# FAKE PEOPLE
## I FALSI PROFILI AUTOMATIZZATI

**Maurizio Tesconi**

*In Rete, la probabilità di incontrare un profilo falso è altissima: solo di recente Facebook ne ha eliminati mezzo miliardo mentre, secondo le stime, su Twitter ce ne sarebbero ancora decine di milioni. Molti di questi sono talmente sofisticati da essere scambiati per persone reali, ma nascondono algoritmi programmati per eseguire in automatico una serie di azioni per raggiungere gli scopi più vari. Alcuni sistemi indirizzati al marketing hanno il compito di lanciare prodotti o attirare clic su una determinata pagina internet, mentre altri possono anche condizionare l'opinione pubblica. Quali sono i rischi che si celano dietro la presenza di schiere di account automatizzati?*

Negli ultimi anni sempre più persone nel mondo utilizzano i social network: Facebook oggi conta più di due miliardi di utenti attivi ogni mese, Instagram circa 800 milioni e Twitter più di 300 milioni. Con questi numeri, le piattaforme social sono diventate uno dei più potenti ed efficaci strumenti di comunicazione e hanno ormai assunto un ruolo centrale nell'ecosistema dell'informazione. Sono state largamente impiegate in contesti come movimenti civili, sensibilizzazione politica, interventi nella sanità pubblica e nella gestione delle emergenze. Tuttavia, come tutti gli strumenti con grandi potenzialità, nascondono anche dei rischi: le cronache recenti hanno portato alla ribalta casi in cui i social sono stati sfruttati anche per fini eticamente discutibili, come la propaganda radicale, il reclutamento da parte di gruppi estremisti, la manipolazione del mercato azionario o la diffusione di false notizie. In caso di obiettivi malevoli, come manipolare o falsare le discussioni online, in genere i malintenzionati ricorrono a software in grado di controllare le azioni di un account, i cosiddetti *social bot*. Questi particolari account sono in grado di compiere in modo automatico una serie di azioni – come scrivere un messaggio, condividere un contenuto o richiedere una connessione con un altro utente – tenendo un comportamento molto simile a quello di una persona in carne e ossa.

Il loro utilizzo è stato rilevato anche in contesti politici, per influenzare le opinioni e, di conseguenza, il voto dei cittadini.
Uno studio elaborato dalla Swansea University (Regno Unito) e dalla University of California, Berkeley (Stati Uniti) ha rilevato indizi sul tentativo di condizionare da parte russa la campagna per il referendum sulla Brexit che si tenne il 23 giugno del 2016 e che ha determinato la decisione del Regno Unito di lasciare l'Unione europea. Sembrerebbe, hanno spiegato i ricercatori, che su oltre 28 milioni di tweet sul tema referendario, un numero importante provenisse da decine di migliaia di account russi che, nei giorni intorno al referendum, hanno veicolato messaggi con orientamento positivo nei confronti del «Leave», e hanno contribuito alla massiccia presenza di tweet a favore dell'uscita dall'Unione europea rispetto a quelli prodotti per sostenere il «Remain».

## UN ROBOT SU TWITTER

Da un punto di vista tecnico, i social bot sono una particolare categoria di bot (un'abbreviazione di robot) definibile come un software automatizzato che gestisce un account social, in grado di prendere decisioni senza l'intervento umano e di adattarsi al contesto in cui opera. La parola «bot» ha assunto nel tempo diversi significati, ma una caratteristica accomuna gli applicativi che ricadono sotto questa definizione: il tentativo di simulare il comportamento umano, cercando di apparire come tale agli occhi degli altri utenti.
In passato il termine è stato usato anche come sinonimo di *web scraper, crawler, indexer* – ovvero software per raccogliere e indicizzare dati su web – ma è stato impiegato altresì per descrivere personaggi autonomi presenti nei primi giochi online multiutente, come in *World of Warcraft*. Nell'ambito della cyber security, denota le macchine compromesse e controllate in remoto da malware, mentre si parla di *botnet* quando questi sistemi instaurano collegamenti tra loro, per esempio per eseguire un attacco coordinato al fine di creare un disservizio (Ddos). Se al termine bot si antepone l'aggettivo social, si fa riferimento a quei particolari software che gestiscono un account su un social network mimando il comportamento di un essere umano. La materia è recente e in continua evoluzione, al punto che non esiste ancora un criterio ufficiale per classificare le centinaia di milioni di social bot che popolano le più famose piattaforme, da Facebook a Twitter, e che si mimetizzano tra i nostri contatti.





Distinguere i social bot dagli account reali non è sempre semplice. Alcuni sono agevolmente individuabili, e si limitano a seguire gli altri account e diffondere spam, altri sono invece più complessi, capaci addirittura di sostenere una conversazione con un utente reale. Uno dei primi utilizzi dei social bot in contesto di elezioni politiche risale al 2010, durante le Massachusetts Special Elections negli Usa. In quell'occasione una rete di account automatizzati ha lanciato una campagna di diffamazione contro la candidata democratica Martha Coakley, che ha poi perso le elezioni per una manciata di punti percentuali. Due anni più tardi, in Russia, gli attivisti politici che su Twitter tentavano di mobilitare e discutere le elezioni presidenziali, si sono scontrati con una campagna pro Cremlino avanzata da social bot che ha sovrastato il movimento di dissenso. E ancora, alcuni ricercatori sostengono che i social bot siano stati utilizzati su Twitter per interferire con la mobilitazione politica in Siria e in Messico, fino agli studi più recenti che suggeriscono un'importante attività di political bot (così sono stati battezzati i social bot che si occupano di propaganda elettorale) in vista del referendum Brexit del Regno Unito del 2016, delle elezioni presidenziali negli Usa del 2016 e in Francia del 2017.

## UNO, NESSUNO E CENTOMILA

I social bot non sono tutti uguali e persino per gli addetti ai lavori cercare di mettere ordine in questo vasto e variegato universo non è semplice. Per distinguerli si possono, ad esempio, osservare le intenzioni, che possono essere malevoli (come nella maggior parte dei casi) o meno. Tra i casi virtuosi ci sono i cosiddetti *news bot*, che partecipano alla diffusione di notizie con il fine di offrire un servizio alla comunità, come il nuovo servizio dell'Istituto Nazionale di Geofisica e Vulcanologia che, dagli account istituzionali, invia in automatico un tweet in caso di terremoto.
Essi si differenziano anche per il livello d'intervento umano richiesto per lavorare. Gli account fasulli possono essere completamente automatizzati e non aver bisogno di alcuna attività da parte di un operatore umano, oppure ibridi, presentando un comportamento automatizzato e prevedendo, al tempo stesso, l'ausilio dell'operatore. Questi ultimi sono anche chiamati *cyborg* e sono in grado di compiere un'azione sotto il controllo di una persona, che può dare un comando da far eseguire in sincronia a migliaia di account.

Un tale esercito di bot atto a influenzare l'opinione pubblica apre potenziali scenari molto pericolosi per la libertà dei singoli paesi. Il criterio principale per distinguere i tipi di social bot riguarda il comportamento che assumono all'interno della Rete, dove i bot più elementari che si possono incontrare sono i cosiddetti fake follower. Il loro scopo è di creare connessioni con determinati utenti (spesso sotto pagamento) per aumentarne popolarità e influenza. Come durante la campagna elettorale degli Usa del 2012, quando l'account Twitter di Mitt Romney, al tempo candidato alle elezioni presidenziali, ha visto un improvviso aumento di seguaci, la maggior parte dei quali si è rivelata essere composta da *fake follower*. Quel che è certo, è che oggi esiste un vero e proprio mercato, con siti specializzati che offrono pacchetti di centinaia o migliaia di sostenitori a prezzi decisamente abbordabili.

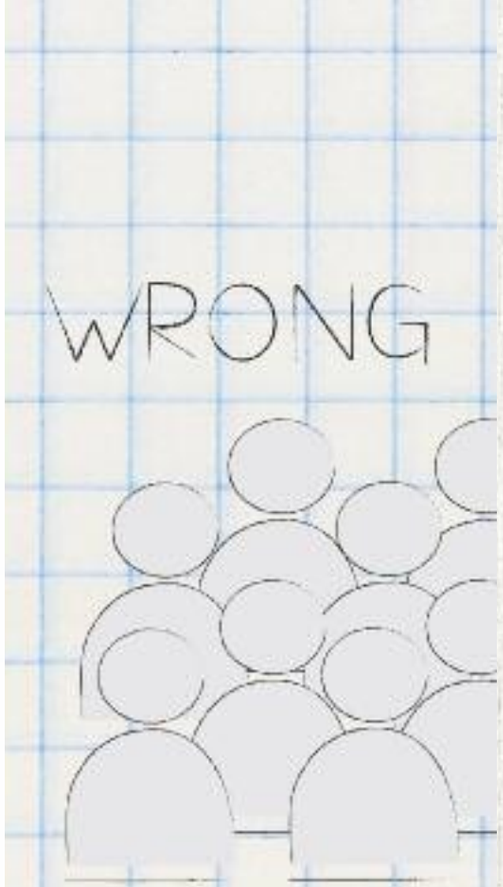
Se ci sono persone disposte a comprarsi fake follower è perché averne tanti paga, non solo in politica. A seconda della platea di seguaci, un influencer nel campo della moda riesce a ottenere un compenso, oppure ad accedere in modo totalmente gratuito a una serie di benefit, come cene in ristoranti di lusso e abbigliamento griffato, in cambio di un post sulla propria pagina.

Sono molti i giovanissimi che provano a lanciarsi in questa impresa, anche se non tutti con buoni risultati: di recente ha fatto scalpore la storia di Elle Darby, un'influencer inglese di 22 anni, che si è vista rifiutare da un albergo di Dublino l'offerta di un pernottamento gratuito in cambio di una copertura social del soggiorno.

In ambito marketing troviamo invece gli *spam bot*, che hanno come obiettivo la condivisione di un determinato tipo di contenuto, in genere una pubblicità. Si possono sottorganizzare in: *displayer*, che condividono i contenuti sulla propria pagina; *bragger*, che compartecipano i contenuti su un feed come, ad esempio, un tweet su Twitter; *poster*, che postano i contenuti sulle pagine degli altri utenti; *whisperer*, che inviano messaggi privati diretti alla propria rete di connessioni.

Ci sono poi i *pay bot* che copiano contenuti da altre fonti e li riportano sotto forma di link a pagine che offrono un pagamento per il traffico generato. In questo modo gli sviluppatori riescono a ottenere un guadagno dall'attività dei loro bot.

In alcuni casi i link possono invece portare l'utente su indirizzi compromessi, che infettano la macchina e consentono l'accesso di terzi alle informazioni personali.







## SE IL BOT GIOCA IN BORSA

Una categoria di bot in grande ascesa è quella dei bot finanziari. Mentre in passato la selezione degli investimenti avveniva all'interno dei ristretti circoli della finanza, aperti solo agli addetti ai lavori, oggi sono già milioni gli investitori finanziari che utilizzano i social media per maturare decisioni sui loro investimenti.

Dal momento che diversi studi suggeriscono che le informazioni provenienti dagli utenti Twitter abbiano un notevole valore per prevedere l'andamento dei mercati, a esse la finanza si sta sempre più affidando attraverso la creazione di algoritmi di trading automatici che acquisiscono e rilevano informazioni da varie fonti (tra cui i social network) per prendere decisioni e fare trading.

Ma la storia recente suggerisce che questi sistemi non sempre agiscono in modo corretto.

Il 6 maggio 2010, noto come Flash Crash Day, si verificò senza alcun apparente motivo un improvviso crollo dell'indice Dow Jones della Borsa valori di New York. Un evento simile non si era mai verificato prima e per questo è finito sotto la lente di un gruppo di ricercatori. Secondo gli studiosi la colpa sarebbe dei sistemi di trading automatici, che avrebbero valutato in modo scorretto le informazioni provenienti dal web.

Il 23 aprile 2013 la Syrian Electronic Army, un gruppo di pirati informatici tra i più ricercati al mondo, hackerò l'account Twitter della Stampa internazionale ufficiale degli Usa, postando una notizia su un fantomatico attacco terroristico alla Casa Bianca, con tanto di ferimento del presidente Obama. Quel giorno i mercati crollarono rovinosamente, per poi risalire una volta scoperto che si trattava di un falso.

Un anno più tardi una finta discussione su Twitter (creata ad hoc), declamava la solidità di un'azienda di servizi tecnologici, la Cynk Technology. Così, mentre un gruppo di utenti sincronizzati postava tweet positivi riguardo la società, gli algoritmi di trading automatici rilevarono la conversazione e iniziarono a investire pesantemente nelle azioni dell'azienda. Tutto questo portò a un incremento nel valore di mercato dell'azienda – che, in realtà, non aveva dipendenti a tempo pieno, né un fatturato – fino a gonfiarne il valore di 5 miliardi di dollari. Quando gli analisti si accorsero dell'orchestrazione e lo stock trading venne interrotto, gli ignari investitori che si erano affidati agli algoritmi si ritrovarono per le mani azioni cartastraccia.

## A CACCIA DI BOT

In Rete la probabilità di incontrare un profilo falso è altissima: solo di recente Facebook ne ha eliminati mezzo miliardo, mentre secondo le stime su Twitter ce ne sono ancora decine di milioni. Il *bot-spotting* è una delle sfide più impegnative. Per smascherare gli account falsi, i ricercatori devono prima di tutto capire quali siano i comportamenti che li distinguono dalle persone reali, le quali, in genere, hanno condotte molto più eterogenee: twittano con discontinuità, non rispettano orari o intervalli regolari e non sono sincronizzati né coordinati.

Partendo da presupposti come questi, nonostante il settore sia nuovo e in espansione, la letteratura scientifica conta già numerosi studi e tecniche per l'individuazione automatica dei social bot che, al momento, seguono due approcci principali.

Nel primo si utilizzano tecniche di *machine learning* basate su classificatori automatici che hanno imparato a riconoscere i bot dalle caratteristiche principali del profilo, dopo un addestramento su milioni di tweet provenienti da umani e bot verificati manualmente. Questo tipo di approccio si concentra sull'analisi di un singolo account alla volta e può sfruttarne tutte le informazioni, dalla foto del profilo, al numero di follower o ai contenuti della timeline dell'utente. Ad esempio, per contare il numero di bot che potrebbero aver influenzato le ultime elezioni africane, l'agenzia inglese di comunicazione Portland ha tenuto traccia di centinaia di caratteristiche, tra cui l'età di un account e l'uso di emoticon. I ricercatori hanno preso in esame anche la distribuzione dei punti esclamativi, visto che gli esseri umani li inseriscono nei post in modo imprevedibile, mentre la maggior parte dei bot o usa molti punti esclamativi oppure non li usa proprio.

Nel secondo approccio, chi va a caccia di fake account analizza gruppi di profili cercando di trovare delle somiglianze nel comportamento o nella descrizione. Le timeline possono essere codificate in sequenze di caratteri, che variano a seconda delle azioni che l'utente compie sul social. Queste sequenze possono poi essere esaminate per cercare di individuare gruppi di bot che presentano sottosequenze in comune abbastanza lunghe. I ricercatori che hanno ideato il metodo parlano di Dna digitale degli utenti, per la somiglianza con le tecniche di analisi biochimica delle sequenze genetiche.

Per riuscire a sopravvivere, e non essere smascherato dagli algoritmi che scoprono automaticamente i comportamenti anomali, il social bot ha la necessità di confondersi nell'ecosistema social e quindi avere un profilo che assomigli il più possibile a quello degli utenti reali. Informazioni come biografia e immagine del profilo possono essere caricate manualmente da un operatore o in modo autonomo dal social bot, che può raccogliere immagini e testi dal web. In altri casi, il social bot è in grado di disporre di un profilo compromesso, ovvero appartenuto in precedenza a un utente reale, che ne ha perso l'accesso come, ad esempio, può accadere a chi abbia subìto un attacco di phishing.

Se la creazione del profilo non viene curata, diventa più semplice riconoscerne la falsità: ad esempio, quando l'immagine del profilo rappresenta una celebrità o un personaggio di fantasia, oppure se immagine del profilo e contenuti condivisi si dimostrano molto discordanti. Gli account di tipo fake follower, avendo come unico obiettivo quello di incrementare il numero di fan di un account, di solito hanno profili molto semplici, vengono creati nello stesso istante insieme ad altre migliaia e possono essere programmati tutti per seguire uno o più account delle persone che hanno comprato il servizio. Se, al contrario, un profilo fasullo è ben curato, talvolta può sfuggire anche al migliore algoritmo di detection. Per questo motivo anche le stesse piattaforme social, nel tentativo di difendersi dal proliferare del fenomeno, si affidano molto anche alle segnalazioni da parte degli utenti. È il caso di Facebook che, dopo la segnalazione, innesca un processo che impegna le persone a rivedere i profili cercando di capire se abbiano comportamenti scorretti o siano dei bot. Twitter, che si affida soprattutto a tecniche di individuazione automatica, al momento non ha ottenuto risultati degni di nota: uno studio dello scorso anno a cura del Consiglio Nazionale delle Ricerche evidenzia che la piattaforma è capace di individuare e bloccare solo il 60% dei bot più semplici, mentre riesce a scovare solo il 4% dei social bot più sofisticati.


## EFFETTI OFFLINE

I social bot sono veramente efficaci? Tra le domande che ruotano intorno al mondo dei fake people, forse questa è quella più interessante. Il dato di fatto è che i bot possono gonfiare l'importanza di un argomento o offuscare la reputazione di un altro, inondando i social network con notizie false e manipolando la valuta di Twitter: like e condivisioni, follow e retweet. Al momento non è chiaro come questo si traduca, per esempio, in voti alle elezioni. Nessuno è riuscito a dimostrare in modo scientificamente rigoroso che i bot abbiano effettivamente cambiato le preferenze degli elettori, né che abbiano influenzato le decisioni sull'opportunità di votare. Nonostante sia chiaro che i political bot, e i social bot più in generale, producano effetti online, non lo è altrettanto se le conseguenze si ripercuotano sul comportamento degli utenti al di fuori della Rete. È difficile riuscire a dimostrare se entrare in contatto con l'attività social bot possa influenzare la vita reale di un utente, perché aspetti come il numero di seguaci o di contenuti condivisi non sono sempre sufficienti a determinare l'influenza di un account. Questa sembra dipendere più dalla fiducia, a sua volta correlata con l'attitudine del social bot di mimare il comportamento umano. Anche la posizione di un account in un grafo sociale sembra incidere sulla sua capacità di influenza, insieme alla lingua utilizzata, mentre studi riguardanti l'interdipendenza tra popolarità di un account e la sua influenza chiariscono che la prima non è un requisito per essere influenti e viceversa. Inoltre, anche se una social botnet riesce a intervenire sulle tendenze di un social, non sempre ci si trova davanti a un atto di manipolazione efficace. Nel frattempo meglio non sottovalutare l'effettiva influenza di un social bot: un individuo non cambierà la propria convinzione politica solo per il fatto di aver letto un semplice messaggio, ma potrebbe verificarsi un condizionamento più sottile che, anche se ancora tutta da dimostrare, potrebbe avere un'ampia portata