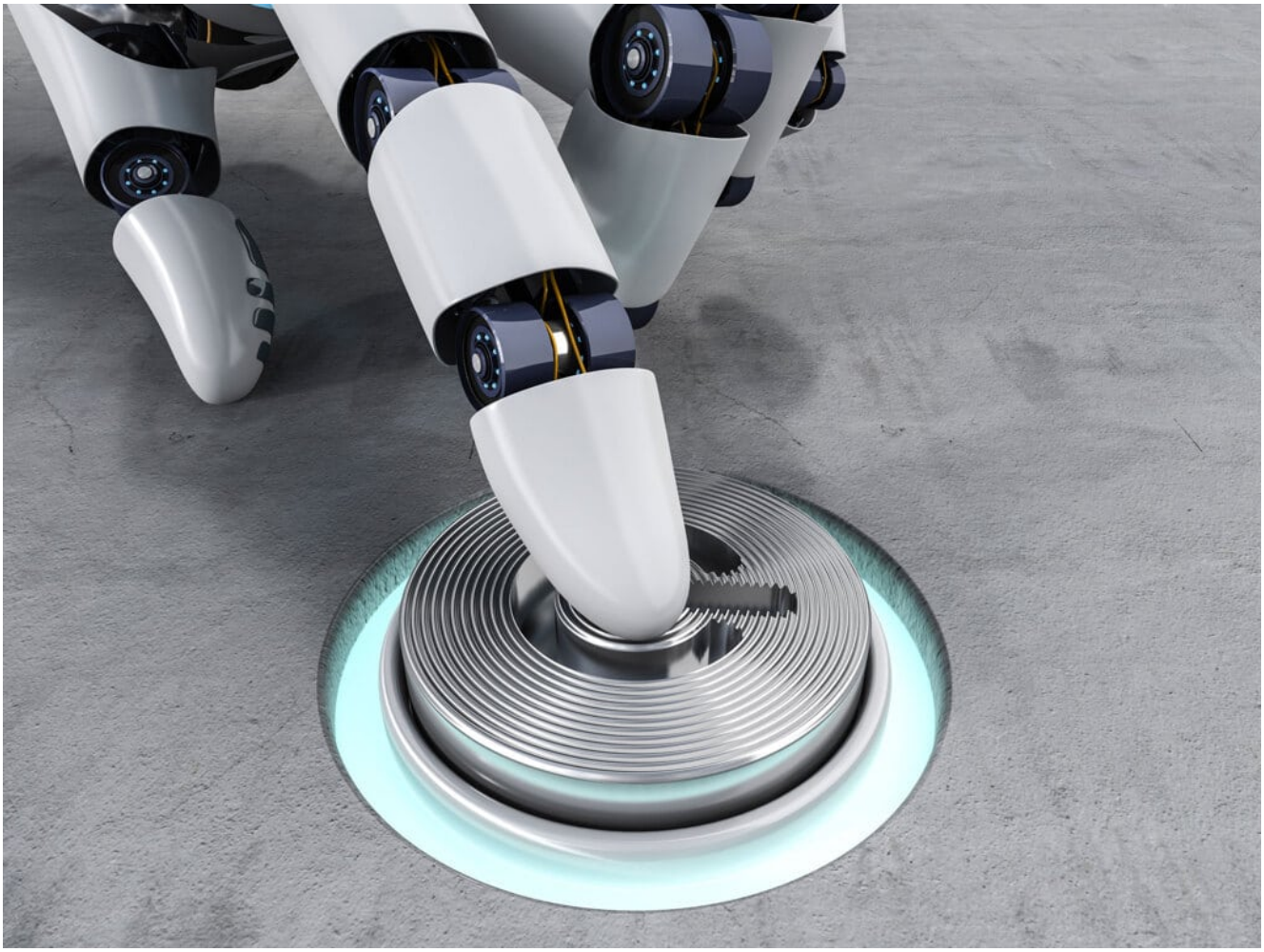
Il futuro della politica è in mano all’intelligenza artificiale



La stagione della campagna presidenziale statunitense è ufficialmente, ma davvero ufficialmente, arrivata, il che significa che è il momento di affrontare gli strani e insidiosi modi in cui la tecnologia sta distorcendo la politica. Una delle principali minacce che si profilano all’orizzonte è l’arrivo di personalità artificiali, destinate a dominare il dibattito politico. Il rischio nasce da due tendenze che si presentano contemporaneamente: la generazione di testi alimentata dall’intelligenza artificiale e le chatbot (i software che simulano una conversazione con un essere umano) sui social network. Queste “persone” generate da computer sommergeranno le discussioni realmente umane su

internet.

I software di generazione di testi sono già abbastanza avanzati da trarre in inganno la maggioranza delle persone, la maggior parte delle volte. Stanno già scrivendo notizie, soprattutto di sport e di finanza. Parlano con i clienti nei siti che vendono prodotti. Scrivono convincenti editoriali su alcuni argomenti d'attualità (anche se esistono dei limiti al riguardo). E sono usati per rafforzare il "giornalismo pink-slime", quei siti concepiti per sembrare fornitori di notizie locali ma che in realtà pubblicano propaganda.

Anche i contenuti generati da algoritmi che si presentano come se fossero scritti da esseri umani hanno raggiunto livelli record. Nel 2017, per un certo periodo, la Commissione federale per le comunicazioni degli Stati Uniti ha aperto ai commenti online il suo piano di porre fine alla neutralità della rete, ricevendo l'impressionante cifra di 22 milioni di commenti. Molti di questi, forse la metà, erano falsi, e si servivano d'identità false. Questi commenti erano anche poco elaborati: 1,3 milioni erano generati a partire dallo stesso modello, semplicemente con alcune parole modificate perché apparissero diversi gli uni dagli altri. Non reggevano neanche di fronte a un'analisi sbrigativa.

Disinformazione e democrazia

Simili azioni saranno sempre più sofisticate. Nel corso di un recente esperimento Max Weiss, un ricercatore di Harvard, ha usato un programma di generazione testi per creare mille commenti in risposta a un appello del governo federale relativo al programma sanitario Medicaid. Ciascuno di tali commenti era diverso dall'altro, e sembrava frutto di persone reali che difendevano una posizione politica specifica. Hanno ingannato gli amministratori del sito Medicaid.gov, che li hanno ritenuti reali preoccupazioni di esseri umani in carne e ossa. Trattandosi di ricerca accademica, Weiss ha successivamente identificato i commenti e ha chiesto che

fossero rimossi, affinché non vi fosse alcuna interferenza irregolare con l'effettivo dibattito sull'argomento. Il prossimo gruppo che tenterà una cosa del genere non sarà altrettanto onesto.

Sono anni che le chatbot distorcono le discussioni sui social network. Circa un quinto dei tweet relativi alle elezioni presidenziali del 2016 è stato pubblicato da bot, secondo una stima. Lo stesso vale per circa un terzo di quelli relativi al voto sulla Brexit dello stesso anno. Un rapporto dell'Oxford internet institute dello scorso anno ha trovato prove dell'utilizzo di bot per diffondere propaganda in cinquanta paesi. Questi tendevano a essere programmi semplici che ripetevano automaticamente slogan, come i 250mila tweet filosauditi "abbiamo tutti fiducia in Mohammed bin Salman" apparsi dopo l'omicidio di Jamal Khashoggi nel 2018.

> *Il nostro futuro sarà fatto di chiassose discussioni politiche, perlopiù tra bot e altri bot*

Individuare molti bot con pochi follower è più difficile che rilevare alcuni bot con molti follower. E misurare l'efficacia di questi bot non è semplice. Le migliori analisi indicano che questi non hanno influenzato le elezioni presidenziali statunitensi del 2016. Più probabilmente distorcono la percezione che le persone hanno dell'opinione pubblica e la loro fiducia nella discussione politica ragionata. Siamo tutti immersi in un nuovo esperimento sociale.

Nel corso degli anni, i bot algoritmici si sono evoluti fino ad avere una propria personalità. Possiedono nomi falsi, false biografie e false foto, talvolta generate dall'intelligenza artificiale. Invece di diffondere propaganda senza sosta, postano solo occasionalmente. I ricercatori possono rilevare che si tratta di bot e non persone in base alla frequenza e all'andamento dei loro post. Ma la tecnologia dei bot continua a migliorare, rendendo difficili i tentativi di rilevamento. I

gruppi futuri non saranno così facili da identificare. Riusciranno a integrarsi meglio nei gruppi sociali di persone in carne e ossa. La loro propaganda sarà più sottile, e si mescolerà all'interno delle discussioni che interessano tali gruppi.

Mettete insieme queste due tendenze e avrete la ricetta per fare in modo che le discussioni non umane prendano il sopravvento sulle discussioni politiche tra esseri umani.

Chi controlla i bot

Presto le personalità alimentate da intelligenza artificiale saranno in grado di scrivere lettere personalizzate a giornali e parlamentari, esprimere il proprio commento nel quadro di processi legislativi pubblici, creando personalità che perdurano e che appaiono reali anche a quanti cercano di smascherarle. Saranno anche in grado di presentarsi come individui sui social network e d'inviare testi personalizzati. Avranno milioni di repliche e discuteranno di tali questioni giorno e notte, inviando miliardi di messaggi, lunghi e brevi. La somma di tutte queste cose gli permetterà di sommergere ogni reale discussione su internet. Non solo sui social network, ma dovunque ci sarà un dibattito.

Magari questi bot dotati di personalità saranno controllati da attori stranieri. Magari da gruppi politici nazionali. Magari dai candidati stessi. Più probabilmente, chiunque potrà farlo. La più importante lezione a proposito della disinformazione nel 2016 non è che ci sia stata disinformazione, bensì quanto sia stato facile e poco costoso disinformare le persone. I futuri miglioramenti della tecnologia renderanno la cosa ancora più economica.

Il nostro futuro sarà fatto di chiassose discussioni politiche, perlopiù tra bot e altri bot. Non è quello che si ha in mente quando si loda il mercato delle idee, o qualsiasi altro processo politico democratico. La democrazia ha bisogno

di due cose per funzionare efficacemente: informazione e rappresentanza. Le personalità artificiali possono privare le persone di entrambe le cose.

Sistemi di difesa

È difficile immaginare delle soluzioni. Possiamo regolamentare l'uso dei bot – una proposta di legge in California imporrebbe ai bot d'identificarsi – ma la cosa sarebbe efficace solo con le campagne d'influenza legittime, come la pubblicità. Sarà molto più difficile rilevare le operazioni d'influenza surrettizia. La difesa più ovvia è sviluppare e standardizzare metodi migliori di autenticazione. Se i social network verificano che dietro ogni account c'è effettivamente una persona reale, allora saranno in grado di eliminare più facilmente le personalità false. Ma account falsi sono già regolarmente creati per persone reali senza che queste lo sappiano o vi acconsentano, e le discussioni anonime sono essenziali per un sano dibattito politico, soprattutto quando chi parla proviene da comunità marginalizzate o penalizzate. Non abbiamo un sistema di autenticazione in grado al contempo di proteggere la privacy e che sia efficace per miliardi di utenti.

Possiamo sperare che la nostra capacità d'identificare le personalità artificiali tenga il passo con la nostra capacità di mascherarle. Se la lotta sempre più feroce tra *deepfake* e rilevatori di *deepfake* può fungere da guida, anche questo non sarà un compito facile. Le tecnologie di offuscamento sembrano sempre un passo avanti alle tecnologie di rilevamento. E le "persone" artificiali saranno progettate per agire esattamente come le persone reali.

In ultima istanza le soluzioni dovranno essere di natura non tecnologica. Dobbiamo riconoscere i limite del dibattito politico in rete, e dare nuovamente priorità alle interazioni di persona, che sono più difficili da automatizzare e ci permettono di sapere che le persone con cui parliamo sono

esseri umani in carne e ossa. Sarebbe una svolta culturale che permetterebbe di prendere le distanze dai testi pubblicati su internet, e di tenersi lontani dai social network e dai thread di commento.

I tentativi di disinformazione sono ormai diffusi in tutto il mondo, e sono praticati in più di settanta paesi. È il metodo ordinario con cui si effettua la propaganda nei paesi con tendenze autoritarie, e sta diventando il modo per portare avanti una campagna politica, che si tratti di un candidato o di una questione specifica.

Le personalità artificiali sono il futuro della propaganda. E anche se potrebbero non essere in grado di spostare il dibattito politico in una direzione o in un’altra, possono facilmente sommergerlo del tutto. Non sappiamo quali siano gli effetti di una simile interferenza sulla democrazia, se non che è nociva, e inevitabile.