lanciato sui social media una

serie di campagne di disinformazione mirate alle elezioni presidenziali statunitensi. Nei sette anni successivi, poi, altri paesi – primi fra tutti Cina e Iran – hanno usato i social media per influenzare le elezioni straniere, sia negli Stati Uniti che in altre parti del mondo.

Questo "modus operandi" – che con ogni probabilità non muterà nei prossimi anni – oggi può però avvalersi di una nuova arma: il potente ausilio dell'IA generativa e dei modelli linguistici di grandi dimensioni. Come è ormai noto, infatti, questi hanno la capacità di produrre in modo rapido e semplice infinite risme di testo su qualsiasi argomento, con qualsiasi tono e da qualsiasi prospettiva. Basti pensare al sistema ChatGPT, introdotto per la prima volta nel novembre 2022, che da pochi mesi (marzo 2023) viene rilasciato nella più potente versione GPT-4. Ma non è l'unico, anzi è in buona compagnia di diverse altre IA per la produzione di linguaggio e immagini.

**Ma come queste tecnologie cambieranno effettivamente la disinformazione? Quanto saranno efficaci? Quali effetti avranno? Non lo sappiamo ancora, ma… lo sapremo presto! La stagione elettorale, infatti, sarà presto in pieno svolgimento in ampia parte del mondo democratico. Basti pensare che, entro la fine del 2024, il 71% delle persone che vivono nelle democrazie voterà in un'elezione nazionale.** Saranno coinvolti ad esempio Argentina e Polonia (ottobre), Taiwan (gennaio), Indonesia (febbraio), India (aprile), Unione Europea e Messico (giugno), Stati Uniti (novembre), oltre a nove democrazie africane (tra cui il Sudafrica). Con molta probabilità, nel 2024 si aggiungeranno anche Australia e Regno Unito. Va da sé che il risultato elettorale di molti di questi Paesi risulta di grande interesse per altri Paesi, che in passato hanno già condotto sui social media operazioni per influenzarlo. Secondo il parere di Bruce Schneier, docente a contratto di politiche pubbliche alla Harvard Kennedy School della Harvard University, ad esempio, alla Cina interessano molto Taiwan, Indonesia, India e molti paesi africani; alla Russia

interessano il Regno Unito, la Polonia, la Germania e l'UE in generale. E poi, tutti hanno a cuore gli Stati Uniti.

**Va anche sottolineato che i costi di un'operazione di disinformazione contro un Paese tende a diminuire se più Paesi vi partecipano. E strumenti come ChatGPT riducono in modo significativo il prezzo della produzione e della distribuzione di propaganda, rendendo questa capacità alla portata di molti più Paesi. Dal punto di vista operativo, la gestione di una campagna di disinformazione non si limita certo alla generazione di contenuti. La vera sfida, infatti, è la distribuzione.**

Per questo, un propagandista necessita di una serie di account falsi su cui pubblicare i contenuti e di altri che li diffondano nel mainstream, dove possono diventare virali. Per contro, aziende come Meta sono diventate molto più brave a identificare questi account e a eliminarli. E' anche vero che, negli anni, mutano molto anche i social media e la loro influenza. Molti organi di propaganda, ad esempio, si sono spostati da Facebook a piattaforme di messaggistica come Telegram e WhatsApp, il che li rende più difficili da identificare e rimuovere. TikTok è una piattaforma più recente, controllata dalla Cina e più adatta a video brevi e provocatori, che l'IA rende molto più facili da produrre. Inoltre, le attuali IA generative sono collegate a strumenti che renderanno più semplice la distribuzione dei contenuti.

Gli strumenti di IA generativa consentono anche nuove tecniche di produzione e distribuzione, come la propaganda di basso livello su larga scala. Immaginiamo un nuovo account personale alimentato dall'IA sui social media. Per la maggior parte, si comporta normalmente. Pubblica la sua finta vita quotidiana, si unisce a gruppi di interesse e commenta i post degli altri, e in generale si comporta come un utente normale. E di tanto in tanto, magari non troppo spesso, dice – o amplifica – qualcosa di politico. Queste "persone" (virtuali) hanno un'influenza trascurabile da soli. Ma se replicati a migliaia

o a milioni, avrebbero un’influenza di gran lunga maggiore. Gli esperti di cyber-sicurezza, poi, sanno molto bene come Paesi come la Russia e la Cina sono soliti testare gli attacchi informatici e le operazioni informative su paesi più piccoli prima di diffonderli su larga scala. Quando ciò accade, è importante essere in grado di rilevare le impronte digitali di queste tattiche.

**Per contrastare le nuove campagne di disinformazione occorre saperle riconoscere, e per riconoscerle occorre cercarle e catalogarle subito. Ma questo futuro scenario “disastroso” non è ineluttabile. Forse le cose andranno meglio di così. Del resto, nell’era dell’IA generativa si sono già svolte alcune importanti elezioni democratiche (primarie in Argentina, elezioni al primo turno in Ecuador ed elezioni nazionali in Thailandia, Turchia, Spagna e Grecia) senza particolari problemi di disinformazione. Ma vigilare con cura serve anche a questo!**

Fonte www.AgenSir.it | 16 ottobre 2023