medialaws

Anticipazioni

Il reato di sostituzione di persona online di fronte a *socialbot* e ritratti *"AI-generated"*. In favore di una interpretazione estensiva dell'art. 494 del codice penale

Alessandro Tedeschi Toschi

Il reato di sostituzione di persona online di fronte a *socialbot* e ritratti "*AI-generated*". In favore di una interpretazione estensiva dell'art. 494 del codice penale

Abstract

La libertà e la facilità di accesso ai *social media* hanno permesso nuove forme di compimento di alcuni reati, quali la diffamazione, l'adescamento e la sostituzione di persona. Ulteriore aggravio dei possibili danni realizzabili sulle piattaforme è la presenza di *software* in grado di potenziare ulteriormente la diffusione di contenuti lesivi. Tra questi vi sono i *socialbot*, dei programmi capaci di interagire con altri utenti, dando l'impressione di essere delle persone vere. In aggiunta, le più moderne estensioni dell'intelligenza artificiale sono oggi in grado di realizzare dei ritratti realistici di persone inesistenti. La combinazione di queste tecnologie permette la creazione di innumerevoli profili *social*, ciascuno dotato di un'identità fittizia ma credibile, facilmente e velocemente gestibili ed orientabili. La possibilità di fingere efficacemente l'esistenza di un ampio numero di persone può essere sfruttata per molteplici fini ma, in ogni caso, essa causa una lesione della fede pubblica, un bene giuridico tutelato in Italia, tra gli altri, dall'art. 494 c.p.

The freedom and ease of access to social media have enabled new forms of committing certain offences, such as defamation, solicitation and impersonation. Further aggravating the possible harm that can be done on these platforms is the presence of software that can enhance even more the dissemination of harmful content. These include social bots, programs capable of interacting with other users, giving the impression of being real people. In addition, the most modern extensions of artificial intelligence are now able to create realistic portraits of non-existent people. The combination of these technologies allows the creation of innumerable social profiles – each with a fictitious but credible identity – which can be easily and quickly managed and oriented. The possibility of effectively faking the existence of many people can be exploited for various purposes, but in any case, it causes damage to public faith, which is protected in Italy by – among others – Article 494 of the Criminal Code.

Sommario

1. Social media e *socialbot*, pericoli interconnessi

La recente diffusione dell'utilizzo dei *social media* come strumenti di comunicazione delle opinioni ha portato a dei veri e propri sconvolgimenti nei modi e nei tempi di propagazione del proprio pensiero. Grazie alle architetture digitali ospitate su infrastrutture *hardware* capaci di gestire in autonomia e con precisione l'immensa mole di

dati generati dalle interazioni degli utenti (si calcola, ad esempio, che nel solo corso dell'agosto 2021 siano stati pubblicati 575.000 *tweet* ogni minuto),[1] gli utenti dei *social* sono in grado di condividere i propri contenuti senza alcun controllo *ex ante* realmente efficace (gli algoritmi di verifica dei contenuti implementati da alcuni dei gestori di queste piattaforme si sono dimostrati spesso incapaci di contrastare il fenomeno delle cosiddette *fake news*).[2]

All'assenza di forme di verifica dei contenuti pubblicati su queste piattaforme si aggiunge anche la mancanza di vere forme di accertamento dell'identità dei loro utenti. Tale libertà di accesso ai servizi digitali di condivisione, che ha permesso in contesti non democratici l'espressione del proprio dissenso senza timore di censure o ritorsioni, ha anche reso possibile il compimento di reati contro altri diritti fondamentali dell'individuo. In molti hanno potuto approfittare delle forme di anonimato e di mascheramento dell'identità garantite da internet – e dai *social media* – per compiere delitti quali la diffamazione, l'adescamento e la sostituzione di persona (spesso con l'ulteriore obiettivo di realizzare delle truffe).

Le tecnologie digitali di interazione online sono già ritenute degli utili strumenti con cui ampliare ed aggravare la portata lesiva di certi crimini. In particolar modo per i reati di diffamazione, i giudici della Corte di cassazione hanno da tempo riconosciuto la natura delle piattaforme *social* di moltiplicatori dell'estensione della diffusione online di contenuti e, quindi, il potenziale aggravamento delle lesioni della reputazione della persona offesa.[3] In tempi più recenti persino la Corte costituzionale ha sottolineato come i caratteri di velocità ed estensione della trasmissione su Internet di un messaggio lesivo dell'onore altrui possano condurre a dei pregiudizi per la vittima estremamente maggiori rispetto a quelli causati dal compimento del medesimo delitto in ambito non digitale.[4]

Ad ulteriore aggravio delle possibilità di danneggiare le persone tramite gli strumenti digitali di diffusione del proprio pensiero vi sono alcuni *software* in grado di potenziare ulteriormente la comunicazione tramite Internet di contenuti pericolosi o lesivi dei diritti altrui. Tra questi la recente cronaca politica ha evidenziato la categoria dei *bot* – un termine molto ampio e da molti lamentato come foriero di confusione – ossia dei programmi *software* capaci, senza alcun intervento umano diretto, di attivarsi a seguito

---

1 *I dati statistici sono disponibili al sito statista.com*.

2 Si veda, a titolo esemplificativo, K. Shu - A. Sliva - S. Wang - J. Tang - H. Liu, *Fake news detection on social media: A data mining perspective*, in *ACM SIGKDD explorations newsletter*, 19(1), 2017, 22 ss.

3 Si veda, ad esempio, Cass. pen., sez. I, 2 gennaio 2017, n. 50, in cui i membri della corte hanno dichiarato che «la diffusione di un messaggio diffamatorio attraverso l'uso di una bacheca "*Facebook*" integra un'ipotesi di diffamazione aggravata». Negli stessi termini si sono poi espressi anche i giudici in Cass. pen., sez. V, 23 gennaio 2017, n. 8482, e Cass. pen., sez. V, 6 settembre 2018, n. 40083. In merito all'attribuzione del carattere di "moltiplicatore" ad internet e ai *social media* si veda M.R. Allegri, *Alcune considerazioni sulla responsabilità degli intermediari digitali, e particolarmente dei social network provider, per i contenuti prodotti dagli utenti*, in *Informatica e Diritto*, 26(1), 2017, 104-105.

4 Corte cost., ord. 26 giugno 2020 n. 132, in cui viene rilevata la «rapidissima e duratura amplificazione degli addebiti diffamatori determinata dai *social networks* e dai motori di ricerca in internet, il cui carattere lesivo per la vittima - in termini di sofferenza psicologica e di concreti pregiudizi alla propria vita privata, familiare, sociale, professionale, politica - e per tutte le persone a essa affettivamente legate risulta grandemente potenziato rispetto a quanto accadeva anche solo in un recente passato».

determinate condizioni, di individuare le azioni da compiere per portare a termini i compiti affidatigli e di eseguirle, riconoscendo l'ambiente informatico in cui operano e adattandosi ai cambiamenti di questo.[5] Seguendo le distinzioni compiute dalla più autorevole letteratura in materia,[6] si farà qui riferimento ai soli *bot* che sono capaci di agire su internet e di controllare un profilo di un *social medium* imitando il modo in cui un essere umano interagisce e dissimulando la propria vera natura, i quali vengono chiamati *socialbot*.[7]

1.1. Interazioni sui *social media* e replicabilità da parte dei *socialbot*

Al fine di meglio comprendere l'incisività della presenza dei *socialbot*, è importante sottolineare come tutte le attività "*social*" siano delle forme di interazione preimpostate e standardizzate per le quali esistono appositi e ben identificati pulsanti virtuali. Risulta, quindi, semplice istruire un programma (il *socialbot*) ad individuare le azioni da compiere (ossia i tasti virtuale da premere) e a compierle solo al verificarsi di determinate condizioni che esso è ben capace di riconoscere (come, ad esempio, il fatto che ad un determinato *post* su Facebook di qualcuno questo particolare *bot* non abbia ancora messo un "Mi piace").

Il ruolo di "elevatori alla potenza" della diffusione di contenuti all'interno dei *social network* (che a loro volta vengono considerati come dei moltiplicatori della diffusività) viene svolto dai *socialbot* attraverso diverse condotte, che possono variare anche a seconda della strutturazione del *medium* considerato. Ad ogni modo, si può far rientrare le interazioni compiute all'interno delle reti sociali online in tre categorie: (A) la produzione di contenuti originali (come la pubblicazione di *post* e commenti); (B) la condivisione di contenuti di altri (quali le azioni "Condividi" o *retweet* dei contenuti di terzi, anche esterni alla piattaforma, su Facebook e Twitter); (C) l'apposizione di "reazioni" ai contenuti di altri (ad esempio i "Consiglia" di LinkedIn o gli *upvote* di Reddit). La distinzione proposta delle azioni che possono essere compiute sui *social media* permette una maggiore comprensione delle dinamiche interne ai *social network* e una più accurata distinzione delle ragioni della rilevanza di ciascuna di esse per il compimento

[5] Per una definizione approfondita di *bot* si vedano S. Franklin - A. Graesser, *Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents*, in J.P. Müller - M.J. Wooldridge - N.R. Jennings (a cura di), *Intelligent Agents III Agent Theories, Architectures, and Languages*, Heidelberg, 1996, 21 ss.; M. Tsvetkova - R. García – Gavilanes - L. Floridi - T. Yasseri, *Even good bots fight: The case of Wikipedia*, in *PLoS ONE*, 12(2), 2017.

[6] N. Abokhodair - D. Yoo - D.W. McDonald, *Dissecting a social botnet: Growth, content and influence in Twitter*, in *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*, 2015, 839 ss.; E. Ferrara - O. Varol - C. Davis - F. Menczer - A. Flammini, *The rise of social bots, in Communications of the ACM*, 59(7), 2016, 96 ss.; S. Stieglitz - F. Brachten - B. Ross - A.K. Jung, *Do social bots dream of electric sheep? A categorisation of social media bot accounts*, in *Proceedings of the 28th Australasian Conference on Information Systems (ACIS)*, 89, 2017; M. Tsvetkova - R. García-Gavilanes - L. Floridi - T. Yasseri, *Even good bots fight: The case of Wikipedia*, cit.

[7] Per una definizione approfondita di *socialbot* si veda Y. Boshmaf - I. Muslukhov - K. Beznosov - M. Ripeanu, *The socialbot network: when bots socialize for fame and money*, in *Proceedings of the 27th annual computer security applications conference*, 2011, 93 ss.

di attività di diffusione potenziata dai *socialbot*. Sebbene l'attenzione rispetto a questi agenti *software* autonomi sia dovuta prevalentemente al loro impiego per propaganda politica e distorsione dei dibattiti online,[8] non si può ignorare come essi possano facilmente essere usati anche per il compimento di reati contro la persona.[9]

Inoltre, la suddivisione delle possibili azioni che possono essere compiute sui *social media* permette di evidenziare con maggiore facilità come esse, nonostante le peculiarità di ciascuna piattaforma, permettano forme di compimento di un reato – spesso aggravato – già accertate dalla nostra giurisprudenza. Infatti, è possibile rinvenire per ciascuna delle categorie qui proposte delle decisioni della Suprema Corte che ben illustrano le ragioni dell'integrazione di una forma aggravata di reato compiuto all'interno dei *social network*.

(A) Produzione di contenuti originali – in merito a questa categoria di attività la Corte di cassazione ha da tempo mostrato come essa possa portare al compimento di reati. In particolare, gli Ermellini hanno dichiarato che tale condotta «integra un'ipotesi di diffamazione aggravata» ai sensi dell'art. 595 c.p., sia nei casi di pubblicazione di *post* denigratori[10] che di commenti calunniosi ad essi.[11] In quest'ambito hanno sottolineato come le bacheche dei *social network*, «destinate per comune esperienza ad essere consultate da un numero potenzialmente indeterminato di persone, secondo la logica e la funzione propria dello strumento di comunicazione e condivisione telematica, che è quella di incentivare la frequentazione della bacheca da parte degli utenti, allargandone il numero a uno spettro di persone sempre più esteso», siano un mezzo di pubblicità idoneo «a coinvolgere e raggiungere una vasta platea di soggetti, ampliando – e aggravando – in tal modo la capacità diffusiva del messaggio lesivo della reputazione della persona offesa».[12]

Come già detto, le condotte diffamatorie sulle bacheche dei *social network* sono il risultato di interazioni standardizzate ed è la struttura della piattaforma *social* a rendere automaticamente e in modo non supervisionato visibili ad un numero di persone indeterminato – ma potenzialmente molto ampio – le frasi offensive scritte. Risulta, quindi, facile immaginare un *socialbot* che sia capace, tramite opportuna programmazione, di pubblicare in autonomia contenuti dal carattere diffamatorio, in special modo alla luce dei recenti progressi nella generazione di testi da parte di intelligenze artificiali di *machine learning* (IA di ML).[13] In realtà, l'utilizzo di *socialbot* per la diffusione

[8] Si vedano, *ex multis*, P.N. Howard - B. Kollanyi, *Bots, #Strongerin, and #Brexit: Computational Propaganda During the UK-EU Referendum*, 2016; E. Ferrara, *Disinformation and social bot operations in the run up to the 2017 French presidential election*, in *First Monday*, 22(8), 2017; L. Luceri - A. Deb - S. Giordano - E. Ferrara, *Evolution of bot and human behavior during elections*, in *First Monday*, 24(9), 2019.

[9] Si veda, ad esempio, A. Tedeschi Toschi - G. Berni Ferretti, *Social media, profili artificiali e tutela della reputazione. Come l'avvento dei social bot per la gestione dei profili social possa rappresentare una grave minaccia per la reputazione delle persone e quali potrebbero essere le risposte a tale pericolo*, in *Rivista Italiana di Informatica e Diritto*, 3(2), 2021, 107 ss.

[10] Cass. pen., sez. I, 8 giugno 2015, n. 24431; Cass. pen., sez. I, 2 gennaio 2017, n. 50; Cass. pen., sez. V, 23 gennaio 2017, n. 8482; Cass. pen., sez. V, 1 febbraio 2017, n. 4873; Cass. pen., sez. V, 6 settembre 2018, n. 40083.

[11] Cass. pen., sez. V, 22 settembre 2004, n. 47452.

[12] Cass. pen., sez. I, 2 gennaio 2017, n. 50.

[13] Si veda, ad esempio, ChatGPT, un *software* di intelligenza artificiale generativo di testi basato su

automatizzata di contenuti testuali precede l'impiego di forme di intelligenza artificiale e può prescindere da questo.[14]

(B) Condivisione di contenuti di altri – riguardo a questa tipologia di condotta verso i contenuti di soggetti terzi, i giudici di piazza Cavour hanno già ritenuto che essa sia la dimostrazione della volontà di «amplificar[li] attraverso il proprio comportamento»[15] [16] e, quindi, qualora detti contenuti integrino un reato (nel caso di specie qui citato di diffamazione aggravata), essa stessa integrerà, a sua volta, un reato. In particolare, gli Ermellini hanno precisato come la condotta di condivisione delle critiche ad una persona offesa «potrebbe assumere in astratto rilevanza penale soltanto qualora potesse affermarsi che con il proprio messaggio l'imputato aveva consapevolmente rafforzato la volontà dei suoi interlocutori di diffamare».[17] In questo senso l'azione di condivisione di contenuti lesivi della dignità di una persona, ossia di loro ulteriore diffusione tra gli utenti di un *social medium*, integra pienamente un comportamento esteriore idoneo ad arrecare un contributo apprezzabile alla produzione del danno nei confronti della vittima.[18]

---

tecniche di *machine learning* sviluppato dalla OpenAI. Si veda anche il caso dell'IA chiamata Tay, descritto in A. Tedeschi Toschi - G. Berni Ferretti, *La responsabilità per la diffamazione compiuta da un'Intelligenza artificiale. Possibili scenari costruiti partendo dall'esempio dell'IA Tay*, in *Ciberspazio e diritto*, 24(74), 2023, 173 ss.

[14] Per fare un esempio, in S. Cresci - R. Di Pietro - M. Petrocchi - A. Spognardi - M. Tesconi, *Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling*, in *IEEE Transactions on Dependable and Secure Computing*, 15(4), 2018, 561 ss., viene riportato come dei profili gestiti da questi agenti *software* pubblicassero periodicamente aforismi portando ad ipotizzare che essi si limitassero a pescare da un elenco preimpostato di frasi a cui avevano accesso. Ad ogni modo, gli attuali avanzamenti tecnologici delle IA di ML permettono la creazione di contenuti testuali sempre nuovi, spesso adattandoli a quanto scritto in precedenza dagli interlocutori.

[15] Cass. pen., sez. V, 29 gennaio 2016, n. 3981. In questa particolare sentenza è necessario compiere una fondamentale distinzione lessicale, in quanto i giudici nell'utilizzare il termine "condividere" intendono riferirsi alla partecipazione alla discussione concordando sulla criticabilità del soggetto considerato ma «senza ricorrere alle espressioni offensive utilizzate da altri, né dimostrando di volerle amplificare attraverso il proprio comportamento», mentre nel gergo delle attività sui *social media* tale verbo indica la ripubblicazione delle medesime espressioni offensive, che di fatto ne determina una volontaria amplificazione.

[16] A questa conclusione è giunta praticamente tutta la dottrina che si è interessata alle azioni di ricondivisione di contenuti di altri da parte di utenti di piattaforme *social* online. Si vedano, *ex multis*, N. Anstead - B. O'Loughlin, *The Emerging Viewertariat and BBC Question Time: Television Debate and Real-Time Commenting Online*, in *The International Journal of Press/Politics*, 16(4), 2011, 440 ss.; E. Bakshy - J.M. Hofman - W.A. Mason - D.J. Watts, *Everyone's an influencer: quantifying influence on Twitter*, in *Proceedings of the fourth ACM international conference on Web search and data mining*, 2011, 65 ss.; T.A. Small, *What the hashtag? A content analysis of Canadian politics on Twitter*, in *Information, Communication & Society*, 14(6), 2011, 872 ss.; G. Elmer, *Live research: Twittering an election debate*, in *New Media & Society*, 15(1), 2013, 18 ss.; T. Highfiel - S. Harrington - A. Bruns, *Twitter as a Technology for Audiencing and Fandom*, in *Information, Communication & Society*, 16(3), 2013, 315 ss.; A.O. Larsson - H. Moe, *Studying political microblogging: Twitter users in the 2010 Swedish election campaign*, in *New Media & Society*, 14-5, 2012, 729 ss.; S. Meraz - Z. Papacharissi, *Networked gatekeeping and networked framing on #Egypt*, in *The International Journal of Press/Politics*, 18(2), 2013, 138 ss.; X.W. Zhao - J. Wang - Y. He - J. Nie - X. Li, *Originator or propagator? Incorporating social role theory into topic models for twitter content analysis*, in *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, 2014, 1649 ss.

[17] Cass. pen., sez. V, 29 gennaio 2016, n. 3981.

[18] In quest'ambito si veda il principio ribadito in Cass. pen., sez. V, 10 gennaio 2022, n. 319. In essa, tra le varie questioni, il collegio giudicante ha statuito che il direttore di un quotidiano online, pubblicando integralmente le dichiarazioni diffamatorie di un intervistato, ne ha agevolato in modo apprezzabile

Questa condotta risulta ancora più elementare da essere eseguita, dato che la maggior parte dei *social media* possiede degli appositi pulsanti virtuali con cui poter facilmente propagare i contenuti creati da terzi (si pensi al tasto "Condividi" su Facebook). In questo caso, la delittuosa attività di diffusione di messaggi lesivi del buon nome di qualcuno non abbisogna nemmeno di un minimo di creatività o di abilità argomentativa, basta cliccare un tasto già predisposto dalla piattaforma. Ancora più semplice è, così, programmare un *socialbot* a compiere questa azione.

Vi è da notare come l'automazione di questa attività possa avere, però, dei risvolti inaspettati. In caso di una programmazione priva di analisi del contenuto delle pubblicazioni di terzi da condividere, l'azione di questi agenti *software* viene compiuta senza consapevolezza di quale opinione o condotta venga così rafforzata. In altre parole, questa condotta può avvenire "al buio", solamente sulla base dell'identità del soggetto che li ha originariamente pubblicati ed indifferentemente dalla rilevanza penale delle affermazioni promosse.[19]

(C) Apposizione di "reazioni" ai contenuti di altri – relativamente a quest'ultima categoria di attività online, infine, i Giudici di Piazza Cavour hanno già riconosciuto come essa sia potenzialmente rilevante per la nostra normativa penale, riconoscendo come l'apposizione di un "Mi piace" su Facebook abbia carattere diffusivo, «attesa la […] funzione propalatrice svolta in tale contesto dal *social network*» e sia, quindi, idonea ad avere portata offensiva.[20]

Più recentemente i giudici di cassazione hanno evidenziato come l'apposizione di una semplice "reazione" a contenuti digitali di terzi (nel caso di specie si trattava dei "*like*" ai contenuti presenti sulle bacheche nella piattaforma Facebook) costituisca una manifestazione di «adesione e condivisione dei messaggi»[21] da questi pubblicati e possa integrare pienamente un reato (nel caso di specie quello di istigazione all'odio razziale). Gli Ermellini si sono anche premurati di sottolineare come ciò dipenda dalle modalità di funzionamento della diffusione automatizzata – ed acritica – dei messaggi inseriti nelle bacheche di Facebook. Tale modalità di propagazione automatica di contenuti «dipende dalla maggiore interazione con le pagine interessate da parte degli utenti» e ha un effetto moltiplicativo, dato che sono «le interazioni che consentono la visibilità del messaggio ad un numero maggiore di utenti i quali, a loro volta, hanno la possibi-

l'opera di discredito del buon nome del danneggiato, risultando così concorrente nella sua realizzazione. In altre parole, l'aver contribuito alla diffusione di un messaggio diffamatorio pronunciato da terzi integri il compimento del medesimo reato.

[19] Non si vuole qui avanzare l'ipotesi che una condivisione "al buio" di contenuti precluda di per sé il concorso della volontà del soggetto che ha programmato o amministrato l'agente *software* al verificarsi di determinati eventi lesivi dei diritti di terzi. È importante, però, sottolineare come l'indagine sull'elemento soggettivo di colui che ha creato o gestito il *socialbot*, alla luce di queste particolarità, dovrà essere molto approfondita (soprattutto per determinare se la realizzazione del danno sia stata dolosa o colposa).

[20] Cass. pen., sez. V, 12 dicembre 2017, n. 55418. La decisione presa dalla Corte era relativa alla questione se la pubblicazione sul proprio profilo Facebook di video inneggianti allo Stato Islamico e l'apposizione di "Mi piace" ad altri integrassero o meno il delitto di istigazione a delinquere previsto dall'art. 414 del codice penale.

[21] Cass. pen., sez. I, 9 febbraio 2022, n. 4534, riportata anche in Redazione, *Se un like ad un post razzista è istigazione all'odio*, in *Diritto di Internet.it*, 2022.

lità di rilanciarne il contenuto».

Come già visto per altre interazioni, anche l'apposizione di una "reazione" (come un *like* di Facebook) dipende da strutturazioni del *social medium* preimpostate e facilmente riconoscibili. Di conseguenza, risulta facile intuire come, sfruttando i caratteri dell'algoritmo che regola la continua esibizione agli utenti di nuovi contenuti (il cosiddetto "*newsfeed*" delle piattaforme), l'impostazione di un *socialbot* perché interagisca secondo queste modalità con i contenuti di terzi possa integrare una condotta di loro diffusione tra un numero indeterminato di persone.

Vi è da fare un'importante precisazione in merito a questa terza tipologia di azioni. In essa si vuole far rientrare anche le cosiddette "visualizzazioni" di contenuti multimediali, anche se tale scelta potrebbe apparire come una forzatura. L'inserimento in questa categoria di una condotta che si potrebbe definire come "passiva" e non "di reazione" ai contenuti di altri deriva dal fatto che essa rimane, comunque, registrata come un'azione di interesse per detti contenuti e viene utilizzata come indicatore delle loro popolarità. L'interesse in questa sede per tale azione passiva deriva dal fatto che, come già egregiamente notato dalla Corte di cassazione, gli algoritmi delle piattaforme sono impostati in modo da diffondere ulteriormente – e senza controlli *ex ante* realmente efficaci – tra gli utenti quei contenuti che sono considerati come popolari, ossia che hanno un alto numero di visualizzazioni (si riconosce, comunque, che il tema delle visualizzazioni dei contenuti ha bisogno di particolari cautele e contestualizzazioni).[22]

1.2. Alcuni caratteri rilevanti dei *socialbot*

La rilevanza della presenza (occultata) dei *socialbot* sulle piattaforme di *social networking* è acuita anche da alcune delle caratteristiche che questi agenti *software* autonomi condividono con la maggior parte delle tecnologie informatiche. Le più rilevanti da riportare in questa sede sono le seguenti:

(A) Localizzazione – Un primo carattere dei *socialbot* che è importante evidenziare in questo contesto è che, per poter operare correttamente, essi necessitano solamente di una connessione internet, senza che tale allaccio avvenga in un luogo specifico. Il risultato di questa particolarità è la possibilità per gli agenti *software* autonomi impiegati, ad esempio, in campagne propagandistiche o diffamatorie di portare avanti la propria opera anche al di fuori dei confini dei paesi in cui si trovano le vittime e rimanendo, quindi, al di fuori della giurisdizione delle Autorità pubbliche deputate alla tutela dei diritti di queste ultime.[23]

---

[22] Il problema dell'utilizzo di *bot* per incrementare in numero di visualizzazioni su Internet e, quindi, incrementare artificialmente i loro indici di popolarità è già stato notato in S. Weissmann, *How Not to Regulate Social Media*, in *The New Atlantis*, 58, 2019, 58 ss.; B. Stricke, *People v. Robots: A Roadmap for Enforcing California's New Online Bot Disclosure Act*, in *Vanderbilt Journal of Entertainment & Technology Law*, 22, 2020, 838 ss.

[23] Come già perspicacemente rilevato anche dal deputato Seán Kyne nel corso del dibattito del 13 dicembre 2017 sulla proposta di legge irlandese di contrasto alle attività dei *socialbot*, la *Online Advertising and Social Media (Transparency) Bill 2017*. In quest'occasione egli dichiarò che «l'idea che possiamo rintracciare i *bot* e perseguire le persone o le organizzazioni dietro di loro ignora la realtà che se la persona che gestisce i falsi *account* ha sede fuori dallo Stato, come la maggior parte di loro, allora sarà al

(B) Gestione di gruppi – Secondo fattore di importanza per la questione dei *socialbot* è che la gestione di ciascuno di essi non deve necessariamente essere compiuta direttamente da un utente. Esistono programmi, chiamati *botmaster*, che sono in grado di impartire ordini in contemporanea ad una moltitudine di agenti *software*.[24] Questo genere di programma di gestione dei *socialbot* permette ad un utente umano, chiamato *botherder*, di controllare un'intera folla di profili automatizzati descrivendo il singolo comando ad esso e lasciando che sia questo materialmente ad impartirlo ad ogni singolo *software* di gestione dei profili. I *botmaster* fungono, insomma, da strumento di automazione della gestione automatizzata e dissimulata dei profili *social*, finendo col diluire ulteriormente il grado di supervisione delle dinamiche che possono essere scatenate dall'impiego dei *socialbot*.[25]

Sebbene da un punto di vista giuridico la questione non rivesta particolari difficoltà interpretative in merito all'attribuzione di eventuali reati compiuti tramite questi "strumenti di gestione di strumenti di gestione", essa ha comunque rilevanza. Infatti, l'impiego di strumenti atti ad incrementare ulteriormente la portata dannosa di un messaggio diffamatorio o di istigazione all'odio ben finirà col costituire un'aggravante della condotta posta in essere, dato l'incremento della diffusione del contenuto.

Nemmeno la creazione dei profili gestiti in modo automatizzato, poi, deve essere necessariamente compiuta manualmente da un individuo. Esistono, infatti, dei programmi (che possono essere fatti rientrare all'interno della categoria dei *bot*) strutturati in modo da creare in automatico nuovi profili sulle piattaforme di *social networking*.[26] Non esistono, poi, particolari ostacoli alla combinazione di questa tecnologia con quella dei *botmaster*, permettendo la creazione e la gestione automatizzata di folle oceaniche di sostenitori artificiali da parte di chiunque conosca questi programmi o possa permettersi di finanziare questo procedimento.

(C) Flessibilità – Terzo aspetto di preminente rilevanza della natura dei *socialbot* è il fatto essi sono programmati per operare su di una specifica piattaforma ma non per gestire un determinato profilo. Questo perché, come accennato in precedenza, ciascun *social medium* è composto da elementi preimpostati e standardizzati, che questi agenti *software* sono ben in grado di riconoscere e di gestire. Tutto ciò significa che, in caso di sospensione (temporanea o perpetua) di un particolare profilo, non esistono impedimenti tecnici alla consegna al *socialbot* delle credenziali di accesso ad un altro *account*

di là della portata del disegno di legge». La trascrizione del dibattito è disponibile all'indirizzo *oireachtas.ie*.

[24] Per un'analisi delle dinamiche di gestioni di grandi gruppi di *socialbot* si veda, ad esempio, Y. Boshmaf - I. Muslukhov - K. Beznosov - M. Ripeanu, *The socialbot network: when bots socialize for fame and money*, cit., 93 ss.

[25] Si pensi, ad esempio, a dei *socialbot* programmati per rilanciare tutti i messaggi pubblicati dal profilo di una persona di chiara fama e attiva nella beneficenza. Nel caso in cui questo *account* venisse violato ed utilizzato per promuovere una truffa, i *socialbot* continuerebbero a promuovere i messaggi pubblicati da questo senza valutare minimamente il loro contenuto e il loro amministratore potrebbe impiegare diverso tempo per accorgersene, rischiando nel frattempo di indurre numerose persone a cadere vittime di un disegno criminoso. Un utilizzo non adeguatamente supervisionato di strumenti per automatizzare e accelerare la diffusione di contenuti potrebbe causare seri danni.

[26] N. Perlroth, *Fake Twitter Followers Becomes Multimillion Dollar Business*, in *The New York Times*, 5 aprile 2013, in cui viene fatto riferimento ad un *software* che «potrebbe creare fino a 100.000 nuovi account in cinque giorni».

con cui esso possa continuare in autonomia la propria opera di diffusione potenziata di contenuti.

(D) Natura strumentale – Infine, tra i caratteri dei *socialbot* e delle piattaforme su cui operano c'è da evidenziare con enfasi uno più generale che contrassegna questo tipo di tecnologie: la loro natura strumentale.[27] Infatti, come già detto in precedenza, essi possono essere impiegati per la diffusione di qualsiasi tipo di contenuto sulle piattaforme *social* (dalle informazioni sulle variazioni metereologiche fino a teorie cospirazioniste di stampo antisemita). In altre parole, i *socialbot* sono dei mezzi, piuttosto sofisticati, per compiere una serie di attività che possono anche – ma non solo – recare danni ad altri. La conseguenza della somma di tutte le caratteristiche dei *socialbot* qui evidenziate è che essi possono facilmente essere impiegati per creare su di un *social medium* un'ampia schiera di utenti artificiali che sono in grado di dare l'impressione di essere persone profondamente convinte della validità delle opinioni espresse online ed assiduamente concentrate ad incrementarne la diffusione.

2. Forme di dissimulazione della natura dei *socialbot* nelle interazioni online con le persone

La definizione di *socialbot* basa la propria distinzione rispetto ad altri programmi *software* autonomi sull'occultamento della sua natura ai destinatari delle interazioni che compie tramite un profilo *social*. Quest'opera di mimetizzazione viene portata avanti tramite diversi accorgimenti che alcuni legislatori hanno anche cercato di catalogare, addirittura inserendo all'interno di nuove normative di rango primario l'elencazione degli elementi da considerare. L'esempio principale di cristallizzazione legislativa degli indici di valutazione è quello del *Protection from Online Falsehoods and Manipulation Act 2019* (*POFMA*) della Repubblica di Singapore. Alla sezione 40 (4) vengono indicati i fattori che le Autorità pubbliche devono prendere in considerazione per determinare se il profilo investigato sia controllato da un *socialbot*.[28] Un'esplicitazione in sede normativa degli aspetti da valutare ha senz'altro il pregio di fornire una prima serie di indici che permetta alle Autorità di avere una direzione verso cui orientare lo svolgimento delle proprie indagini. Tuttavia, essa non fornisce indicazioni su come ponderare il valore

[27] Questa osservazione viene avanzata anche in M. Lamo-R. Calo, *Regulating Bot Speech*, in *UCLA Law Review*, 66, 2019, 988 ss; J. Horder, *Online Free Speech and the Suppression of False Political Claims*, in *Journal of International and Comparative Law*, 8, 2021, 15 ss. Il riconoscimento della strumentalità di questi agenti *software* autonomi avviene anche – in modo – implicito nella Sezione 7 del *Protection from Online Falsehoods and Manipulation Act 2019* (*POFMA*) della Repubblica di Singapore. In merito a questa normativa si veda A. Tedeschi Toschi-G. Berni Ferretti, *Il contrasto legislativo ai socialbot e le soluzioni avanzate nella Repubblica di Singapore e nella Repubblica d'Irlanda*, in questa *Rivista*, 3, 2022, 352 ss.

[28] La Sezione, infatti, dispone che «nel determinare se un *account online* è un *account online* non autentico o è controllato da un *bot*, il Ministro competente deve tenere conto dei seguenti fattori: (a) se le informazioni utilizzate nella creazione del profilo *online* si riferiscono a un paese o un territorio diverso dal paese o territorio da cui presumibilmente proviene il titolare del profilo; (b) se esista uno schema di attività sospette svolte utilizzando l'*account online*; (c) la data in cui è stato creato il profilo *online*; (d) qualsiasi altro fattore che il Ministro competente consideri rilevante». È importante qui evidenziare come la lettera (d) di questa Sezione permetta all'autorità pubblica investita del compito di svolgere l'indagine di includere qualsiasi altro fattore che consideri «rilevante» in ciascun caso concreto.

di ciascuno degli elementi indicati e ha anche il notevole svantaggio di rendere meno flessibili le loro attività in un ambito – quello informatico – caratterizzato dall'introduzione di profonde innovazioni in tempi rapidissimi.

La principale letteratura scientifica sull'argomento, invece, pone l'accento sulla corrispondenza o meno delle interazioni compiute con dei modelli comportamentali standard estrapolati dalle condotte di un gran numero di profili usati da utenti umani.[29] In altre parole, le metodologie di controllo della possibile presenza di *socialbot* oggi si basano principalmente sull'analisi dei comportamenti tenuti dai profili all'interno dei *social network* e sulla corrispondenza con le condotte abitualmente tenute in essi dalle persone. Questa metodologia di valutazione è operata prendendo in considerazione diversi elementi, quali la ricchezza di informazioni dei profili, la struttura delle reti sociali entro le quali sono attivi, i contenuti da questi pubblicati, il tono delle opinioni espresse ed il tempismo delle loro reazioni.

C'è da evidenziare come venga attribuita una rilevanza relativa agli aspetti della ricchezza di informazioni dei profili, come la presenza di ritratti o di fotografie negli "spazi personali" di questi. La scarsa considerazione riservata a questi elementi – considerati come essenziali dalla nostra giurisprudenza per il compimento del reato di sostituzione di persona sui *social media*[30] – trova giustificazione nel fatto che, come ben evidenziato da Cresci, Di Pietro, Petrocchi, Spognardi e Tesconi nelle loro indagini sui profili che hanno sostenuto uno dei candidati alla contesa elettorale capitolina del 2014,[31] l'opera di mimesi degli *account* consiste non solo nel corredare i profili forniti ai *socialbot* di ritratti e descrizioni credibili ma anche – e soprattutto – nell'assemblare il loro codice sorgente in modo da imitare le condotte medie tipiche di un essere umano.

La ragione della maggior importanza attribuita alle modalità di interazione di un profilo su di una piattaforma *social* deriva anche dal fatto che la strutturazione di internet ha da sempre reso estremamente semplice ricercare ed ottenere copia di immagini ritraenti altre persone. Lo stesso gruppo di ricercatori dell'IIT-CNR guidato da Cresci ha riportato come i profili automatizzati impiegati per la promozione di uno dei candidati all'elezione del sindaco di Roma fossero corredati da fotografie rubate.[32] Ad ulteriore giustificazione della limitata rilevanza della presenza di fotografie vi è anche il fatto che ormai le attuali tecnologie permettono di prescindere integralmente dall'utilizzo di caratteri e generalità di altre persone esistenti od esistite.[33] Infatti, grazie all'utilizzo

[29] L. Alvisi - A. Clement - A. Epasto - S. Lattanzi - A. Panconesi, *Sok: The evolution of sybil defense via social networks*, in *2013 IEEE symposium on security and privacy*, 2013, 382 ss.; L. Luceri - A. Deb - S. Giordano - E. Ferrara, Evolution of bot and human behavior during elections, cit.; S. Cresci, *A decade of social bot detection*, in *Communications of the ACM*, 63(10), 2020, 72 ss.

[30] Si vedano *infra* le sentenze della Corte di Cassazione relativo all'uso abusivo dell'immagine e delle generalità di altre persone inconsapevoli.

[31] S. Cresci - R. Di Pietro - M. Petrocchi - A. Spognardi - M. Tesconi, *The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race*, in *Proceedings of the 26th international conference on world wide web companion*, 2017, 963 ss.; Id., *Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling*, cit., 561 ss.

[32] Ivi, 563 ss.

[33] S. Bond, *That smiling LinkedIn profile face might be a computer-generated fake*, in *npr.org*, 27 marzo 2022, in cui vengono riportate in modo dettagliato le ricerche condotte da Renée DiResta, technical research manager dello Stanford Internet Observatory.

di intelligenze artificiali è oggi possibile generare un numero quasi infinito di ritratti realistici di persone inesistenti (ritratti cosiddetti *AI-generated*).[34] Quindi, ormai può essere abbandonato il cosiddetto "furto d'identità" altrui ed essere comunque facilmente compiuta su internet l'attribuzione a sé o ad altri di un falso nome, di un falso stato ovvero di qualità a cui la legge attribuisce effetti giuridici.

3. La dissimulazione della natura di un *socialbot* come forma di attribuzione a sé di caratteri ai quali la legge attribuisce rilevanza giuridica

All'interno del nostro ordinamento viene attribuita rilevanza penale alle condotte di dissimulazione o di attribuzione di caratteri personali che possano indurre taluno in errore sulla effettiva identità di una persona.[35] Dalla diffusione di internet e dei *social media* questo crimine è diventato estremamente più frequente di quanto non fosse ai tempi della promulgazione del Codice e, di conseguenza, è oggetto della giurisprudenza di Cassazione con sempre maggiore cadenza.

Nelle loro numerose decisioni gli Ermellini hanno avuto modo di precisare come questo delitto venga integrato anche dalla condotta di colui che crei ed utilizzi un *account* online attribuendo ad esso le generalità di un diverso soggetto ed inducendo in errore altre persone attive su internet nei confronti delle quali le false generalità siano state declinate.[36] In altre occasioni hanno specificato come la sostituzione di persona sia integrata anche dalla creazione su di un *social medium* di un profilo con l'utilizzo abusivo dell'immagine di una persona del tutto inconsapevole, associata a generalità o soprannomi di fantasia, e dal suo impiego per condividere contenuti ed indurre in errore i

[34] Si vedano, ad esempio, A. Ramesh - P. Dhariwal - A. Nichol - C. Chu - M. Chen, *Hierarchical Text-Conditional Image Generation with CLIP Latents*, in *arXiv.org*, 2022; T. Karras - S. Laine - M. Aittala - J. Hellsten - J. Lehtinen - T. Aila, *Analyzing and Improving the Image Quality of StyleGAN*, in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, 8110 ss.; S.J. Nightingale - H. Farid, *AI-synthesized faces are indistinguishable from real faces and more trustworthy*, in *Proceedings of the National Academy of Sciences*, 119(8), 2022. Negli articoli citati gli autori illustrano i percorsi di sviluppo di intelligenze artificiali in grado di creare immagini veritiere di alta qualità di persone, animali, oggetti e paesaggi inesistenti (il secondo e il terzo articolo, in particolare, contengono un ampio numero di esempi di ritratti di persone generati artificialmente).

[35] L'art. 494 c.p. stabilisce che: «Chiunque, al fine di procurare a sé o ad altri un vantaggio o di recare ad altri un danno, induce taluno in errore, sostituendo illegittimamente la propria all'altrui persona, o attribuendo a sé o ad altri un falso nome, o un falso stato, ovvero una qualità a cui la legge attribuisce effetti giuridici, è punito, se il fatto non costituisce un altro delitto contro la fede pubblica, con la reclusione fino a un anno».

[36] Si vedano, ad esempio, Cass. pen., sez. V, 14 dicembre 2007, n. 46674, in cui i giudici hanno confermato la condanna all'imputato che aveva creato un *account* di posta elettronica utilizzando il nome della vittima e ledendone l'immagine e la dignità attraverso l'allacciamento di rapporti con altre persone; Cass. pen., sez. III, 3 aprile 2012, n. 12479, in cui è stato riconosciuto sussistere il delitto nella condotta di colui che crei ed utilizzi un *account* ed una casella di posta elettronica, servendosi dei dati anagrafici di un diverso soggetto, inconsapevole, con il fine di far ricadere su quest'ultimo l'inadempimento delle obbligazioni conseguenti all'avvenuto acquisto di beni; Cass. pen., sez. V, 6 luglio 2020, n. 22049, e 5 febbraio 2021, n. 12062.

suoi interlocutori.[37]

La giurisprudenza di Cassazione ha anche valorizzato il fatto che il dolo specifico del delitto di cui alla norma considerata può consistere in varie finalità, anche profondamente diverse tra loro. I giudici di piazza Cavour hanno, infatti, avuto modo di ritenere integrato il reato di sostituzione di persona sia quando il perpetratore lo abbia compiuto per procurarsi un ingiusto profitto economico (con danno del titolare dell'identità abusivamente utilizzata),[38] sia quando ne abbia tratto un vantaggio non patrimoniale, come la possibilità di intrattenere rapporti con altre persone o il soddisfacimento di una propria vanità.[39] Inoltre, le sentenze della Suprema Corte hanno chiarito che la violazione della norma può realizzarsi anche quando, con le condotte tenute, sia stato solamente causato un danno (anche non patrimoniale) ad altri, come la lesione dell'immagine o della dignità delle vittime.[40]

Inoltre, gli Ermellini hanno avuto modo in più occasioni di precisare che l'integrazione del reato – e non solo del suo tentativo – si verifica addirittura nel caso in cui il vantaggio perseguito dall'agente non sia da questo effettivamente raggiunto ma l'uso di mezzi fraudolenti abbia comunque indotto in errore il soggetto passivo del reato.[41]

Sebbene questo articolo del codice penale possa dare l'impressione di essere sufficiente da solo a costituire un valido strumento di difesa della pubblica fede dall'illecita induzione in errore degli interlocutori online dei *socialbot* – soprattutto alla luce dell'opinione recentemente espressa dalla Corte di Cassazione che «oggetto della tutela penale è l'interesse riguardante la pubblica fede»[42] – vi sono da tenere in considerazione diversi peculiari aspetti della tecnologia di questi agenti *software* autonomi che rendono l'applicazione della norma meno scontata di quanto non possa apparire a prima vista.

(A) Attribuzione a sé o ad altri – La prima problematica che riguarda l'applicazione dell'art. 494 c.p. anche alle attività compiute dai *socialbot* è che la lettera della norma prevede che si integri il reato sostituendo «la propria all'altrui persona, o attribuendo a sé o ad altri» identità o caratteri non propri. Il dubbio interpretativo che potrebbe sorgere qui deriva dal fatto che detti caratteri non sembrano *prima facie* attribuiti ad una persona fisica ma, invece, ad un *software* caricato su di una macchina, ossia ad un bene immateriale che non costituisce nemmeno un autonomo centro di imputazione di rapporti giuridici. Vi è, infatti, un terreno ancora nebbioso del diritto entro i cui confini ad alcuni sembra possibile argomentare che, dal momento che è il *socialbot* a portare avanti in modo autonomo le interazioni sulle piattaforme *social* (quali pubblicare contenuti o apporre "reazioni" ai contenuti di altri), le qualità «a cui la legge attribuisce effetti giuridici» debbano essere riferite ad esso.

---

[37] Cass. pen., sez. V, 29 aprile 2013, n. 18826, 16 giugno 2014, n. 25774 (in cui gli Ermellini hanno rilevato come il profilo creato fosse corredato da «una descrizione tutt'altro che lusinghiera»), e 08 giugno 2018, n. 33862.

[38] Si vedano, ad esempio, Cass. pen., sez. II, 17 maggio 2019, n. 21705, e 02 luglio 2020, n. 23760.

[39] Cass. pen., sez. V, 16 giugno 2014, n. 25774.

[40] Cass. pen., sez. V, 14 dicembre 2007, n. 46674; Cass. pen., sez. V, 16 giugno 2014, n. 2577.

[41] Cass. pen., sez. V, 19 marzo 1985, n. 2542; Cass. pen., sez. V, 16 marzo 2015, n. 11087, e *a contrariis* Cass. pen., sez. V, 18 dicembre 2020, n. 5432.

[42] Cass. pen., sez. V, 16 giugno 2014, n. 25774, che riprende quanto statuito precedentemente in Cass. pen., sez. V, 14 dicembre 2007, n. 46674.

Una prima avventata – per non dire errata – interpretazione di questa situazione potrebbe essere di ritenere che l'utilizzo di un *socialbot* impedisca l'integrazione della fattispecie criminosa, in quanto identità o caratteri verrebbero attribuiti ad un oggetto e non «a sé o ad altri». Un'altra interpretazione – parimenti incauta – potrebbe essere quella di confondere le capacità di azione autonoma della macchina con il possesso di autonomia decisionale da parte di essa e, quindi, di reclamare una qualche forma di imputabilità nei suoi confronti.

Davanti al rischio di attribuire erroneamente delle effettive capacità decisionali – ed una possibile conseguente responsabilità penale – a quelli che, in fin dei conti, sono solo degli elaboratori dati che si limitano a seguire un elenco piuttosto flessibile di comandi preimpostanti (ossia, ad eseguire un codice sorgente particolarmente complesso), si deve dare risalto al fatto che essi si limitano ad essere degli strumenti – sebbene particolarmente sofisticati – che svolgono solamente le azioni che un programmatore ha inserito nel loro codice sorgente o i comandi che il loro amministratore gli dà. Lo stesso concetto di autonomia di azione di questi agenti *software* poggia, infatti, su di una loro programmazione che li abbia dotati di un ampio bagaglio di parametri che gli permettono di identificare numerosi ambienti informatici e ad interagire con essi.[43] La loro indipendenza di azione non deve, però, essere confusa con libertà di giudizio o di iniziativa: una volta attivati sono, sì, in grado di portare a termine il compito affidatogli senza bisogno di assistenza ma non sono capaci di decidere da soli il loro scopo.

In questa sede è anche importante evidenziare come il creatore di un *socialbot* non deve necessariamente rispondere di tutte le azioni compiute per mezzo del suo strumento *software*. Infatti, sebbene sia lui a decidere quali siano le interazioni che il suo prodotto debba poter compiere, è ben possibile che non sia lui ad indicargli quali condizioni debbano fargli intraprendere le sue attività. È solo colui che impiega attivamente il *socialbot* a dover rispondere degli eventi realizzati per il tramite di questo agente *software*.[44] Questo perché è solo l'utilizzatore effettivo del *socialbot* ad attivarlo e ad impostare gli elementi che, se rilevati dal programma, lo faranno agire all'interno del *social network*.[45]

---

[43] Un *socialbot* correttamente programmato contiene all'interno del suo codice sorgente tutti i parametri per identificare correttamente gli elementi contenuti nel sito internet di un *social medium* e le indicazioni di come interagire con essi. Quindi, quando agisce all'interno di un *social network*, un *socialbot* si limita ad interagire con gli elementi che è in grado di riconoscere grazie alle librerie di cui è dotato (ad esempio, può essere in grado di cliccare il pulsante "Mi piace" di un *post* su Facebook perché riconosce gli elementi della piattaforma) mentre, quando incontra un elemento che non riconosce o non trova l'elemento che sta cercando, esso potrebbe saltare l'operazione che avrebbe dovuto compiere. Il grado di adattabilità e di autonomia d'azione di un *socialbot* dipende, quindi, dalla qualità della sua programmazione.

[44] Al programmatore che abbia creato il *socialbot* e lo abbia fornito al suo utilizzatore potrebbe essere sempre contestata la partecipazione alla realizzazione di un reato nella posizione di ausiliatore, qualora fosse provata la sua volontà di apportare un aiuto materiale alla realizzazione del conosciuto progetto delittuoso. In merito alla figura del cosiddetto "complice ausiliatore" si veda quanto riportato in G. Fiandaca - M. Musco, *Diritto penale. Parte generale*, Bologna, 2009, 504, e le recenti precisazioni in Cass. pen., sez. V, 9 novembre 2021, n. 8973.

[45] Si immagini un *socialbot* programmato per agire su Facebook e programmato per mettere un "Mi piace" solamente a quei *post* di altri utenti che contengano al proprio interno una o più parole-chiave impostate dal suo amministratore. In questo esempio il programmatore ha solamente impostato la funzionalità mentre è l'amministratore ad indicare al *software* quale sia la parola-chiave che, se rilevata, gli fa avviare il processo di apposizione del "Mi piace". Differenti copie del medesimo *socialbot* ben potrebbero essere usate da persone diverse per promuovere messaggi di tolleranza oppure di discriminazione senza che il

Ciò non toglie, comunque, che molto spesso il creatore di un *socialbot* sia anche il suo utilizzatore finale ma è importante sottolineare come non sia necessariamente sempre così.

Insomma, l'attribuzione di qualità rilevanti per la nostra legge penale non deve immaginarsi avvenire in capo ad un bene immateriale (del quale alcuni potrebbero erroneamente arrivare a congetturare una qualche forma di personalità giuridica). Essa, invece, deve essere compiuta in capo all'utilizzatore del *socialbot*, con quest'ultimo che funge da strumento particolarmente raffinato e complesso per l'induzione in errore dei destinatari delle interazioni sulle piattaforme *social*. In altre parole, l'impiego di questo agente *software* costituisce una forma particolarmente intricata di attribuzione a sé di determinate qualità, dal momento che esso è semplicemente uno strumento sottoposto in ogni momento al controllo del suo amministratore (il quale può intervenire a suo piacimento per direzionarlo o fermarlo).

Le più recenti innovazioni in campo informatico creano, poi, un ulteriore nodo che, allo stato attuale, è ammantato da pesanti incertezze interpretative. L'attuale sviluppo dei *software* di intelligenza artificiale (IA) è ormai giunto al punto in cui essi sono in grado di produrre in autonomia testi di senso compiuto. Ben poco impedisce oggi ad un *socialbot* di essere progettato in modo da interagire su di un *social medium* e da produrre un testo – o anche di una immagine – grazie a forme di IA.[46] In altre parole, un *socialbot* ben potrebbe essere progettato in modo da integrare funzionalità di intelligenza artificiale di tipo generativo che gli permettono di produrre contenuti e di pubblicarli poi su di una piattaforma *social* per rendere ancora di più l'impressione di essere una persona vera.

In merito al particolare – ma plausibile – scenario di un *socialbot* capace di generare in autonomia dei contenuti artificiali è importante evidenziare come le legislazioni vigenti siano in uno stato di grave arretratezza ma, allo stesso tempo, come sia possibile individuare alcuni appigli per una loro interpretazione che permetta di dare risposta a qualsiasi richiesta di giustizia. Sebbene, infatti, il tema della produzione da parte delle IA di testi ed immagini stia generando un acceso dibattito, è comunque possibile individuare un orientamento sempre più forte – seguito anche all'interno dei dibattiti degli organi normativi dell'Unione europea[47] – che attribuisce la responsabilità per i danni cagionati a coloro che avevano i privilegi di amministratore di questi *software* e che non hanno supervisionato con sufficiente diligenza i loro processi di produzione di *output*.[48]

loro creatore abbia modo di intervenire.

[46] Già nel 2016 la Microsoft aveva creato una IA in grado sia di produrre testi che interagire con le persone su diversi *social media* (ma specificando che si trattava di una IA). Il risultato dell'esperimento di interazione di una IA con gli utenti di più piattaforme *social* aveva dato esiti disastrosi e l'impresa di Redmond aveva sospeso gli *account* della sua creatura digitale. In merito a questa vicenda si veda A. Tedeschi Toschi - G. Berni Ferretti, *La responsabilità per la diffamazione compiuta da un'Intelligenza artificiale*, cit., 173 ss.

[47] Si vedano, in particolare, la proposta di regolamento COM(2021) 206 final, che stabilisce regole armonizzate sulla intelligenza artificiale (legge sull'intelligenza artificiale) e la proposta di direttiva COM(2022) 496 final, relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale.

[48] Su questo tema si veda, ad esempio, A. Tedeschi Toschi - G. Berni Ferretti, *La responsabilità per la diffamazione compiuta da un'Intelligenza artificiale*, cit., 173 ss.

Quindi, anche nel caso di un *socialbot* con caratteri di IA, sarà il suo amministratore a dover rispondere degli eventi causati da questo (oltre che dell'illecita attribuzione di qualità a cui la legge attribuisce rilevanza).

Quindi, dalle considerazioni appena esposte in merito alla natura strumentale dei *socialbot* si può facilmente ricavare che l'azione di sostituire «illegittimamente la propria all'altrui persona» viene compiuta, nella realizzazione del reato di cui all'art. 494 c.p., dal soggetto che possiede i privilegi di amministratore di questi agenti *software* autonomi. Costui, infatti, si limita ad utilizzare un sofisticato strumento digitale (il *socialbot*) per far apparire sé stesso come un'altra persona su di una piattaforma di *social networking*.

Chiarito come sia l'amministratore del *socialbot* l'effettivo perpetratore del reato, conviene analizzare ulteriormente come tale condotta di mascheramento della propria identità possa essere realizzata. Infatti, all'interno del campo della giurisprudenza relativa all'art. 494 c.p. la valutazione dell'illegittima sostituzione della propria all'altrui persona su internet è giunta ormai ad affrontare numerose e disparate declinazioni. In particolare, molte delle recenti decisioni della Corte di cassazione hanno avuto ad oggetto la creazione di profili sui principali *social media* corredati di generalità di altre persone e di fotografie "trafugate" che le ritraevano. Questi profili sono stati usati per contattare, interagire e stringere relazioni con altri utenti ed indurli a credere che il loro interlocutore avesse le caratteristiche che venivano loro mostrato tramite il profilo così costruito e tramite i messaggi ed altre forme di interazione (come l'invio di fotografie o l'apposizione di reazioni ai loro contenuti).

Tra le varie declinazioni della sostituzione della propria all'altrui persona che sono giunte all'attenzione dei giudici della Suprema Corte pare qui importante riportare come questi siano stati chiamati a valutare anche il caso in cui ad essere "trafugata" era stata solo la fotografia di un'altra persona mentre le sue generalità non venivano impiegate (tale condotta è stata, comunque, ritenuta integrante il delitto descritto dal 494 c.p.).[49] Addirittura, è stato portato davanti alla Corte di cassazione il caso, poi sanzionato come illecito, dell'utilizzo non di una fotografia ma di un disegno caricaturale di un'altra persona.[50]

La particolare attenzione riservata dall'Autorità giurisdizionale per l'illecito impiego di elementi grafici che ritraggono altre persone (spesso inconsapevoli) può trovare una valida argomentazione nel fatto che detto utilizzo di raffigurazioni tra le condotte realizzate in violazione dell'art. 494 c.p. arreca un danno non solo al destinatario delle interazioni *online* compiute dal reo ma anche alla persona effettivamente ritratta nell'immagine. La perpetrazione del delitto in questa forma, insomma, comporta una lesione non solo alla fede pubblica[51] ma anche del diritto all'identità e al riconoscimento del giusto credito sociale del "derubato" dell'immagine (il quale facilmente può vedersi, così, attribuiti fatti e parole non propri con un conseguente potenziale detrimento della

[49] Si veda, ad esempio, Cass. pen., sez. V, 16 giugno 2014, n. 25774, in cui gli Ermellini hanno ritenuto che «integra il delitto di sostituzione di persona (art. 494 c.p.) la condotta di colui che crea ed utilizza un "profilo" su *social network*, utilizzando abusivamente l'immagine di una persona del tutto inconsapevole, associata ad un "*nickname*" di fantasia ed a caratteristiche personali negative».

[50] Cass. pen., sez. V, 23 luglio 2020, n. 22049.

[51] Bene giuridico espressamente e direttamente tutelato dalla norma, come specificato in Cass. pen., sez. II, 11 settembre 2020, n. 26589.

sua considerazione da parte di altri). Tuttavia, c'è da notare che le attuali capacità delle tecnologie digitali permettono di superare la causazione del duplice danno appena descritto. Infatti, come detto, le più recenti tecnologie digitali permettono la produzione in massa e quasi immediata di ritratti *AI-generated* realistici e sufficientemente credibili. Così, in combinazione con l'utilizzo di *nickname* privi di allusioni o riferimenti ad altre persone, è possibile creare profili *social* che non rientrano nell'ambito delle summenzionate interpretazioni della norma offerte dalla Suprema Corte.

La creazione di un profilo secondo gli accorgimenti appena descritti (ritratto generato da una IA e *nickname* di fantasia) eviterebbe la lesione dell'immagine o della dignità di una persona (dal momento che non vi è alcun soggetto esistente che sia effettivamente titolare delle generalità spese) ed eviterebbe, quindi, interventi delle autorità a tutela del diritto all'identità personale. Ne deriva, così, il rischio che la gravità delle condotte poste in essere da chi mascheri la gestione di un *account* con caratteri falsificati affidato ad un *socialbot* sia percepita come estremamente inferiore. In fondo, quando non esiste una persona che sia stata "derubata" della propria identità e che rischi di vedersi ingiustamente attribuire fatti o dichiarazioni, viene meno uno dei soggetti che possono essere offesi dal perpetratore del reato.[52]

(B) Induzione in errore – A solido bastione contro le preoccupazioni appena esposte vi è il fatto che, ai fini dell'applicazione dell'art. 494 c.p., tale norma non considera fondamentale l'illegittimo impiego dell'identità di una persona realmente esistente. La realizzazione del reato viene riconosciuta anche quando avviene attraverso l'attribuzione di generalità, stati o qualità che abbiano rilevanza giuridica che non siano corrispondenti al vero. Questo perché il momento consumativo del reato è stato individuato nell'altrui induzione in errore e perché «oggetto della tutela penale è l'interesse riguardante la pubblica fede, in quanto questa può essere sorpresa da inganni relativi alla vera essenza di una persona o alla sua identità o ai suoi attributi sociali; siccome si tratta di inganni che possono superare la ristretta cerchia d'un determinato destinatario».[53]

Giova ribadire che, se il momento consumativo del reato viene identificato nell'induzione in errore circa l'essenza di chi – o di cosa – si nasconde realmente dietro la maschera del profilo *social*, deve essere tenuto ben presente che gli elementi che possono condurre a tale sbaglio sono sempre direttamente o indirettamente il risultato dell'operato di una o più persone che agiscono per il tramite di un *socialbot* (anche qualora tale azione avvenga con l'ausilio di un *botmaster*), alle quali saranno imputabili gli effetti generati da tali elementi. In particolare, in questa ipotesi il disegno criminoso volto alla realizzazione della sostituzione di persona su di un *social medium* vede, da una parte, la combinazione delle attività di creazione di un profilo che mostri un ritratto (anche generato da un'intelligenza artificiale), delle generalità e delle informazioni che siano diverse da quelle di colui che ne ha le credenziali di accesso e, dall'altro, delle attività di interazione, quali inviare messaggi, apporre reazioni ai contenuti di terzi e pubblicarne

[52] Riguardo all'importanza attribuita alla persona le cui generalità sono state illegittimamente spese online si veda, ad esempio, Cass. pen., sez. V, 16 giugno 2014, n. 25774, in cui viene fatto riferimento, *ex plurimis*, alle sentenze 29 aprile 2013, n. 18826, 27 marzo 2009, n. 21574, 9 dicembre 2008, n. 7187, e 25 ottobre 2007, n. 237855.

[53] Cass. pen., sez. V, 16 giugno 2014, n. 25774, che riprende quanto statuito in Cass. pen., sez. V, 14 dicembre 2007, n. 46674.

di propri, compiute dal *socialbot* su indicazione del proprio amministratore. È quest'ultimo soggetto a fornire all'agente *software* la chiave di accesso (le credenziali) da utilizzare per accedere all'*account* creato e ad indicargli quali attività svolgere sulla piattaforma o con quali destinatari interagire.

Dall'inquadramento appena compiuto dei complessi mezzi fraudolenti utilizzabili per nascondere la vera identità di colui che gestisce – anche indirettamente tramite l'uso di un *socialbot* – un profilo su di un *social medium* nasce, però, un ulteriore problema: quando si possa parlare di effettiva induzione in errore di soggetti terzi non destinatari delle interazioni degli agenti *software* autonomi. Se, infatti, un'interazione continuata nel tempo tra un *socialbot* (che, si ribadisce, è sempre sotto il controllo di chiunque ne abbia i privilegi di amministratore) e una persona, che induca quest'ultima a credere di star intrattenendo un qualche tipo di relazione (amicale o addirittura sentimentale),[54] può facilmente rientrare nell'alveo dell'induzione in errore, altre situazioni sono meno chiare.

La prima situazione che può sollevare delle problematiche interpretative è legata alla strutturazione stessa delle piattaforme di *social networking* e alle modalità con cui esse diffondono i contenuti degli utenti, se questi sono oggetto di apposizione di reazioni da parte di altri. Infatti, come ben illustrato dalla Corte di cassazione, gli algoritmi delle piattaforme *social* sono – nella maggior parte dei casi – volti a mostrare agli utenti un «continuo aggiornamento delle notizie e delle attività sviluppate dai [loro] contatti» e tale modalità di informazione è «condizionata dal maggior numero di interazioni che riceve ogni singolo messaggio».[55] I sistemi dei *social media*, quindi, si limitano a mostrare agli utenti terzi rispetto a questa prima interazione i contenuti dell'autore originale corredati dall'indicazione dell'avvenuta apposizione di reazioni ad essi. Il problema interpretativo che sorge dalla situazione appena descritta è che chi ha i privilegi di amministratore di un *socialbot* e le credenziali di accesso a dei profili *social* risulta in grado di indurre in errore degli utenti terzi rispetto all'interazione iniziale senza interagire direttamente con loro. Vi è addirittura chi potrebbe provare a ipotizzare che questi terzi soggetti, non essendo mai stati i destinatari della interazione iniziale, siano stati, al massimo, colposamente indotti in errore.[56]

Questa prima situazione di incertezza può essere facilmente superata considerando sotto una luce leggermente diversa la medesima metodologia di diffusione di contenuti

---

[54] Grazie all'implementazione di forme di IA questa situazione non è affatto un'ipotesi ma una situazione reale e ormai ben documentata. Si vedano a questo riguardo, ad esempio, J. Zitser, *A tech bro used AI to create a virtual version of his girlfriend*, in *Insider.com*, 1 giugno 2023, e A. Sternlicht, *A 23-year-old Snapchat influencer used OpenAI's technology to create an A.I. version of herself that will be your girlfriend for $1 per minute*, in *Fortune.com*, 9 maggio 2023.

[55] Cass. pen., sez. I, 9 febbraio 2022, n. 4534.

[56] Si pensi, ad esempio, ad una situazione in cui Tizio abbia il controllo di un *socialbot* e di numerosi profili falsi su Facebook e disponga che il suo agente *software* metta dei "Mi piace" con tutti i profili che gestisce ad un *post* pubblicato da Caio, il quale è completamente ignaro dei sotterfugi di Tizio. Dato che gli algoritmi della piattaforma *social* sono progettati in modo da "rilanciare" i contenuti popolari, il *post* di Caio – che ha ricevuto diversi "Mi piace" dai profili automatizzati di Tizio – verrebbe mostrato dalla piattaforma nel *newsfeed* di Sempronio, anche lui totalmente ignaro della natura sintetica dei profili che ne hanno artificialmente accresciuto la popolarità. In questo caso Tizio – attraverso il suo *socialbot* – non ha mai interagito direttamente con Sempronio ma quest'ultimo ha, comunque, avuto l'impressione che un alto numero di persone abbia espresso il proprio apprezzamento per il *post* di Caio.

che viene adottata dai gestori delle piattaforme di *social networking*. È lo stesso principio di aggiornamento continuo ed automatico del *newsfeed* mostrato agli utenti, ossia quello di mostrare loro i contenuti che ricevono il maggior numero di interazioni, a fungere da base di appoggio per un utilizzo efficace dei mezzi fraudolenti per la realizzazione della sostituzione di persona nei *social network*. Data, quindi, la basilare strutturazione dei *social media* secondo questo principio, l'apposizione di una reazione ha una intrinseca funzione propalatrice, come già riconosciuto dalla Corte di cassazione,[57] nei confronti di soggetti terzi che, di conseguenza, sono destinatari voluti e ricercati di tale interazione. In altre parole, con la promozione artificiale dei contenuti del destinatario delle interazioni dei *socialbot* – che viene compiuta per il tramite di uno o più profili *social* caratterizzati da qualità (di altri o inventate) rilevanti agli occhi della legge – il loro amministratore sfrutta consapevolmente la strutturazione delle piattaforme per raggiungere il maggior numero possibile di utenti reali e indurli in errore circa la reale identità dell'utilizzatore di ciascuno dei profili che ha automatizzato. Il fine ultimo di questa operazione può essere molto vario e può consistere sia in un vantaggio per l'amministratore del *socialbot* che per altre persone come in un danno per altri ma la finalità intermedia di essa, almeno in questo scenario, rimane sempre quella di dare a utenti terzi una falsa impressione di popolarità dei messaggi veicolati dai contenuti oggetto di interazioni dirette.

Insomma, l'utilizzo dei *socialbot* per dare false impressioni di popolarità si traduce, nei fatti, nell'induzione in errore di un alto numero dei destinatari degli aggiornamenti automatici del loro *newsfeed* circa le reali generalità dell'utilizzatore di numerosi profili attivi sulla piattaforma di *social networking*. L'effetto è, così, quello di ledere la pubblica fede in merito all'effettiva identità degli utenti dei *social media*. In aggiunta, tale impiego viene fatto dall'amministratore di questi agenti *software* autonomi con la finalità di procurare a sé o ad altri un vantaggio (come nel caso della promozione di una determinata ideologia) o di recare ad altri un danno (come nell'ipotesi della diffamazione aggravata). Pare, quindi, difficile negare che venga così integrato il reato di sostituzione di persona attraverso l'impiego di *socialbot*.

Ancora più complicata dal punto di vista interpretativo è la situazione in cui non vi siano delle "reazioni" apposte ai contenuti di un utente (come i "Mi piace" di Facebook o i "Consiglia" di LinkedIn) ma solo delle condotte più "passive", come il numero di visualizzazioni ad un contenuto video o il numero di ascolti di un contenuto audio.[58] Tale scenario, già egregiamente rilevato da Weissmann,[59] presenta rispetto a quello precedente l'ancor più complessa problematica del fatto che la maggior parte dei sistemi di conteggio delle visualizzazioni o degli ascolti di contenuti multimediali non mostra l'identità di coloro che li hanno guardati o ascoltati, non dando così alcuna indicazione

[57] Cass. pen., sez. V, 12 dicembre 2017, n. 55418. La decisione presa dalla Corte era relativa alla questione se la pubblicazione sul proprio profilo Facebook di video inneggianti allo Stato Islamico e l'apposizione di "Mi piace" ad altri integrassero o meno il delitto di istigazione a delinquere previsto dall'art. 414 c.p.

[58] Si vedano la descrizione e le considerazioni in merito a questa forma di interazione fatte *supra* alla lettera (C) del paragrafo 1.1. di questo articolo.

[59] S. Weissmann, *How Not to Regulate Social Media*, cit., 58 ss., dove l'autrice fa particolare riferimento all'uso di *socialbot* fatto da agenti della disinformazione russa per incrementare il numero di visualizzazioni di contenuti video e indurre società di pubblicità a spendere milioni di dollari in annunci ad essi collegati.

ad altri utenti di chi abbia interagito con essi. Per di più, queste azioni non riguardano condotte immediatamente riconducibili a forme di consenso e di adesione al contenuto ma, invece, afferiscono solamente al livello di fama raggiunto da esso (la quale ben potrebbe essere caratterizzata da una comune disapprovazione). Questo scenario sembra posizionare l'uso di profili falsi o di *socialbot* per incrementare artificialmente il conteggio del consumo "passivo" di contenuti multimediali al di fuori dei confini del reato di sostituzione di persona. Questo perché, data la mancata spendita di alcuna forma di generalità – di altri od inventate – verso delle altre persone, non sembra possibile argomentare che queste ultime siano state indotte in errore tramite l'utilizzo illegittimo di qualità a cui la legge attribuisce effetti giuridici.

Vi è comunque da notare come, quando l'incremento artificiale degli indici di popolarità di contenuti sui *social media* sia volto ad ingannare qualcuno per ottenere un guadagno e causi al contempo un danno economico alla vittima, sarà sempre ipotizzabile una forma di tutela di quest'ultima tramite previsioni diverse da quelle di tutela della fede pubblica (come l'art. 640 c.p. relativo alla truffa).[60] Tuttavia, quando gli artifizi ed i raggiri consistano nella creazione di false identità online e siano finalizzati al soddisfacimento di una propria vanità[61] o a rendere false impressioni di popolarità di opinioni o idee non necessariamente estremiste o discriminatorie (finalità che poco si discostano da forme di vanità), ripiegare sull'applicazione del solo reato di frode informatica (previsto dall'art. 640-*ter* c.p.) non sembra – almeno agli occhi di chi scrive – fornire la giusta tutela ai beni giuridici aggredibili da tali condotte. Infatti, in questo secondo caso si ha, comunque, la creazione di una o più false identità digitali che vengono utilizzate per dare ad un vasto pubblico l'impressione che venga prestata ampia attenzione alle posizioni veicolate dai contenuti pubblicati *online*, ossia per indurre in errore la fede pubblica circa l'effettivo numero di persone che abbiano guardato od ascoltato dei contenuti multimediali. Insomma, anche in assenza della spendita di un'identità fittizia o di terzi, l'utilizzo di *socialbot* per incrementare il numero di reazioni "passive", come le visualizzazioni di contenuti multimediali, finisce per indurre in errore la fede pubblica in merito ad un dato: l'esistenza di un ampio numero di persone interessate a detti contenuti.

L'insistenza sulla rilevanza di condotte che possono inizialmente sembrare secondarie dipende dal fatto che, come mostrato da alcuni importanti studi,[62] la formazione delle

[60] Si pensi, ad esempio, ad un *influencer* che si "compri" dei *follower* fittizi. Ossia, un creatore di contenuti video che utilizzi dei *socialbot* per incrementare artificialmente il numero di visualizzazioni dei suoi prodotti, dando in questo modo l'impressione di essere più conosciuto e popolare tra le persone di quanto non sia in realtà e riuscendo, grazie a questi artifizi, ad ottenere delle lucrose sponsorizzazioni da parte di imprese che altrimenti non avrebbero mai sottoscritto un contratto con lui. In questo caso l'utilizzo di agenti *software* autonomi è finalizzato a procurarsi un ingiusto profitto con contestuale danno economico per l'impresa sponsorizzatrice, finendo col configurare almeno una truffa.

[61] In merito alla rilevanza penale del perseguimento di vantaggi non patrimoniali nel reato di sostituzione di persona si veda, ad esempio, Cass. pen., sez. V, 16 giugno 2014, n. 25774.

[62] P.F. Lazarsfeld - B. Berelson - H. Gaudet, *The People's Choice. How the Voter Makes Up His Mind in a Presidential Campaign*, New York, 1944; B. Berelson - P.F. Lazarsfeld - W.N. McPhee, *Voting: a study of opinion formation in a presidential campaign*, Chicago, 1954; E. Katz - P.F. Lazarsfeld, *Personal influence: The part played by people in the flow of mass communications*, Glencoe, 1955; M.E.J. Newman, *Networks: An Introduction*, New York, 2010; G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, Oxford, 2010, 211 ss.

opinioni degli individui è influenzata dalle reti sociali di cui fanno parte e costoro tendono a fidarsi delle informazioni che circolano al loro interno[63] e a considerare come veritiere le opinioni che sono ampiamente diffuse.[64] Alla luce di queste considerazioni non si può, quindi, giudicare come irrilevante la malevola diffusione potenziata di contenuti sui *social media*, dal momento che essa è in grado di condizionare il comportamento o la psicologia di un vasto pubblico. In altre parole, la diffusione artificialmente aumentata di contenuti è una condotta che, dati gli innati meccanismi di inconscia e tendenziale adesione ai messaggi ampiamente diffusi nel proprio contesto sociale, già di per sé contiene quel *quid pluris* che la rende una forma di divulgazione di opinioni finalizzata ad influenzare il comportamento o la psicologia di un vasto pubblico e a raccogliere adesioni.[65]

Insomma, l'uso di *socialbot* per aumentare la portata della diffusione di contenuti online, sfruttando la conformazione degli algoritmi dei *newsfeed* delle piattaforme *social*, deve essere visto come una forma di induzione in errore di un vasto numero di persone finalizzata a procurare dei vantaggi – anche di natura non patrimoniale – che ha l'effetto di trarre in inganno la fede pubblica su aspetti che, come mostrato da autorevoli e consolidati studi, condizionano la realtà sociale.

Quindi, sebbene l'effetto di tale utilizzo dei *socialbot* possa risolversi in un riorientamento marginale delle posizioni delle persone da esso investite, non può essere ignorato come questo malevolo impiego sia caratterizzato da una capacità di influenza non nulla sui molteplici destinatari – anche indiretti – delle interazioni con i contenuti pubblicati sulle piattaforme di *social networking* e come tale ascendente sia raggiunto tramite l'induzione in errore della fede pubblica sull'esistenza di un ampio numero di persone.

(C) Particolare tenuità del fatto – Non può essere tralasciato il fatto che, a minaccia della tesi della necessità di un'applicazione estensiva delle tutele penali contro gli atti di induzione in errore della fede pubblica appena illustrata, la Corte di Cassazione abbia già in passato ritenuto che alcune concrete condotte di sostituzione di persona possano non richiedere una sanzione. In particolare, in una loro recente sentenza, i giudici della Suprema corte hanno statuito che l'aver utilizzato per la creazione di un profilo online le generalità di un'altra persona del tutto inconsapevole possa non essere punito per ef-

[63] Y. Jun - R. Meng - G.V. Johar, *Perceived social presence reduces fact-checking*, in *Proceedings of the National Academy of Sciences*, 2017, 5976 ss.; K.C. Yang - O. Varol - C.A. Davis - E. Ferrara - A. Flammini - F. Menczer, *Arming the public with artificial intelligence to counter social bots*, in *Human Behavior and Emerging Technologies*, 2019, 48 ss.; H. Wolters - K. Stricklin - N. Carey - M.K. McBride, *The Psychology of (Dis) information: A Primer on Key Psychological Mechanisms*, in *CNA Research Memorandum*, 2021, 1 ss.

[64] M. Luo - J.T. Hancock - D.M. Markowitz, *Credibility perceptions and detection accuracy of fake news headlines on social media: Effects of truth-bias and endorsement cues*, in *Communication Research*, 2022, 171 ss., in cui gli autori hanno messo in luce come su Facebook contenuti con un elevato numero di "Mi piace" vengano percepiti come maggiormente credibili.

[65] In merito alla natura della "propaganda di idee" rilevante agli occhi della legge si vedano Cass. pen., sez. I, 9 febbraio 2022, n. 4534 e Cass. pen., sez. V, 22 luglio 2019, n. 32862. Invero, le sentenze citate sono relative al reato di propaganda e istigazione a delinquere per motivi di discriminazione razziale, etnica e religiosa (descritto all'art. 604-*bis* c.p.) e – si riconosce – lo spettro di un'applicazione analogica sembra aleggiare sopra al ragionamento qui sviluppato. Tuttavia, come visto, la condotta dell'illegittimo utilizzo di *socialbot* possiede un'alta intensità offensiva, dato che non si limita ad indurre in errore degli utenti di un *social medium* circa il numero – e, quindi, le identità – di persone che diffondono contenuti sulla piattaforma ma è anche in grado di influenzare il comportamento o la psicologia di questi.

fetto delle previsioni dell'art. 131-*bis* del codice penale, ossia quando le peculiarità della fattispecie concreta – le modalità della condotta, il grado di colpevolezza da esse desumibile e l'entità del danno o del pericolo – ammantino il fatto di particolare tenuità.[66]
Addirittura, gli Ermellini hanno ritenuto che il grado di offensività delle condotte valutate fosse particolarmente tenue, nonostante al cosiddetto "furto di identità" fossero seguite ulteriori condotte illecite (nel caso di specie delle molestie integranti il reato di cui all'art. 660 c.p.). In altre parole, essi hanno ritenuto che, nonostante la plurima offensività della condotta tenuta in concreto, l'aggressione ai beni giuridici compiuta non fosse caratterizzata da un grado di dannosità rilevante. In aggiunta, i giudici di piazza Cavour hanno stabilito che le osservazioni del Procuratore Generale presso la Corte di appello di Milano circa la concreta portata offensiva del delitto di sostituzione di persona facessero leva su «un approccio astrattizzante, non in linea con il necessario ancoraggio alla fattispecie concreta che connota la causa di non punibilità» e ne hanno così, appunto, escluso la punibilità.[67]
Quindi, non sembra neanche irrealistico ipotizzare che la Suprema Corte possa giungere alle medesime conclusioni per un caso simile a quello descritto da Cresci, Di Pietro, Petrocchi, Spognardi e Tesconi,[68] ossia quando una persona denunci di essere stata vittima di un furto di identità sui *social* consistente nell'utilizzo di un suo ritratto per la creazione di un profilo avente altre generalità (che non ne permettano la sua identificazione) e che questo venga impiegato solamente per ri-condividere su una piattaforma *social* i contenuti pubblicati da un candidato politico. Infatti, in quest'ultima ipotesi all'indebito utilizzo di una immagine della vittima non farebbe seguito alcun altro illecito comportamento diretto nei suoi confronti (come avvenuto, invece, nel caso del 2020 ritenuto di particolare tenuità dai giudici di Cassazione).
non pare irrealistico immaginare che la medesima considerazione del valore particolarmente tenue del crimine possa essere applicata anche ad attività di propaganda o di cosiddetto *astroturfing* potenziate dall'uso di *socialbot*.[69] Dopotutto, la creazione di profili

[66] Cass. pen., sez. V, 10 gennaio 2020, n. 652, in cui i membri del collegio giudicante hanno ritenuto che il carattere del tutto isolato dell'episodio di creazione di un profilo online utilizzando le generalità della vittima fosse causa di esclusione della punibilità per particolare tenuità del fatto.

[67] Pur riconoscendo che «l'avvalersi di una piattaforma multimediale accessibile ad un numero indeterminato di utilizzatori comporta un danno importante e astrattamente senza termini nel tempo e nello spazio».

[68] S. Cresci - R. Di Pietro - M. Petrocchi - A. Spognardi - M. Tesconi, *Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling*, cit., 561 ss.

[69] L'*astroturfing* è la pratica di nascondere l'identità del vero patrocinatore di un determinato messaggio tramite l'organizzazione di false campagne "nate dal basso" che diano l'impressione che vi sia una vasta pletora di persone – e non solo alcune imprese commerciali o dei ristretti gruppi di interesse politico od economico – a sostenerlo. Questa attività si è tradotta su internet e sui *social media* nella creazione di reti di profili che promuovono e condividono incessantemente contenuti relativi ad un determinato argomento ed orientati verso il sostegno di una precisa posizione al riguardo. In merito a questa pratica si vedano, J. Ratkiewicz - M. Conover - M. Meiss - B. Gonçalves - S. Patil - A. Flammini - F. Menczer, *Truthy: Mapping the spread of astroturf in microblog streams*, in *WWW '11: Proceedings of the 20th International Conference Companion on World Wide Web*, 2011, 249 ss.; J. Zhang - D. Carpenter - M. Ko, *Online astroturfing: A theoretical perspective*, in *Proceedings of the Nineteenth Americas Conference on Information Systems, Chicago, Illinois, August 15-17, 2013*, 2013, 1 ss.; S.C. Woolley, *Automating power: Social bot interference in global politics*, in *First Monday*, 21, 2016; M. Kovic - A. Rauchfleisch - M. Sele - C. Caspar, *Digital astroturfing in politics: Definition, typology, and countermeasures*, in *Studies in Communication Sciences*, 18(1), 2018, 69 ss.; E. Dubois

social tramite la completa invenzione delle generalità degli stessi e l'utilizzo di ritratti di persone inesistenti priverebbe la condotta di qualsiasi potenziale lesivo del diritto all'identità personale, all'immagine o alla dignità di qualcuno, diminuendo sensibilmente il senso di urgenza di un intervento di tutela giurisdizionale. Permarrebbero, comunque, il danno alla pubblica fede e il fine di «procurare a sé o ad altri un vantaggio» di natura non patrimoniale ma, in assenza di una chiara e condivisa illustrazione che espliciti come detto bene giuridico debba essere difeso anche da artificiosi incrementi della diffusione di contenuti e da inganni volti a dare impressioni di popolarità in realtà inesistente, non pare potersi escludere il rischio che strategie di propaganda volte a deviare il naturale percorso di formazione delle opinioni vengano erroneamente viste come prive di particolare gravità.

4. Conclusioni

La presenza dei *socialbot* all'interno dei *social network* è allo stato attuale una caratteristica inestirpabile di queste piattaforme. Sebbene da alcuni possa non essere vista come un problema rilevante per l'ordinamento di uno Stato, si fa sempre più forte la consapevolezza che questi agenti *software* rappresentano, invece, un rischio per la stabilità delle istituzioni democratiche di una nazione.[70] Ben facilmente, infatti, essi possono essere utilizzati per potenziare la diffusione di discorsi incitanti all'odio e alla discriminazione, di false informazioni su fatti, qualità personali ed eventi e per accrescere artificiosamente le impressioni di popolarità di certe persone od opinioni.[71]

In ambito europeo sono attualmente in corso delle azioni legislative volte a contrastare l'impiego dei *socialbot*, dato che essi possono essere impiegati per «la manipolazione intenzionale e spesso coordinata del servizio della piattaforma [*social*], con effetti prevedibili sulla salute pubblica, sul dibattito civico, sui processi elettorali, sulla sicurezza pubblica e sulla tutela dei minori».[72]

In attesa che anche in Italia vi sia una presa di coscienza della pericolosità delle pratiche di propaganda computazionale (tra le quali rientra l'impiego dei *socialbot*) e ci si renda conto che sono ormai necessari degli interventi normativi, dato che è già stato evidenziato come anche il nostro panorama digitale sia piagato da questo fenomeno,[73] ci si

- F. McKelvey, *Political Bots: Disrupting Canada's Democracy*, in *Canadian Journal of Communication*, 44(2), 2019, 27 ss.; F.B. Keller - D. Schoch - S. Stier - J. Yang, *Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign*, in *Political Communication*, 37(2), 2020, 256 ss.

[70] Si vedano a questo riguardo le preoccupazioni espresse dalle istituzioni europee, poi confluite nelle considerazioni del COM(2020) 825 final, *Proposta di regolamento del Parlamento europeo e del Consiglio relativo a un mercato unico dei servizi digitali (legge sui servizi digitali)*. In particolare, al Considerando 57 i *socialbot* vengono inseriti tra i «rischi sistemici» che minacciano la tenuta delle istituzioni dell'Unione e dei suoi Stati membri.

[71] Si vedano, ad esempio, le considerazioni contenute in COM(2018) 236 final, *comunicazione della Commissione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni, Contrastare la disinformazione online: un approccio europeo*.

[72] Considerando 57 della COM(2020) 825.

[73] A. Vogt, *Hot or bot? Italian professor casts doubt on politician's Twitter popularity*, in *The Guardian*, 22 luglio 2012; S. Cresci - R. Di Pietro - M. Petrocchi - A. Spognardi - M. Tesconi, *DNA-inspired online behavioural*

deve rivolgere alla legislazione già in vigore per tutelare quei beni giuridici che sono minacciati da questa tecnologia. In altra sede si è già avuto modo di evidenziare come l'impiego di *socialbot* per danneggiare la reputazione delle persone costituisca una forma aggravata del reato di diffamazione.[74] Qui, invece, si è voluto concentrarsi sull'applicazione della normativa esistente relativa al reato di sostituzione di persona per la tutela di un altro bene giuridico di estrema rilevanza che viene aggredito dal malevolo impiego dei *socialbot*: quello della pubblica fede.

Al fine di evitare fraintendimenti in merito all'individuazione del responsabile degli eventi causati da questa particolare categoria di agenti *software* autonomi, si è sottolineato come essi non siano intelligenti, capaci di intendere e volere e neppure realmente autonomi. Essi sono solamente dei sofisticati strumenti digitali sottoposti al controllo di colui che ne ha i privilegi di amministratore. Ne consegue che l'attribuzione di un'identità (di terzi o inventata), di qualità o di stati a cui la legge attribuisce effetti giuridici non avviene in capo a questi beni immateriali ma solo ai loro utilizzatori (i loro amministratori), che finiscono col commettere una sostituzione di persona.

Nell'analisi sull'integrazione di questo crimine tramite l'uso dei *socialbot* occupano una posizione centrale il bene giuridico tutelato, ossia la fede pubblica, e il suo momento consumativo, cioè l'induzione in errore di altre persone. Questo perché questi agenti *software*, quando dotati di numerosi profili *social*, non sono solo in grado di indurre in errore le persone sulla reale identità del loro utilizzatore ma possono addirittura condizionare il comportamento o la psicologia di un vasto pubblico, a causa della naturale tendenza inconscia delle persone ad aderire alle posizioni ampiamente diffuse nel proprio contesto sociale (anche in quelli online).[75] Sebbene gli effetti di campagne propagandistiche volte ad influenzare l'opinione delle persone tramite false impressioni di popolarità siano difficili da misurare e possano risultare marginali, si è visto come essi siano comunque in grado di generare risultati rilevanti per le istituzioni democratiche e la tenuta di interi sistemi sovranazionali.[76]

---

modelling and its application to spambot detection, in *IEEE Intelligent Systems*, 31(5), 2016, 58 ss.; Id., *The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race*, cit., 963 ss.; Id., *Exploiting digital DNA for the analysis of similarities in Twitter behaviours*, in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2017, 686 ss.; Id., *Social fingerprinting: detection of spambot groups through DNA-inspired behavioral modeling*, cit., 561 ss.

[74] A. Tedeschi Toschi - G. Berni Ferretti, *Social media, profili artificiali e tutela della reputazione*, cit., 107 ss.

[75] P.F. Lazarsfeld - B. Berelson - H. Gaudet, *The People's Choice. How the Voter Makes Up His Mind in a Presidential Campaign*, cit.; B. Berelson - P.F. Lazarsfeld - W.N. McPhee, *Voting: a study of opinion formation in a presidential campaign*, cit.; E. Katz - P.F. Lazarsfeld, *Personal influence: The part played by people in the flow of mass communications*, cit.; M.E.J. Newman, *Networks: An Introduction*, cit.; G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, cit.; Y. Jun - R. Meng - G.V. Johar, *Perceived social presence reduces fact-checking*, cit., 5976 ss.; K.C. Yang - O. Varol - C.A. Davis - E. Ferrara - A. Flammini - F. Menczer, *Arming the public with artificial intelligence to counter social bots*, cit., 48 ss.; H. Wolters - K. Stricklin - N. Carey - M.K. McBride, *The Psychology of (Dis)information: A Primer on Key Psychological Mechanisms*, cit., 1 ss.

[76] M. Hänska - S. Bauchowitz, *Tweeting for Brexit: how social media influenced the referendum*, in J. Mair - T. Clark - N. Fowler - R. Snoddy - R. Tait (a cura di), *Brexit, Trump and the Media*, Bury St Edmunds, 2017, 31 ss.; N. Persily, *The 2016 US Election: Can democracy survive the Internet?*", in *Journal of democracy*, 28(2), 2017, 63 ss.; J.A. Tucker - Y. Theocharis - M.E. Roberts - P. Barberá, *From liberation to turmoil: Social media and democracy*", in *Journal of democracy*, 28, 2017, 46 ss.; W. Hall - R. Tinati - W. Jennings, *From Brexit to Trump: Social media's role in democracy*, in *Computer*, 51(1), 2018, 18 ss.; S. Sanovich - D. Stukal - J.A. Tucker, *Turning the virtual tables: Government strategies for addressing online opposition with an application to Russia*, in *Comparative*

In ogni caso, l'impiego dei *socialbot* comporta la lesione della pubblica fede circa la reale identità dei soggetti attivi sui *social media* (i quali, usano questi agenti *software* come strumenti per automatizzare l'attribuzione a sé di altre identità) e, a seconda delle concrete modalità d'azione, finisce per aggredire altri beni giuridici meritevoli di tutela, quali l'identità personale e l'onore (nel caso che i profili *social* gestiti siano il frutto di un cosiddetto "furto d'identità") ovvero la libertà di ricevere informazioni o idee senza ingerenze e condizionamenti (nel caso di campagne di propaganda computazionale o di *astroturfing*). Quindi, nel contesto normativo vigente nel nostro paese si è individuato nella normativa di contrasto al reato di sostituzione di persona un valido strumento di tutela di questi beni giuridici, dal momento che «oggetto della tutela penale è l'interesse riguardante la pubblica fede […] siccome si tratta di inganni che possono superare la ristretta cerchia d'un determinato destinatario».[77]

Politics, 50(3), 2018, 435 ss.; Y. Gorodnichenko - T. Pham - O. Talavera, *Social media, sentiment and public opinions: Evidence from #Brexit and #USElection*, in *European Economic Review*, 136, 2021, 103772 ss.

[77] Cass. pen., sez. V, 16 giugno 2014, n. 25774, che riprende quanto statuito precedentemente in Cass. pen., sez. V, 14 dicembre 2007, n. 46674.