Il contrasto legislativo ai *socialbot* e le soluzioni avanzate nella Repubblica di Singapore e nella Repubblica d'Irlanda*

Alessandro Tedeschi Toschi - Giampaolo Berni Ferretti

Abstract

L'utilizzo dei *social media* ha permesso a chiunque di diffondere il proprio pensiero con una velocità ed una capillarità che eclissano quelle dei grandi organi di stampa tradizionali. All'interno di queste "piazze virtuali" operano anche i cosiddetti "*social bot*", dei programmi che, una volta forniti delle credenziali di accesso di un *account*, sono in grado di gestirlo in autonomia dando, però, l'impressione di essere una persona vera. La loro rapidità e precisione di reazione sui *social network* rendono questi "strumenti digitali" pericolosamente utili per la diffusione di contenuti favorevoli od ostili. Il recente impiego dei *socialbot* all'interno del dibattito elettorale e politico di diverse nazioni e le sue conseguenze hanno spinto i legislatori di diversi Stati ad intervenire contro questo fenomeno. Vengono qui prese in considerazione le riforme – e le proposte di riforma – avanzate nella Repubblica di Singapore e nella Repubblica d'Irlanda che specificamente mirano a contrastare l'uso di *social bot* con intenti malevoli o fuorvianti.

The use of social media has allowed anyone to spread their thoughts with a speed and a capillarity that eclipse those of the large traditional press companies. Within these "virtual squares" also operate the so-called "*social bot*," programs that, once provided with the login credentials of an account, can manage it independently, but giving the impression of being a real person. Their rapidity and precision of reaction on social networks make these "digital tools" dangerously useful for the dissemination of favourable or hostile content. The recent use of *socialbots* within the electoral and political debate of different nations and its consequences have prompted the legislators of various States to act against this phenomenon. We took into consideration here the reforms – and reform proposals – made in the Republic of Singapore and in the Republic of Ireland which specifically aim to counter the use of social bots with malicious or misleading intent.

* Su determinazione della direzione, il contributo è stato sottoposto a referaggio anonimo in conformità all'art. 15 del regolamento della Rivista

Sommario

Keywords

social network - automazione - *bot* - propaganda computazionale - disinformazione.

1. I *social bot* all'interno dei *social media*: diffusione e pericoli

La recente diffusione dell'utilizzo dei *social media* come strumenti di comunicazione e di divulgazione delle proprie opinioni ha portato a dei veri e propri sconvolgimenti nei modi e nei tempi di propagazione del pensiero. Grazie alle architetture digitali sviluppate negli ultimi anni e ospitate su infrastrutture capaci di gestire in autonomia e con perfetta precisione l'immensa mole di dati generati dalle azioni e dalle interazioni dei loro utenti (si calcola, ad esempio, che nel solo corso dell'agosto 2021 siano stati pubblicati 575.000 *tweet* ogni minuto)[1], la diffusione di contenuti e idee avviene oggi con una velocità ed una capillarità che eclissano quelle dei grandi organi di stampa tradizionali. Tuttavia, negli ultimi anni essi sono diventati anche un terreno fertile per la propaganda ed il proselitismo di tutti coloro che ne hanno compreso le potenzialità e hanno un minimo di cognizione del loro funzionamento[2].

Ad attrarre l'attenzione verso questi strumenti di comunicazione anche di soggetti pubblici e politici sono stati la facilità e la gratuità del loro utilizzo, il loro crescente impiego da parte delle persone come fonte di notizie ed il parallelo declino della carta stampata[3], nonché il conseguente avvicinamento dei *mass media* tradizionali a queste

[1] I dati statistici sono disponibili in S. Dixon, *Media usage in an online minute 2022*, in *statista.com*, 4 ottobre 2022.

[2] Si vedano, *ex multis*, S. Utz, *The (Potential) Benefits of Campaigning via Social Network Sites*, in *Journal of Computer-Mediated Communication*, 14(2), 2009, 221 ss.; S. Stieglitz – L. Dang-Xuan, *Social media and political communication: a social media analytics framework*, in *Social Network Analysis and Mining*, 3(4), 2013, 1277 ss.; G.S. Enli – E. Skogerbø, *Personalized Campaigns in Party-centred Politics Twitter and Facebook as Arenas for Political Communication*, in *Information, Communication & Society*, 16(5), 2013, 757 ss.; G.S. Enli, *Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election*, in *European Journal of Communication*, 32(1), 2017, 50 ss.; B. Mclaughlin – T. Macafee, *Becoming a Presidential Candidate: Social Media Following and Politician Identification*, in *Mass Communication and Society*, 22(5), 2019, 584 ss.

[3] Si vedano, *ex multis*, S.A. Myers – A. Sharma – P. Gupta – J. Lin, *Information network or social network? The structure of the Twitter follow graph*, in *Proceedings of the 23rd International Conference on World Wide Web*, 2014, 493 ss.; H. Allcott – M. Gentzkow, *Social media and fake news in the 2016 election*, in *Journal of economic perspectives*, 31(2), 2017, 211 ss.; K. Shu – H.R. Bernard – H. Liu, *Studying fake news via network analysis: detection and mitigation*, in N. Agarwal – N. Dokoohaki – S. Tokdemir (a cura di), *Emerging research challenges and opportunities in computational social network analysis and mining*, Berlino, 2019, 43 ss. Inoltre, anche il *Digital News Report* del 2016 redatto dal Reuters Institute in collaborazione con l'Università di Oxford ha sottolineato la crescente importanza dei *social media* come fonte di notizie. Il rapporto è liberamente

piattaforme. A tali ragioni, poi, si aggiunge il fatto che essi garantiscono una forma di interazione non mediata ed istantanea con il pubblico, la quale permette di evitare verifiche puntigliose dei fatti, contestualizzazioni e valutazioni sulla fondatezza delle rivendicazioni da parte di critici, esperti e professionisti. Per di più, l'architettura fisica di qualsiasi *social network* non è quella di una grande infrastruttura "ministeriale", dove migliaia di persone devono coordinarsi assiduamente tramite una minuziosa ed intricata burocrazia per far sì che anche un solo slogan possa giungere ad una vasta platea di ascoltatori. Essa consiste, invece, unicamente in una rete di *server* connessi tra loro e capaci di svolgere in autonomia miliardi di operazioni in una frazione di secondo.
Soprattutto negli ultimi anni, in particolare durante la campagna referendaria sulla permanenza del Regno Unito di Gran Bretagna e Irlanda del Nord nell'Unione europea e le elezioni presidenziali negli Stati Uniti d'America del 2016, il ruolo che queste piattaforme hanno avuto come catalizzatori e motori del dibattito all'interno della società civile è risultato particolarmente evidente[4]. Dopo l'approvazione della Brexit e l'elezione di Donald J. Trump, l'ampio catalogo di studi sull'incidenza sulle dinamiche sociali – e perfino sulle votazioni politiche – delle discussioni e delle interazioni che avvengono sui *social network* ha posto anche in luce come attività ben studiate ed organizzate di creazione e condivisione di contenuti sulle piattaforme siano in grado di creare l'illusione di un ampio consenso su argomenti che in realtà non sono così popolari all'interno della società per così dire "non digitale"[5]. In particolare, riguardo al referendum britannico alcuni studi hanno messo in evidenza come il ruolo dei *social media* nella determinazione del suo esito sia stato marginale e allo stesso tempo fondamentale, data una suddivisione delle preferenze dei votanti abbastanza vicina alla parità[6]. Inoltre, il dibattito sviluppatosi *online* circa l'evento referendario d'oltremanica

disponibile in *Digital News Report 2016*, in *reutersinstitute.politic.ox.ac*, 2016.

[4] Si vedano, *ex multis*, 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., in cui gli autori fanno anche riferimento ad uno studio del 2015 condotto dalla Ofcom (l'autorità indipendente per la regolazione delle società di comunicazione del Regno Unito), dal quale è risultato che il 43% di coloro che consultavano notizie su Internet lo facevano tramite i *social media*, e ad uno del 2017 dal quale è risultato che della fascia d'età 16-24 anni il 61% degli appartenenti usava i *social* come primaria–ma non unica–fonte d'informazione; W. Hall – R. Tinati – W. Jennings, *From Brexit to Trump: Social Media's Role in Democracy*, in *Computer*, 51(1), 2018, 18 ss.; L. Luceri – A. Deb – S. Giordano – E. Ferrara, *Evolution of bot and human behavior during elections*, in *firstmonday.org*, 2019; N. Grinberg – K. Joseph – L. Friedland – B. Swire-Thompson – D. Lazer, *Fake news on Twitter during the 2016 US presidential election*, in *Science*, 363(6425), 2019, 374 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.

[5] Si vedano, *ex multis*, M. Hänska – S. Bauchowitz, *Tweeting for Brexit: how social media influenced the referendum*, cit.; N. Persily, *The 2016 U.S. 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(2), 2017, 46 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 Politics*, 50(3), 2018, 435 ss.; J.A. Tucker – A. Guess – P. Barberá – C. Vaccari – A. Siegel – S. Sanovich – D. Stukal – B. Nyham, *Social media, political polarization, and political disinformation: A review of the scientific literature*, in *ssrn.com*, 2018, in cui gli autori svolgono un'approfondita rassegna degli studi compiuti sull'argomento; Y. Gorodnichenko – T. Pham – O. Talavera, *Social Media, Sentiment and Public Opinions: Evidence from #Brexit and #Uselection*, cit.

[6] Il "*Leave*" ha ottenuto il 51,8% dei voti mentre il "*Remain*" il 48,2%. I dati statistici sono stati

è stato anche il palcoscenico di un'estesa e strutturata operazione di propaganda sui *social* messa in atto da soggetti che di certo non avevano il diritto di esprimere il proprio voto nel corso della consultazione[7]: è stato ampiamente – ma solo successivamente – riportato come oltre 150.000 *account* localizzati all'interno della Federazione Russa abbiano pubblicato su Twitter contenuti inerenti alla Brexit, diffondendo tra gli elettori britannici informazioni non veritiere e dissimulandone la provenienza[8].

Insomma, queste "piazze digitali", ossia dei gruppi di *server* governati da *software* acritici e volti a permettere ai loro utenti la condivisione del maggior numero possibile di contenuti, hanno consentito a soggetti in possesso di capitali relativamente ridotti e di un certo grado di conoscenza del loro funzionamento di raggiungere una platea di proporzioni altrimenti inarrivabili e di crearsi, così, un seguito sempre pronto a promuovere e a condividere i contenuti che essi pubblicano sui *social*.

Quella che a prima vista pare una schiera di fedelissimi seguaci profondamente convinti della validità delle opinioni espresse da questi soggetti può, però, non essere effettivamente composta da esseri umani assiduamente presenti su queste piattaforme. Può trattarsi, invece, dei cosiddetti "*social bot*" o nella vulgata più comune – criticata da diversi ricercatori per la sua portata confusiva all'interno del dibattito – semplicemente "*bot*" (dalla contrazione delle parole *software robot*)[9], ossia dei programmi *software* autonomi progettati per agire su di un *social medium* simulando il modo in cui un utente umano si comporta su di esso. Tra questi vi è anche chi distingue la sottocategoria dei cosiddetti "*socialbot*" (una parola unica), spesso chiamati "*sybil*", ossia dei programmi *software* autonomi progettati per agire su di un *social medium* che, dissimulando la propria natura ed identità, si infiltrano nelle comunità *online* di utenti umani per diffondere contenuti (spesso mistificatori ed ingannevoli)[10].

L'utilizzo dei *socialbot* come strumenti di propaganda sui *social network*, oltre a giovarsi delle già viste caratteristiche di diffusione accelerata dei contenuti insite nell'architettura stessa di queste piattaforme, beneficia anche di un'altra importantissima caratte-

ottenuti consultando il sito della *Electoral Commission* britannica, l'ente indipendente di verifica dei risultati elettorali del Regno Unito, e sono disponibili in *Results and turnout at the EU referendum*, in *electoralcommission.org.uk*, 16 luglio 2019.

7 P.N. Howard – B. Kollanyi, *Bots, #Strongerin, and #Brexit: Computational Propaganda During the UK-EU Referendum*, in *ssrn.com*, 2016; M. Del Vicario – F. Zollo – G. Caldarelli – A. Scala – W. Quattrociocchi, *Mapping social dynamics on Facebook: The brexit debate*, in *Social Networks*, 50, 2017, 6 ss.; Digital, Culture, Media and Sport Committee della House of Commons, *Disinformation and 'fake news': Final Report*, HC1791, 2019; Y. Gorodnichenko – T. Pham – O. Talavera, *Social Media, Sentiment and Public Opinions: Evidence from #Brexit and #Uselection*, cit.

8 C. François, *Actors, Behaviours, Content: A Disinformation ABC*, in *Algorithms*, 2020; D. Wolchover – A. Robinson, *Is Brexit a Russia-backed Coup?*, in *New Law Journal*, 2020; J. Horder, *Online Free Speech and the Suppression of False Political Claims*, in *ILSA Journal of International and Comparative Law*, 2021.

9 Per un'accurata analisi della terminologia impiegata nella descrizione di questi particolari programmi *software* si veda quanto riportato in A. Tedeschi Toschi – G. Berni Ferretti, *Il contrasto legislativo ai socialbot e le soluzioni avanzate negli Stati Uniti d'America a livello federale e statale*, in *Diritto di Internet*, 3, 2022, 451 ss.

10 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 (ACSAC '11)*, 2011, 93 ss.; 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.; O. Varol – E. Ferrara – C. Davis – F. Menczer – A. Flammini, *Online human-bot interactions: Detection, estimation, and characterization*, in *Proceedings of the international AAAI conference on web and social media*, 11(1), 2017, 280 ss.

ristica comune alla maggior parte delle tecnologie digitali: la ripetibilità[11]. Infatti, una volta creato un *socialbot* – un programma piuttosto "leggero" per le capacità di calcolo dei moderni computer – è possibile copiarlo e dare al suo clone le credenziali di un altro *account* nel giro di pochissimi minuti, creando così un secondo *follower*, che, come il primo, sarà un fedelissimo sostenitore di qualsiasi profilo sia programmato a seguire. Questo processo può essere quindi ripetuto un'altra volta e un'altra ancora, tanti quanti sono i profili che si hanno a disposizione, creando così folle oceaniche di sostenitori artificiali di chiunque possa permettersi di finanziare questo procedimento.

L'aspetto più problematico di un tale processo è il fatto che, come accennato, chiunque abbia o sia in grado di reperire un capitale adeguato (magari anche grazie all'aiuto di gruppi di interesse stranieri)[12] sarà in grado di crearsi una rete di fedelissimi seguaci artificiali programmati per osannarlo e per ripetere prontamente qualunque cosa dica, senza che questa sia effettivamente convincente, veritiera o anche solo rispettosa dei più basilari principi del rispetto della persona (finendo, così, col configurare questi seguaci sintetici come dei *socialbot*)[13], e che diano ai meno esperti l'illusione che sia invece il grande pubblico a condividere e a seguire le sue opinioni[14].

Oltre all'insidioso aspetto della dilatazione artificiale dell'apprezzamento esibito sui *social media* per alcuni soggetti (sia pubblici che politici), è stato osservato come i *socialbot* agiscano anche nell'alterazione dei temi e delle discussioni *online* tramite modalità meno celebrative dei loro "padroni" e delle dichiarazioni che questi fanno. Una parte non indifferente delle attività di questi agenti *software* viene, infatti, occupata dalla propagazione o dal supporto di contenuti, interpretazioni od opinioni legati a temi al

[11] O. Varol – E. Ferrara – C. Davis – F. Menczer – A. Flammini, *Online human-bot interactions: Detection, estimation, and characterization*, cit.; L. Luceri – A. Deb – S. Giordano – E. Ferrara, *Evolution of bot and human behavior during elections*, cit., in cui gli autori fanno anche riferimento a B. Mønsted – P. Sapiezynski – E. Ferrara – S. Lehmann, *Evidence of complex contagion of information in social media: An experiment using twitter bots*, in PLoS *ONE*,12, 2017, e0184148; O. Boichak – S. Jackson – J. Hemsley – S. Tanupabrungsun, *Automated diffusion? bots and their influence during the 2016 us presidential election*, in *International Conference on Information*, 2018, 17 ss.; C. Shao – G.L. Ciampaglia – O. Varol – K.C. Yang – A. Flammini – F. Menczer, *The spread of low-credibility content by social bots*, in *Nature communications*, 9(1), 2018, 47 ss.

[12] A questo riguardo si veda, in particolare, quanto accuratamente descritto in T.C. Helmus – E. Bodine-Baron – A. Radin – M. Magnuson – J. Mendelsohn – W. Marcellino – A. Bega – Z. Winkelman, *Russian social media influence: Understanding Russian propaganda in Eastern Europe*, Santa Monica, 2018; F. Giglietto – N. Righetti – G. Marino, *Understanding coordinated and inauthentic link sharing behavior on Facebook in the run-up to 2018 general election and 2019 European election in Italy*, in *osf.io*, 20 settembre 2019.

[13] Si veda a questo proposito l'analisi sul rischio di lesioni irreparabili della reputazione amplificate dall'uso di *socialbot* compiuta in 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.

[14] Y. Xie – F. Yu – Q. Ke – M. Abadi – E. Gillum – K. Vitaldevaria – J. Walter – J. Huang – Z.M. Mao, *Innocent by association: Early recognition of legitimate users*, in *Proceedings of the 2012 ACM Conference on Computer and Communications*, 2012, 353 ss.; D. Westerman – P.R. Spence – B. Van Der Heide, *A social network as information: The effect of system generated reports of connectedness on credibility and health care information on Twitter*, in *Computers in Human Behavior*, 28(1), 2012, 199 ss.; M. Forelle – P.N. Howard – A. Monroy-Hernández – S. Savage, *Political bots and the manipulation of public opinion in Venezuela*, in *ssrn.com*, 1 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.C. Woolley, *Automating power: Social bot interference in global politics*, in *First Monday*, 21(4), 2016; D. Guilbeault, *Growing Bot Security: An Ecological View of Bot Agency*, in *International Journal of Communication*, 10, 2016, 5003 ss.

centro dell'interesse pubblico senza dare l'impressione che siano emanazione diretta di un determinato soggetto ma, invece, di un gruppo più o meno ampio e spontaneo. Questo approccio alla diffusione potenziata di determinati contenuti al centro del dibattito può essere spiegato dal fatto che, come già mostrato da alcuni importanti studi compiuti a metà del secolo scorso[15], la formazione del pensiero politico degli individui è influenzata da più reti interpersonali di natura diversa (come amicizie, colleghi di lavoro e famiglia), che essi tendono a prestare attenzione a ciò che sembra popolare e a fidarsi delle informazioni che circolano all'interno del loro contesto sociale[16] e che oggigiorno anche le cerchie virtuali di connessioni giocano un ruolo importante nella formazione delle convinzioni personali e delle opinioni politiche[17]. Così, l'uso dei *socialbot* è divenuto per tutti coloro che ne siano entrati in possesso grazie all'impiego di capitali propri o di terzi, da alcuni soprannominati anche «*augmented humans*» (ossia «umani potenziati»)[18], un valido strumento di sviamento del naturale orientamento delle discussioni *online* tra utenti ed un ancor più valido mezzo per indirizzare le persone attive sui *social* verso le loro posizioni, dando, però, l'illusione che queste siano popolari e condivisibili per via di una loro intrinseca validità e non a causa di un'intensa e ben finanziata campagna propagandistica.

Di fatto, quindi, chi sappia creare e ben orchestrare – oppure, più semplicemente, sia in grado di comprarsi – queste folle di utenti artificiali ben potrebbe fingersi un semplice interprete dei sentimenti delle masse attive *online*, anche se in verità ne sta dettando gli orientamenti in maniera unilaterale, o anche apparire completamente estraneo agli umori e alle direzioni presi da esse, pur essendone in realtà l'unico vero direttore. In altre parole, grazie ai *socialbot*, i soggetti in possesso delle giuste conoscenze possono presentarsi come dei moderni vati della società, quando in realtà, per usare le parole di Ferrara, Varol, Davis, Menczer e Flamini[19], sono dei burattinai che muovono a proprio piacimento i fili di migliaia di pupazzi digitali, liberi di orientarli nella direzione che più gli aggrada.

---

[15] 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.R. 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.

[16] Y. Jun – R. Meng – G.V. Johar, *Perceived social presence reduces fact-checking*, in *Proceedings of the National Academy of Sciences*, 114(23), 2017, 5976 ss.; K.C. Yang – O. Varol – C. Davis – E. Ferrara – A. Flammini – F. Menczer, *Arming the public with artificial intelligence to counter social bots*, in *Human Behavior and Emerging Technologies*, 1(1), 2019, 48 ss.

[17] D. Centola, *The spread of behavior in an online social network experiment*, in *Science*, 329(5996), 2010, 1194 ss.; R.M. Bond – C.J. Fariss – J.J. Jones – A.D. Kramer – C. Marlow – J.E. Settle – J.H. Fowler, *A 61-million-person experiment in social influence and political mobilization*, in *Nature*, 489(7415), 2012, 295 ss.; L. Muchnik – S. Aral – S.J. Taylor, *Social influence bias: A randomized experiment*, in *Science*, 341(6146), 2013, 647 ss.; C. Becatti – G. Caldarelli – R. Lambiotte – F. Saracco, *Extracting significant signal of news consumption from social networks: the case of Twitter in Italian political elections*, in *Palgrave Communications*, 5(1), 2019, 1 ss., in cui gli autori fanno anche riferimento a G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*, Oxford, 2007; M.E.J. Newman, *Networks: an introduction*, New York, 2010.

[18] M. Stella – M. Cristoforetti – M. De Domenico, *Influence of augmented humans in online interactions during voting events*, in *PLoS ONE*, 14(5), 2019, e0214210, in cui gli autori usano questi termini per "*per indicare specificamente quegli* account *umani che sfruttano i* bot *per aumentare artificialmente, cioè* potenziare*, la loro influenza nel mondo digitale*".

[19] O. Varol – E. Ferrara – C. Davis – F. Menczer – A. Flammini, *The Rise of Social Bots*, cit., 103.

2. I tentativi di riforma legislativa nel mondo

Di fronte alle conseguenze delle influenze straniere sperimentate sui *social media* ed ai risultati concreti – soprattutto in ambito elettorale – di cui sono stati testimoni, i governi di diverse nazioni hanno istituito commissioni di indagine per comprendere a fondo quali fossero state le dinamiche all'interno delle reti sociali *online* che avevano avuto un'ascendente su questi risultati[20]. Le conclusioni a cui queste sono giunte si sono accompagnate, poi, ad un impegno da parte degli organi legislativi ad adottare delle riforme volte ad ostacolare le attività di disinformazione e di mistificazione avvenute sulle piattaforme digitali e che, come detto, avevano avuto riflessi incontrollati anche sulla società *off-line*. Un breve spazio di intervento legislativo è stato anche dedicato specificamente al contrasto delle attività di diffusione di contenuti compiute dai *socialbot* con, tuttavia, un'effettiva incidenza sul fenomeno che da alcuni è stata messa in dubbio[21].

Tra le poche nazioni a reagire attivamente – e legislativamente – alle attività di questi agenti *software* si sono distinti gli Stati Uniti d'America, i quali, però, nonostante la profondità degli effetti che la propaganda digitale di nazioni ad essi avverse aveva avuto sul loro tessuto sociale e politico, non sono ancora riusciti ad approvare una normativa federale[22]. Anche la piccola Repubblica di Singapore ha adottate alcuni provvedimenti di riforma del proprio ordinamento al fine di contrastare la diffusione – e soprattutto la credibilità – di contenuti fuorvianti, mistificatori o semplicemente falsi tramite Internet, i *social media* ed anche i *socialbot*. Invece, nella Repubblica d'Irlanda è in discussione al Dáil Éireann (la camera bassa del Parlamento) una proposta di legge per «garantire la trasparenza nella divulgazione delle informazioni nella pubblicità politica online». Avendo già in altra sede compiuto un esame delle normative statunitensi proposte od approvate di regolamentazione dell'uso di *social bot* e *socialbot*[23], di seguito vengono analizzate le disposizioni contenute nel *Protection from Online Falsehoods and Manipulation Act 2019* della Città dei Leoni e nell'*Online Advertising and Social Media (Transparency) Bill 2017* della repubblica dell'Isola di Smeraldo.

[20] Si vedano, ad esempio, le conclusioni a cui è giunta la U.S. Senate Select Committee on Intelligence sulle interferenze della Federazione Russa nelle elezioni presidenziali americane del 2016, la relazione (divisa in cinque parti) è disponibile per la consultazione in *Russian active measures campaigns and interference in the 2016 U.S. election*, *intelligence.senate.gov*, 10 novembre 2020, e i risultati a cui è pervenuta la Digital, Culture, Media and Sport Committee della House of Commons britannica sul ruolo occupato dagli apparati di disinformazione russi nell'influenzare il referendum sulla Brexit, il rapporto è già stato citato *supra* alla nota 10.

[21] S. Woolley – N. Monaco, *Amplify the party, suppress the opposition: social media, bots, and electoral fraud*, in *Georgetown Law Technology Review*, 4, 2020, 447 ss.

[22] Per una disamina delle normative proposte od approvate all'interno dell'ordinamento degli Stati Uniti d'America per contrastare le attività dei *socialbot* si veda A. Tedeschi Toschi – G. Berni Ferretti, *Il contrasto legislativo ai socialbot e le soluzioni avanzate negli Stati Uniti d'America a livello federale e statale*, cit.

[23] Ivi.

2.1. Il *Protection from Online Falsehoods and Manipulation Act 2019* della Repubblica di Singapore

Il 02 ottobre 2019 è entrato in vigore il *Protection from Online Falsehoods and Manipulation Act 2019* (*POFMA*), approvato dal Parlamento di Singapore il 03 giugno 2019[24]. Questa estesa normativa – composta da ben 62 Sezioni – è stata adottata dalla Città dei Leoni al fine di affrontare non solo l'impiego di *bot* per potenziare la propagazione di contenuti *online* ma soprattutto il più ampio problema della diffusione delle "*fake news*", ossia, come esplicitato nel preambolo della normativa stessa: «Per impedire la comunicazione elettronica a Singapore di false dichiarazioni su di un fatto, per sopprimere il supporto a e contrastare gli effetti di tale comunicazione, per salvaguardare dall'uso di *account online* per tale comunicazione e per la manipolazione delle informazioni, per consentire l'adozione di misure per migliorare la trasparenza degli annunci politici *online* e per questioni correlate»[25].

La volontà della Repubblica di Singapore di tutelare i suoi cittadini dall'esposizione a contenuti digitali falsi e manipolatori attraverso una normativa di ampio respiro è stata oggetto di numerose critiche da parte di gruppi internazionali di difesa dei diritti umani, quali la Commissione Internazionale dei Giuristi (ICJ)[26], Reporters Sans Frontiers[27] e Human Rights Watch[28], che hanno sostenuto che la legge sopprimerebbe indebitamente la libertà di parola e potrebbe essere utilizzata in modo abusivo dal governo per ridurre al silenzio qualsiasi forma di dissenso espresso *online*. Anche in dottrina è stato messo in evidenza da molti come sia evidente il potenziale effetto di diminuire il senso di libertà nell'esprimersi (per il quale viene spesso utilizzata in inglese la locuzione «*chilling effect on the freedom of speech*») e di indurre forme di auto-censura[29], nonostante le

[24] Il testo della legge assieme ai suoi emendamenti è disponibile è liberamente consultabile sul sito istituzionale del Governo di Singapore in *Protection from Online Falsehoods and Manipulation Act 2019*, in *sso.agc.gov.sg*, 2 ottobre 2019.

[25] Preambolo del *POFMA*.

[26] International Commission of Jurists, *Singapore: ICJ calls on government not to adopt online regulation bill in current form*, in *icj.org*, 12 aprile 2019.

[27] Reporters Sans Frontiers, *RSF explains why Singapore's anti-fake news bill is terrible*, in *rsf.org*, 8 aprile 2019.

[28] Human Rights Watch, *Singapore: Reject Sweeping 'Fake News' Bill Proposed Law Would Excessively Restrict Online Freedom of Speech*, in *hrw.org*, 3 aprile 2019.

[29] S. Chen – C.W. Chia, *Singapore's latest efforts at regulating online hate speech*, in *Research Collection School Of Law*, 6, 2019, 1 ss.; K. Han, *Big Brother's regional ripple effect: Singapore's recent "fake news" law which gives ministers the right to ban content they do not like, may encourage other regimes in south-east Asia to follow suit*, in *Index on Censorship*, 48(2), 2019, 67 ss.; H. Lee – T. Lee, *From contempt of court to fake news: public legitimisation and governance in mediated Singapore*, in *Media International Australia*, 173(1), 2019, 81 ss.; S. Neubronner, *Bill to Protect from Online Falsehoods: Refinements Needed*, in *RSIS Commentaries*, 85, 2019, 1 ss.; J. Horder, *Online Free Speech and the Suppression of False Political Claims*, cit.; R.K.C. Ren, *Suppressing Fake News or Chilling Free Speech: Are the Regulatory Regimes of Malaysia and Singapore Compatible With International Law?*, in *JMCL*, 47, 2020, 25 ss.; A. Carson – L. Fallon, *Fighting fake news: a study of online misinformation regulation in the Asia Pacific*, in *opal.latrobe.edu.au*, 2021, 1 ss.; R.K. Helm – H. Nasu, *Regulatory responses to 'fake news' and freedom of expression: normative and empirical evaluation*, in *Human Rights Law Review*, 21, 2021, 302 ss.; J.Y. Tay, *No news is good news, but "fake news" is bad news: A comparative analysis of Singapore's and Australia's measures to combat misinformation on social media*, in *Singapore Academy of Law Journal*, 33(2), 2021, 600 ss.; M. Mohan, *Singapore and the Universal Periodic Review: An Ongoing Human Rights Assessment*, in E. Yong Joong Lee (a

molteplici rassicurazioni ed i chiarimenti fatti dal Ministero di Giustizia singaporiano (il *Ministry of Law*) sul fatto che l'obiettivo della legge è di prendere di mira le falsità e non le opinioni, le critiche, le satire e le parodie, che il principale provvedimento adottabile dall'Autorità Competente è volto alla correzione e non alla rimozione dei contenuti e che qualsiasi decisione del Governo su ciò che debba considerarsi come «una falsa dichiarazione su di un fatto» potrà essere annullata da un tribunale[30].
Come accennato, il *POFMA* è una normativa piuttosto ampia che contiene definizioni e chiarimenti in merito a determinati termini (Parte 1), divieti di condotte relative alla diffusione di contenuti (Parte 2), indicazioni ("*directions*") alle autorità competenti su quando e come intervenire per contrastare la diffusione di affermazioni false (Parte 3), sanzioni comminabili a coloro che non ottemperano agli ordini delle autorità competenti (Parte 3), disposizioni procedurali per fare ricorso («*appeal*») presso la *General Division* della *High Court* contro le sanzioni irrogate (Parte 3), prescrizioni per gli *Internet Services Providers* e per i «*Providers of Mass Media Services*» (Parte 4), disposizioni in merito alla attestazione delle localizzazioni *online* ("*declaration of online locations*"), indicazioni per contrastare *account online* non autentici e comportamenti coordinati non autentici (Parte 6), disposizioni per gli intermediari di pubblicità *online* e di servizi Internet (Parte 7), indicazioni in merito alle autorità competenti alternative in periodo elettorale (Parte 8), indicazioni ("*directions*") sui soggetti che dovrebbero applicare gli ordini di interruzione dei servizi di accesso a Internet (Parte 9), disposizioni in merito all'ipotesi in cui sia un'impresa a compiere atti di diffusione di affermazioni false (Parte 9).
All'interno di questo vasto ed eterogeneo insieme di disposizioni che è il *Protection from Online Falsehoods and Manipulation Act 2019* ve ne sono anche diverse – e sparse – relative all'impiego dei *bot* per la diffusione di «false dichiarazioni su di un fatto» ed alle sanzioni applicabili per queste condotte.

(A) Definizione di *bot* – Nella prima parte della legge viene affermato che il termine «*bot*» indica: «Un programma per computer creato o alterato con lo scopo di eseguire attività automatizzate» e che gli *account* controllati da uno di questi agenti *software* non rientrano nella definizione contenuta nel *POFMA* di «*account online* non autentico»[31] (in originale «*inauthentic online account*»).
Pare qui opportuno iniziare precisando che l'esclusione degli *account* controllati da un *bot* dal novero di quelli definiti come «*account online* non autentici» non deriva dall'idea che si possa attribuire una qualche forma di autenticità alle condotte tenute su di un *social medium* da questa categoria di programmi *software* automatizzati, quanto, piuttosto, dal fatto che la stessa normativa di Singapore definisce un «*account online* non autentico» (in originale un «*inauthentic online account*») come un: «*Account online* che è controllato da una persona diversa dalla persona rappresentata (sia [che ciò avvenga] tramite il profilo

cura di), *ASEAN International Law*, Singapore, 2022, 427 ss.

[30] Si vedano la trascrizione del dibattito parlamentare avvenuto in data 8 maggio 2019 in sede di seconda lettura del *POFMA*, disponibile in *Official Reports – Parliamentary Debate (Hansard)*, in *sprs.parl.gov.sg*, ed il comunicato stampa del Ministry of Law di Singapore, *New Bill to Protect Society from Online Falsehoods and Malicious Actors*, in *mlaw.gov.sg*, 01 aprile 2019.

[31] Sezione 2 (1) del *POFMA*.

utente, l'identificatore univoco o altre informazioni) come il suo titolare, e la rappresentazione è effettuata con l'obiettivo di ingannare gli utenti finali a Singapore di un qualsiasi servizio di intermediazione di Internet sull'identità del titolare»[32].

In altre parole, il legislatore singaporiano ha attribuito alla definizione di «*account online* non autentico» gli stessi caratteri che la nostra giurisprudenza nazionale ha riconosciuto come integranti il reato di sostituzione di persona all'interno di un *social network*[33], distinguendo l'ipotesi in cui una persona usi un profilo *online* dotato di questo caratteri di inautenticità da quella in cui essa usi un *bot* per gestire un profilo ed attribuendo ad entrambe le eventualità il carattere di aggravio del disvalore delle condotte di diffusione di «false affermazioni su di un fatto» a Singapore (senza, però, prendere esplicitamente in considerazione l'eventualità che queste ipotesi vengano in concreto combinate)[34]. Questa strutturazione della normativa trova la propria motivazione nel fatto che il *POFMA* non ha come primario obiettivo quello di impedire l'uso di strumenti di automazione dei profili *online* ma solo di inibire il loro impiego per la propagazione "potenziata" di determinati contenuti, permettendo, quindi, al titolare effettivo di un *account* di lasciarlo in gestione ad un *bot*, purché con esso non compia degli illeciti. Riprendendo un esempio già fatto da altri[35], si pensi ad un giornalista finanziario che, per poter immediatamente informare i suoi lettori sulle variazioni di borsa tramite il proprio profilo *social*, abbia dato ad un *social bot* le credenziali di accesso ad esso e gli abbia dato il compito di pubblicare automaticamente un *post* ad ogni fluttuazione rilevante degli indici azionari. In base alla normativa della Città dei Leoni ciò sarebbe lecito perché il titolare dell'*account* è colui che lo gestisce – *rectius* che gestisce il *software* che a sua volta lo gestisce – e perché i contenuti pubblicati non verrebbero considerati dalle autorità come delle «false dichiarazioni su di un fatto».

Inoltre, è da sottolineare come la definizione di «*bot*» data all'interno del *POFMA* sia estremamente ampia e permetta, quindi, di sanzionare non solo il suo illecito impiego per la diffusione di «false dichiarazioni su di un fatto» all'interno di una determinata piattaforma come un *social medium* o, più in generale, su Internet ma addirittura attraverso l'uso di altre infrastrutture di comunicazione digitale. Infatti, la Sezione 3 (2)(b)

---

[32] Sezione 2 (1) del *POFMA*.

[33] Si vedano, ad esempio, Cass. Pen., sez. V, 23 luglio 2020, n. 22049, in cui viene specificamente affermato che «il reato di sostituzione di persona è integrato da colui che crea ed utilizza un profilo su *social network*, utilizzando abusivamente l'immagine di una persona del tutto inconsapevole, trattandosi di condotta idonea alla rappresentazione di una identità digitale non corrispondente al soggetto che lo utilizza», richiamando anche la precedente Cass. Pen., sez. V, 08 giugno 2018, n. 33862. Prima ancora Cass. Pen., sez. V, 16 giungo 2014, n. 25774, aveva sancito 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».

[34] Più volte, infatti, all'interno della normativa vengono considerati come ipotesi separate l'uso di un «*account online* non autentico» o l'impiego di un *bot* per comunicare una falsa affermazione su di un fatto o accelerarne la diffusione. Si vedano ad esempio le Sezioni 7 (3) e 40 (4) del *POFMA*. Tuttavia, non si esclude che l'Autorità Competente possa ritenere di comminare il massimo delle sanzioni previste dalla legge, nel caso di combinazione dell'uso di un «*account online* non autentico» e di un *bot* per la sua gestione.

[35] Si veda, *ex multis*, G. Caldarelli – R. De Nicola – F. Del Vigna – M. Petrocchi – F. Saracco, *The role of bot squads in the political propaganda on Twitter*, in *Communication Physics*, 3(1), 2020, 1 ss.

esplicitamente dichiara che il divieto di compiere atti «per comunicare a Singapore una dichiarazione sapendo o avendo motivo di credere che (a) si tratti di una falsa dichiarazione su di fatto» si estende anche all'uso di MMS e SMS.
Infine, pare opportuno rilevare come dalla lettura della normativa scaturisca il dubbio se essa possa trovare applicazione anche quando la comunicazione di «false dichiarazioni su di un fatto» avvenga all'interno di una *intranet*[36]. Infatti, sebbene da un'interpretazione formalistica della lettera del *POFMA* sembra potersi escludere questa eventualità (la stessa definizione di "comunicare" contenuta nella Sezione 3 fa riferimento solo all'uso di Internet o di MMS e SMS), l'esclusione dell'uso di una struttura di comunicazione chiusa ma accessibile ad un numero potenzialmente molto alto di utenti – ed anche di *bot* – per la commissione dell'illecito al centro della normativa stessa pare più una svista del legislatore che una decisione ponderata[37].

(B) Fattori per determinare se un profilo *online* è controllato da un *bot* – Di fronte alla diffusione *online* di contenuti falsi su di un fatto, il *POFMA* indica anche una serie di fattori da prendere in considerazione per determinare se il profilo con cui è stata compiuta questa violazione sia gestito da un essere umano ovvero da un *bot*. La Sezione 40 (4), 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*[38]; (c) la data in cui è stato creato il profilo *online*; (d) qualsiasi altro fattore che il Ministro [competente] consideri rilevante».
Al fine, poi, di arginare l'eccessiva discrezionalità dell'Autorità nel determinare se dalle azioni compiute dal profilo considerato risulti evidente l'utilizzo di un *bot*, la successiva sottosezione (5) impone che qualsiasi provvedimento di restrizione dell'uso dell'*account* adottato (ossia qualsiasi "*Account Restriction Direction*") «deve identificare in modo sufficientemente dettagliato l'affermazione dell'oggetto o il comportamento del soggetto, a seconda dei casi».
L'inclusione nella normativa di indicazioni circa gli aspetti da valutare per determinare se la diffusione di «false dichiarazioni su di un fatto» sia stata compiuta tramite l'impiego di un *bot*, nonostante un'evidente vaghezza, ha il pregio di fornire una prima serie

[36] Per un'analisi del funzionamento di una *intranet* e dell'opportunità per le aziende di adottarne una si vedano, ad esempio, M. Hills, *Intranet business strategies*, Hoboken, 1996; S. Baker, *Getting the most from your intranet and extranet strategies*, in *Journal of Business Strategy*, 21, 2000, 40 ss.

[37] Si pensi, ad esempio, ad una persona che utilizzi un *bot*, ossia un programma per «eseguire attività automatizzate», per comunicare tramite la rete interna di un'impresa, ossia una *intranet*, con sede a Singapore delle «false dichiarazioni su di un fatto» a tutti i dipendenti di questa. Pur non essendo questo agente *software* connesso alla rete Internet mondiale e non potendo, quindi, diffondere "*online*" questi contenuti, il suo utilizzo potrebbe essere interpretato come una violazione del *POFMA*, dato che, riprendendo le parole della Sezione 7 (1), esso verrebbe impiegato «per comunicare a Singapore una affermazione sapendo o avendo motivo di credere che (a) si tratti di una falsa affermazione su di fatto» ad un numero potenzialmente molto alto di utenti che possono accedere all'infrastruttura digitale privata sulla quale avviene la trasmissione delle false informazioni.

[38] In originale alla Sezione 40 (4)(b) viene impiegata la locuzione «*pattern of suspicious activity*».

di indici che permetta alle Autorità di avere una direzione verso cui orientare lo svolgimento delle proprie indagini, mantenendo, comunque, grazie alla disposizione di cui alla lettera (d) della Sezione 40 (4), un certo grado di discrezionalità nella valutazione e soprattutto nell'inclusione degli elementi da cui desumere questo utilizzo per il compimento di un illecito. Inoltre, pare opportuno evidenziare come gli indici esplicitati nella normativa siano – in parte – quelli che già da tempo vengono stati utilizzati da molti ricercatori per stimare la presenza di *socialbot* sui *social media*[39]. Infatti, dal momento che per l'apertura di un profilo su di una piattaforma *social* non è richiesta alcuna certificazione da parte di una autorità centrale che attesti l'identità del creatore o dell'utilizzatore di questo, la maggior parte degli strumenti di rilevazione dei *socialbot* si sono strutturati in modo da giudicare la natura di chi gestisce un profilo *social* valutando la corrispondenza delle interazioni compiute con dei modelli comportamentali standard estrapolati dai dati statistici relativi ai comportamenti di un gran numero di profili usati da utenti umani. In altre parole, gli strumenti di controllo di "attività *bot*" oggi si basano principalmente sull'analisi dei comportamenti tenuti dai profili all'interno dei *social network* e sulla corrispondenza con gli atteggiamenti abitualmente tenuti 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. Tuttavia, come già ben messo in evidenza in diversi studi[40], questi agenti *software* "malevoli" sono attivi già da molto tempo all'interno delle piattaforme *social* ed hanno subito – e continuano a subire – un affinamento ininterrotto delle loro *routine* interne di mimica dei comportamenti degli utenti umani[41]. Di conseguenza, ormai l'efficacia degli strumenti di rilevazione dipende in gran parte dal grado di sofisticatezza dei *socialbot* analizzati e, quindi, quelli maggiormente capaci di imitare le azioni compiute mediamente dagli utenti umani sono in grado di passare senza problemi il vaglio di questi strumenti di controllo[42]. Tutto ciò comporta che la

[39] Si veda, ad esempio, S. Cresci, *A decade of social bot detection*, in *Communications of the ACM*, 63(10), 2020, 72 ss., in cui l'autore svolge un'approfonditissima analisi dell'evoluzione dei *social bot* e dei metodi di analisi e rilevamento di questi.

[40] Si veda, *ex multis*, quanto dichiarato in Alvisi – Clement – Epasto – Lattanzi – Panconesi, *Sok: The evolution of sybil defense via social networks*, cit.; S. Cresci – R. Di Pietro – M. Petrocchi – A. Spognardi – M. Tesconi, *The paradigm-shift of social spambots*, in *Proceedings of the 26th international conference on world wide web companion*, 2017, 963 ss.; S. Cresci – R. Di Pietro – M. Petrocchi – A. Spognardi – M. Tesconi, *Social Fingerprinting: detection of spambot groups through DNA-inspired behavioural modelling*, in *IEEE Transactions on Dependable and Secure Computing*,15(4), 2017, 561 ss.; L. Luceri – A. Deb – S. Giordano – E. Ferrara, *Evolution of bot and human behavior during elections*, cit.; R.J. Schuchard – A.T. Crooks, *Insights into elections: An ensemble bot detection coverage framework applied to the 2018 U.S. midterm elections*, in *PLoS ONE*, 16(1), 2021, e0244309.

[41] K.C. Yang – R. Harkreader – G. Gu, *Empirical evaluation and new design for fighting evolving twitter spammers*, in *IEEE Transactions on Information Forensics and Security*, 8(8), 2013, 1280 ss.; K.C. Yang – O. Varol – C. Davis – E. Ferrara – A. Flammini – F. Menczer, *Arming the public with artificial intelligence to counter social bots*, cit.; S. Cresci, *A decade of social bot detection*, cit., in cui l'autore svolge un'approfonditissima analisi dell'evoluzione dei *social bot* e dei metodi di analisi e rilevamento di questi; M. Himelein-Wachowiak – S. Giorgi – A. Devoto – M. Rahman – L. Ungar – H.A. Schwartz – D.H. Epstein – L. Leggio – B. Curtis, *Bots and Misinformation Spread on Social Media: Implications for COVID-19*, in *Journal of Medical Internet Research*, 23(5), 2021, e26933.

[42] Si vedano, ad esempio, le conclusioni raggiunte in S. Cresci – R. Di Pietro – M. Petrocchi – A.

previsione di cui alla Sezione 40 (4)(b) del *POFMA* rischia di perdere di efficacia nella pratica, dato che le attività svolte dall'*account online* considerato rischiano di non destare sospetti per quanto riguarda aspetti comportamentali e di reattività.
Inoltre, anche la previsione della lettera (c) della medesima Sezione, relativa alla data di creazione del profilo, presenta alcune criticità, dato che è già stata ipotizzata da tempo l'esistenza di un mercato nero di *socialbot* riutilizzabili per campagne di disinformazione politica, i quali si appoggerebbero a profili vecchi di anni e già in precedenza molto attivi[43].
In definitiva, perché venga svolta un'accurata valutazione che permetta effettivamente di riconoscere se un *account* è gestito da un *bot* e, di conseguenza, possa essere adottato dall'Autorità Competente un provvedimento di restrizione del suo utilizzo (in originale una «*Account Restriction Direction*»), è di estrema importanza non limitarsi alla considerazione dei primi fattori elencati dalla Sezione 40 (4) del *POFMA* e dare, invece, il massimo risalto alla Sezione 40 (4)(d), che permette l'inclusione di elementi ulteriori (si spera, riprendendo gli studi e le ricerche più avanzate compiuti in ambito informatico)[44]. Questo perché, come sopra accennato in merito allo stato dell'arte della progettazione di questi programmi *software* automatizzati, già adesso le loro evoluzioni più sofisticate sono in grado di imitare le condotte degli utenti umani in modo sufficientemente credibile da non essere riconosciuti.

(C) Divieto di creazione o di alterazione di un *bot* per la comunicazione di false affermazioni su di un fatto a Singapore – Alla Sezione 8 (1) del *POFMA* viene sancito che: «Una persona non deve, sia dentro che fuori Singapore, creare o alterare un *bot* con l'intenzione di (a) comunicare a Singapore, tramite l'uso del *bot*, una falsa dichiarazione su di un fatto; oppure (b) consentire a qualsiasi altra persona di comunicare a Singapore, tramite l'uso del *bot*, una falsa dichiarazione su di un fatto».
In merito a questa disposizione pare opportuno preliminarmente sottolineare come essa determini l'integrazione di un reato non nel caso di realizzazione di un evento dannoso ma, invece, al semplice compimento di una condotta motivata dall'intenzione di commettere, poi, un altro atto illecito. Infatti, perché l'attività di creazione o di alterazione di un programma *software* autonomo assuma il carattere dell'antigiuridicità, la norma non prevede che con essa il suo autore abbia causato alcun danno ad altri ma solo che abbia tenuto dei comportamenti che il legislatore considera come idonei a porre in pericolo l'ordine pubblico di Singapore e la sicurezza dei suoi cittadini.

Spognardi – M. Tesconi, *The paradigm-shift of social spambots*, cit.

[43] E. Ferrara, *Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election*, in *First Monday*, 22(8), 2017, dove l'autore riferisce di aver riscontrato nei suoi studi dei *pattern* dell'uso di *account* anomali, dato che essi erano risultati molto attivi durante la campagna presidenziale statunitense del 2016 (supportando Donald J. Trump) e poi erano rimasti inattivi fino alle elezioni presidenziali francesi del 2017, durante le quali avevano contribuito alla campagna di disinformazione soprannominata *MacronLeaks* (attaccando Emmanuel Macron).

[44] Si vedano, ad esempio, le (dichiarate) promettenti opere di analisi dei comportamenti collettivi menzionate in S. Cresci – R. Di Pietro – M. Petrocchi – A. Spognardi – M. Tesconi, *The paradigm-shift of social spambots*, cit. e la tecnica di rilevamento di gruppi di *spambot* attraverso modelli comportamentali ispirati alla sequenziazione del DNA descritta in S. Cresci – R. Di Pietro – M. Petrocchi – A. Spognardi – M. Tesconi, *Social Fingerprinting: detection of spambot groups through DNA-inspired behavioural modelling*, cit.

Vi è, poi, da evidenziare come la Sezione 8 (1) non punisca la mera creazione di un agente *software* o la sua modifica – anche se con quest'ultima condotta è comunque possibile violare il diritto d'autore del suo creatore originale – ma richieda, invece, che il compimento di queste condotte sia motivato dalla finalità di utilizzare gli strumenti informatici creati o innovati per la diffusione di contenuti che contengano falsità su di un fatto perché esse vengano sanzionate come illecite. In altre parole, è solo in presenza dell'intenzione di impiegarne il prodotto per il compimento di un ben determinato crimine, ossia il comunicare ai destinatari di un'interazione *online* delle informazioni non veritiere (sancito come illecito dalla Sezione 7 del *POFMA*), che la creazione o la modifica di un programma *software* autonomo, fatto di per sé lecito (tranne nella già vista ipotesi di violazione del diritto d'autore di terzi), integrano il reato previsto dalla Sezione 8 (1) della normativa.

L'applicazione pratica delle disposizioni di questa Sezione rischia di risultare meno agevole di quanto non possa apparire a prima vista. Infatti, la comunicazione *online* di contenuti ritenuti affetti da falsità può avvenire sostanzialmente secondo due modalità: (I) tramite la pubblicazione in un determinato ambiente informatico di testi, immagini e video (ad esempio, su Facebook ciò può avvenire tramite la pubblicazione di un "*post*") oppure (II) attraverso la condivisione di contenuti già pubblicati altrove (quando la struttura dell'ambiente in cui il *bot* opera lo permette). Entrambe queste attività si basano sull'interazione di un *bot* con elementi solitamente esterni al proprio codice sorgente (nulla impedisce che gli elementi da pubblicare siano, invece, contenuti nel codice sorgente ma è buona pratica strutturare un programma *software* autonomo in modo che i dati che deve diffondere siano esterni ad esso) e che, quindi, possono essere compiute con finalità diverse nel tempo oppure anche esulando dal controllo del suo creatore[45].

Di conseguenza, per quanto riguarda la condotta di cui alla Sezione 8 (1)(a), ossia l'utilizzo di un *bot* da parte del suo creatore per la comunicazione di una falsa dichiarazione su di un fatto, è possibile immaginare che quest'ultimo si difenda dichiarando di averlo inizialmente realizzato solamente per potenziare la diffusione di contenuti leciti su Internet e che solo successivamente lo avesse impiegato per la pubblicazione di uno o più contenuti ritenuti falsi – e, quindi, illeciti – dall'Autorità Competente. Di fronte ad un'ipotesi di questo tipo, è facile vedere come non sia possibile desumere la sussistenza dell'intenzione di creare un *bot* per il compimento di un reato dalla sua semplice realizzazione, data la sua natura strumentale che ne permette l'impiego sia per

[45] Si pensi, ad esempio, ad un *social bot* programmato per pubblicare su di un *social medium* i contenuti presenti all'interno di un determinato *file* di testo presente nel computer su cui anch'esso è installato. In questo caso il suo programmatore lo ha sì creato per pubblicare un ampio numero di contenuti ma questi possono essere facilmente cambiati da chiunque. In altre parole, questo particolare tipo di *bot*, che può essere un semplice programma *software* in formato *.exe*, potrebbe contenere nel proprio codice sorgente le istruzioni per cercare all'interno della cartella in cui è installato un documento di testo, come un *file* in formato *.doc* (il classico *file* di Word), e di pubblicare il contenuto di questo documento – qualunque esso sia – in un *post* sul *social medium* a cui può accedere. Una strutturazione del *social bot* in questi termini permetterebbe, quindi, a chiunque di utilizzarlo per le pubblicazioni più disparate per natura e liceità, sia a dei programmatori che a degli utenti meno esperti (i quali, sebbene privi di conoscenze approfondite in campo informatico, sarebbero ben in grado di modificare un documento di testo in formato *.doc* e di avviare un programma in formato *.exe*).

il compimento di attività lecite che illecite[46]. In altre parole, dato che questi programmi *software* autonomi non sono intrinsecamente atti ad offendere, perché possa considerarsi violata la Sezione 8 del *POFMA*, sembra necessario provare la sussistenza della volontà di utilizzarli con finalità (illecite) di disinformazione già durante il processo di scrittura del loro codice sorgente e non semplicemente farla derivare dall'utilizzo da parte del loro creatore per «comunicare a Singapore […] una falsa dichiarazione su di un fatto». Quindi, per valutare se il creatore di un *bot* abbia effettivamente violato la specifica Sezione 8 (1)(a), risulta necessario riferirsi ad elementi ulteriori rispetto alla semplice comunicazione di falsità tramite esso come, ad esempio, lo scarto temporale tra la sua creazione ed il suo impiego per questa illecita attività (un uso criminoso immediatamente successivo alla sua creazione è un valido indizio delle intenzioni del suo creatore) o la natura e la quantità di tutti contenuti diffusi con esso (la pubblicazione di numerosi contenuti leciti e di pochi – o solo uno – affetti da falsità su di un fatto ben potrebbero suggerire un intento produttivo non motivato da intenzioni prettamente disinformative).

Per quanto riguarda, invece, la condotta di cui alla Sezione 8 (1)(b), cioè la creazione o l'alterazione di un *bot* per consentire ad altri la comunicazione tramite esso di una falsa dichiarazione su di un fatto, le osservazioni compiute *supra* in merito al fatto che la comunicazione *online* di contenuti si basa sostanzialmente sull'interazione di un *bot* con elementi solitamente esterni al proprio codice sorgente acquistano una rilevanza ancora maggiore. Se infatti, la previsione di cui alla Sezione precedentemente analizzata contempla la condotta di un solo soggetto, in questo secondo caso assume rilievo anche la condotta di un soggetto diverso rispetto al creatore dello strumento di propagazione dei contenuti illeciti. Ciò deriva dal fatto che, per come è strutturata questa Sezione del *POFMA*, viene attribuito il compimento di un autonomo reato a chiunque crei un *bot* con l'intenzione di far sì che esso venga utilizzato da altri per il materiale compimento delle illecite attività di disinformazione. In altre parole, la legge di Singapore non attribuisce al creatore del *bot* la qualifica di "ausiliatore" al compimento del reato di comunicazione di false dichiarazioni su di un fatto, nonostante egli abbia volontariamente fornito un aiuto materiale determinante all'esecuzione del reato di cui alla Sezione 7[47], ma lo considera come l'autore di un reato distinto. La vertebrazione di una condotta illecita secondo questa formula ha senz'altro il vantaggio di ammantare la legge singaporiana di una maggiore forza preventiva, dato che viene così a formarsi

[46] Questa osservazione viene avanzata anche in J. Horder, *Online Free Speech and the Suppression of False Political Claims*, cit., 16, dove l'autore fa anche riferimento alla difesa del libero uso di *bot* fatta in M. Lamo – R. Calo, *Regulating Bot Speech*, in *UCLA Law Review*, 66, 2019, 988 ss.

[47] In merito alla figura del complice "ausiliatore" all'interno del nostro ordinamento si vedano, invece, G. Fiandaca – E. Musco, *Diritto penale parte generale*, Bologna, 2009, 504, nonché le considerazioni esposte in merito alla posizione del cosiddetto "palo" nel reato di rapina in Cass. pen., sez. II, 9 novembre 1971, n. 2076, in cui gli Ermellini hanno affermato che «il c.d. "palo" deve considerarsi un cooperatore immediato, in quanto egli, pur senza porre in essere l'intera attività tipica, è perfettamente consapevole dell'azione delittuosa che si va compiendo, e partecipa ad essa in maniera efficace e nel momento della perpetrazione, consentendo ai correi di realizzare gli atti materiali di esecuzione con maggiore tranquillità e di sfuggire al pericolo di sorprese di ogni genere», e le recenti precisazioni compiute in Cass. Pen., sez. V, 09 novembre 2021, n. 8973, dove è stato dichiarato che l'ausiliatore è «colui che si limita ad apportare un qualsiasi aiuto materiale nella preparazione o nella esecuzione del reato».

l'antigiuridicità di una condotta prima ancora della causazione di un danno (il quale si può comunque individuare, tramite il preambolo e la dichiarazione dello scopo della legge esplicitata alla Sezione 5, nonché le ipotesi contenute alla Sezione 8 (3), nell'ingenerazione nel destinatario della comunicazione di un'impressione erronea su eventi o qualità di altri) e dato anche che le pene previste per la creazione finalizzata all'illecito utilizzo di un *bot* sono particolarmente severe[48].

Tuttavia, impernare l'attribuzione del carattere dell'antigiuridicità ad una condotta di per sé non bollata come proibita sulla base della semplice volontà di «consentire a qualsiasi altra persona di comunicare a Singapore, tramite l'uso del *bot*, una falsa dichiarazione su di un fatto» rischia di risultare estremamente controversa o inefficace. Infatti, dato che, come *supra* evidenziato, i *bot* non sono intrinsecamente atti ad offendere – come implicitamente riconosciuto dallo stesso legislatore singaporiano – e che possono essere utilizzati per diffondere sia contenuti leciti che illeciti, diventa estremamente importante definire cosa si intenda per «intenzione di […] (b) consentire a qualsiasi altra persona di comunicare» delle affermazioni ingannevoli. Questo perché, se per «intenzione» di permettere a terzi di utilizzare il programma *software* autonomo realizzato, si volesse ricomprendere la consapevolezza che detto programma può essere utilizzato per la diffusione dei più disparati contenuti – anche illeciti – e la conseguente accettazione del rischio che il terzo ricevente il *bot* lo usi, anche solo in un momento successivo e ad anni di distanza, per pubblicare non soltanto contenuti leciti ma anche alcuni mistificatori o falsi, allora, la creazione di qualsiasi *bot* potrebbe essere compiuta in violazione della Sezione 8 del *POFMA*. Se, invece, si volesse dare una più stretta interpretazione del concetto di «intenzione» di permettere a terzi di utilizzare l'agente *software* prodotto, intendendo con esso la sola volontà di apportare un aiuto materiale alla realizzazione del conosciuto progetto delittuoso di comunicare false affermazioni su di un fatto pianificato da colui al quale ha ceduto il *bot* (assumendo, così, una posizione assimilabile a quella dell'ausiliatore di un reato prevista dalla nostra giurisprudenza)[49], allora, sarebbe necessario per l'Autorità Competente provare che il creatore del *bot* fosse a conoscenza del disegno criminoso di colui che ha poi materialmente compiuto la comunicazione che viola la Sezione 7 del *POFMA* e che fosse sua effettiva intenzione agevolarne o consentirne la realizzazione. Questa seconda interpretazione, sebbene possa essere considerata come maggiormente ragionevole, avrebbe, però, per l'Autorità Competente un alto grado di difficoltà pratica, non solo perché in generale la prova della conoscenza da parte del creatore del *bot* dei fini illeciti del comunicatore dei contenuti non è agevole ma soprattutto perché, dato che questo programma *software* autonomo può essere facilmente inviato o condiviso tramite Internet, il grado di

[48] La sezione 8 (2) prevede per il creatore del *bot* che l'abbia usato per comunicare o facilitare ad altri la comunicazione di false affermazioni su di un fatto una pena fino a 500.000 dollari di Singapore e di 3 anni di reclusione (anche in combinazione tra loro) mentre la Sezione 8 (3) dispone, nel caso di comunicazioni che potrebbero ledere degli interessi pubblici, fino ad un milione di dollari di Singapore e 6 anni di reclusione.

[49] Si veda quanto riferito alla nota 47, tenendo, però, ben presente che la struttura dei due ordinamenti differisce profondamente circa l'integrazione del reato di produzione di *bot* per fini illeciti considerato dalla Sezione qui commentata e l'attribuzione del ruolo di ausiliatore al compimento di reati attuabile sulla scorta del nostro art. 110 c.p.

anonimato raggiungibile dalle parti coinvolte nel suo scambio è tale da permettere la ragionevole esclusione della conoscibilità sia dell'identità che delle reali intenzioni del cessionario del programma. A ciò si aggiunga, poi, che sempre a causa della strutturazione di Internet è anche possibile che, nonostante l'identificazione di una delle due parti della cessione del *software*, risulti impraticabile risalire all'identità dell'altra, lasciando così impunito uno dei crimini contemplati dalla normativa singaporiana.
Inoltre, vi è da notare come nella Sezione 8 (1)(b) del *POFMA* è la creazione (o l'alterazione) del *bot* con la finalità di permettere ad altri di impiegarlo in violazione della legge di Singapore a costituire la condotta antigiuridica e non, invece, la sua cessione a qualsiasi titolo a questi. Questa previsione ha il vantaggio per l'Autorità Competente procedente di evitarle di dover provare il trasferimento del *software* prodotto, dato che, come visto, questo può essere facilmente condiviso su Internet con molti mezzi che ben potrebbero anche renderlo irriconoscibile agli occhi di un osservatore esterno[50]. Allo stesso tempo, però, in assenza della cessione la violazione di questa specifica Sezione della normativa sembra estremamente ardua da provare.
Da ultima, appare opportuna una considerazione circa l'ipotesi in cui una persona crei o modifichi un *bot* con l'intenzione di utilizzarlo per comunicare o consentire ad altri di comunicare «una falsa dichiarazione su di un fatto» ma che esso non funzioni a causa di un errore di scrittura del codice sorgente. Sebbene all'atto pratico un agente *software* così configurato sarebbe incapace di svolgere il compito prefissato e, di conseguenza, non sarebbe utile al compimento del disegno criminoso disinformativo desiderato, non si può ritenere la condotta creativa completamente priva di idoneità alla commissione del delitto di comunicazione di false dichiarazioni su di un fatto e, quindi, non punibile e che invece debba essere dato il giusto risalto all'entità dell'errore presente nel codice. Questa conclusione deriva dal fatto che la creazione di programmi *software* autonomi non è semplice ed anche il minimo errore nella scrittura del loro codice sorgente può comportare la compilazione di un *set* di istruzioni incomprensibili per un computer (basta inserire una sola volta una virgola anziché un punto e virgola perché una macchina sia incapace di far funzionare correttamente un programma il cui codice è formato da centinaia di righe). Fonte di una tale conclusione è anche il fatto che esistono diversi programmi chiamati *debugger* specificamente progettati per l'analisi e l'eliminazione degli errori di programmazione (i *bug*) con cui poter tentare di correggere gli errori di scrittura e rendere successivamente – spesso dopo variati tentativi – idonei i *bot* ad agire in violazione della normativa qui considerata.

[50] Non è difficile immaginare una situazione in cui il creatore del *bot* ne invii una copia criptata tramite Internet al destinatario e solo in un momento successivo (ad esempio, in seguito ad un pagamento compiuto tramite criptovalute) gli invii anche la chiave di decriptazione. In questo modo chiunque intercetti la comunicazione o controlli i terminali usati per l'invio della copia dell'agente *software* in assenza della chiave si troverebbe di fronte ad un cumulo di dati incomprensibile con cui non sarebbe in grado di incriminare nessuna delle parti. Un'altra ipotesi non irrealistica, data la dimensione contenuta di questo tipo di *software*, sarebbe quella in cui il creatore nasconda il codice sorgente in una o più immagini tramite le pratiche di steganografia e lo invii all'utilizzatore, rendendo anche in questo caso quasi impossibile per l'Autorità Competente individuare la transazione ed incriminare tutti i soggetti coinvolti.

(D) Aggravante specifica per l'uso di un *bot* per la comunicazione di false affermazioni su di un fatto a Singapore – Come accennato *supra* all'inizio del paragrafo, il fine principale del *POFMA* è quello di «impedire la comunicazione elettronica a Singapore di false dichiarazioni su di un fatto» e di «sopprimere il supporto a e contrastare gli effetti di tale comunicazione». In base a questo obiettivo, la normativa statuisce che: «Una persona non deve compiere alcun atto all'interno o all'esterno di Singapore per comunicare a Singapore una affermazione sapendo o avendo motivo di credere che (a) si tratti di una falsa affermazione su di fatto; e (b) la comunicazione dell'affermazione a Singapore è probabile che (i) sia di pregiudizio per la sicurezza di Singapore o di qualsiasi parte di Singapore; (ii) sia di pregiudizio per la salute pubblica, la sicurezza pubblica, la tranquillità pubblica o le finanze pubbliche; (iii) pregiudicare le relazioni amichevoli di Singapore con altri paesi; (iv) influenzare l'esito di un'elezione alla carica di Presidente, un'elezione generale dei membri del Parlamento, un'elezione suppletiva di un membro del Parlamento o un referendum; (v) incitare sentimenti di inimicizia, odio o ostilità tra diversi gruppi di persone; oppure (vi) diminuire la fiducia del pubblico nell'adempimento di qualsiasi dovere o funzione o nell'esercizio di qualsiasi potere da parte del Governo, di un Organo di Stato, di un consiglio statutario o di una parte del Governo, di un Organo di Stato o di un consiglio statutario»[51].

Alla determinazione dell'antigiuridicità delle condotte appena elencate la Sezione 7 (3) aggiunge il riconoscimento come aggravamento del loro disvalore l'utilizzo di un «*bot*» per «(a) comunicare a Singapore la dichiarazione menzionata alla sottosezione (1); e (b) allo scopo di accelerare tale comunicazione».

Questa previsione di un'aggravante specifica per l'uso dei *bot* riconosce implicitamente la loro natura strumentale nel compimento di atti illeciti e, soprattutto, riconosce come essi siano estremamente utili per potenziare la diffusione di contenuti ritenuti dall'Autorità Competente come falsi e potenzialmente dannosi per l'ordine pubblico di Singapore o la sicurezza dei suoi cittadini[52]. Infatti, sebbene la sintassi impiegata alla Sezione 7 (3) non sia delle più felici, è possibile ricavare la loro identificazione come mezzi «per accelerare» la propagazione di comunicazioni viste come pregiudizievoli per i beni giuridici propri della collettività elencati dalla normativa stessa (Sezione 7 (1)(b) da (i) a (vi)). Inoltre, data la previsione di massimi edittali più alti rispetto alla semplice comunicazione di «una falsa affermazioni su di un fatto», pare evidente che il legislatore dia per scontato che il loro impiego permetta una lesione più grave. Ciò anche in ragione del fatto che, come già *supra* rilevato, questi agenti *software* possono facilmente essere – e spesso sono – impiegati in squadre o, comunque, può essere loro affidata la gestione di molteplici profili *online* al fine di dare ai meno esperti l'illusione che sia un vasto pubblico di utenti a condividere e a seguire le opinioni comunicate, in realtà, da uno o pochi soggetti che tirano i fili di questa compagnia di marionette digitali[53].

[51] Sezione 7 (1) del *POFMA*.

[52] Si tenga ben presente che all'interno della normativa qui considerata, per quanto riguarda il verbo "comunicare" (in originale «*communicate*»), la Sezione 3 (1) del *POFMA* specifica che «una dichiarazione o [un] contenuto viene comunicato a Singapore se è reso disponibile a uno o più utenti finali a Singapore su o tramite Internet».

[53] Si veda al proposito quanto riferito dagli autori citati *supra* alle note 11, 12, 14 e 19.

La particolare gravità con cui viene sanzionato l'impiego di *bot* per la comunicazione di false dichiarazioni su di un fatto che potrebbero compromettere l'armonia sociale di Singapore (la Sezione 7 (3) prevede l'irrogazione di pene con un massimo edittale di 100.000 dollari di Singapore e di 10 anni di carcere ovvero «in ogni altro caso» fino ad un milione di dollari di Singapore) trova la base della propria giustificazione nella considerazione che oggi anche le cerchie virtuali di connessioni giocano un ruolo importante nella formazione delle convinzioni personali e delle opinioni politiche, che già da tempo vien fatto un massiccio impiego dissimulato di questi programmi *software* autonomi per la diffusione di contenuti *online* che vengano percepiti dagli utenti umani come espressione di un genuino e condivisibile sentimento popolare[54] e che proprio questa suggestione di apprezzamento diffuso può avere l'effetto di sviare artificiosamente lo sviluppo delle convinzioni personali, come già messo in luce da alcuni celebri studi[55].

In ultima analisi, la specifica previsione di una circostanza aggravante del reato di cui alla Sezione 7 (1) integrata dall'uso dei *bot*, mostra – almeno agli occhi di scrive – che il legislatore di Singapore abbia compreso la loro natura di moltiplicatori, anzi, di elevatori alla potenza[56] dell'estensione della diffusione *online* di contenuti (in particolare sui *social media*) e, quindi, della potenziale portata offensiva delle condotte disinformative considerate all'interno della normativa qui analizzata.

(E) Provvedimenti di restrizione dell'uso di un *account* gestito da un *bot* – La previsione della Sezione 40 di attribuire all'Autorità Competente il potere di adottare i provvedimenti di restrizione dell'uso dell'*account* (chiamati "*Account Restriction Directions*") che essa abbia riconosciuto come gestito da un *bot* trasmette – almeno a chi scrive – l'impressione che il legislatore di Singapore abbia sottovalutato una delle caratteristiche fondamentali di questo genere di programmi *software* automatizzati: il fatto che essi vengono progettati per funzionare su di una determinata piattaforma e non per gestire un determinato profilo. Infatti, per compiere un accesso ad una qualsiasi infrastruttura informatica, un *bot* si limita ad utilizzare delle credenziali di accesso, ossia una combinazione di termini quali *username* (oppure indirizzo *e-mail*) e *password*, e ad in-

[54] J. Ratkiewicz – M. Conover – M. Meiss – B. Gonçalves – A. Flammini – F. Menczer, *Detecting and tracking political abuse in social media*, in *Fifth international AAAI conference on weblogs and social media*, 5(1), 2011, 297 ss.; 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.; 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 (CSCW '15)*, 2015, 839 ss.; S.C. Woolley, *Automating power: Social bot interference in global politics*, cit.

[55] Si vedano le note 15, 16 e 17.

[56] L'avanzamento della definizione dei *bot* come "elevatori a potenza delle diffusione di contenuti *online*" deriva dal fatto che Internet ed in particolare i *social media* sono già stati definiti come dei «moltiplicatori di discorsi d'odio (*hate speech*) e di notizie false (*fake news*)» in 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, dove l'autrice fa anche riferimento a C.R. Sunstein, *Republic. com. Cittadini informati o consumatori di informazioni?*, Bologna, 2003, 82 ss.; C.R. Sunstein, *Voci, gossip e false dicerie*, Milano, 2010; G. Pitruzzella – O. Pollicino – S. Quintarelli, *Parole e potere. Libertà di espressione, hate speech e fake news*, Milano, 2017; Aa.Vv., in questa *Rivista*, 1, 2017, spec. 11-115.

terfacciarsi con un ambiente standardizzato e per lui ben riconoscibile. Di conseguenza, in caso di applicazione di restrizioni all'*account* gestito da un *bot*, nulla impedisce al suo "padrone" di fornirgli le credenziali di accesso ad un altro profilo[57], rendendo nei fatti inefficace il provvedimento che vorrebbe impedire la prosecuzione dell'opera di diffusione *online* potenziata di contenuti mendaci.

Una soluzione (di poco) più efficace e fin troppo lapidaria sarebbe stata, quindi, quella di imporre delle restrizioni non tanto all'uso dell'*account* utilizzato per la diffusione automatizzata di contenuti illeciti quanto alla possibilità di accedere ad Internet all'indirizzo IP collegato all'*account* incriminato o, per quanto possibile, al soggetto beneficiario dei servizi di intermediazione di Internet che sia stato individuato come titolare del profilo incriminato[58]. Tuttavia, questa ipotesi è stata esclusa dal *POFMA*, come si può desumere dalla lettura combinata delle Sezioni 2 e 40 (1). Infatti, dalla quarantesima Sezione è possibile ricavare che, in caso di condotte criminose compiute con l'ausilio di un «*account online* non autentico» («*inauthentic online account*») o di un *bot* – ipotesi espressamente considerate alla Sezione 40 (2)(c) – a detto profilo il fornitore dei servizi di intermediazione di Internet (in originale «*Internet intermediary*») debba, su provvedimento motivato (ossia una «*direction*») dell'Autorità Competente, impedire che esso possa usufruire dei suoi servizi per «comunicare qualsiasi affermazione a Singapore» (Sezione 40 (1)(a)) e che debba impedire a chiunque di poter utilizzare l'*account* per «interagire con qualsiasi utente finale» dei suoi servizi (Sezione 40 (1)(b)). Dalla lettura della Sezione 2, poi, si ricava che per «fornitore dei servizi di intermediazione di Internet» si intende colui che permette ai propri utenti di «accedere a contenuti provenienti da terzi su o tramite Internet», di trasmettere questi contenuti su o tramite Internet e di fare ricerche *online* e di consultare i contenuti trovati[59] mentre da questa categoria ne viene esplicitamente escluso il fornitore di accesso ad Internet e di servizi di «incremento della capacità di elaborazione o la capacità di archiviazione di un computer», ossia l'*Internet Service Provider* che consente la connessione al *web* o, ad esempio, il fornitore di servizi di *cloud computing*[60]. In altre parole, le restrizioni che l'Autorità Competente può richiedere

[57] Questa operazione può risultare particolarmente facile e veloce. Chi scrive, infatti, è stato capace di creare un *social bot* – invero molto rudimentale – che per compiere l'accesso ad un *social medium* era programmato per cercare delle credenziali di accesso in un file esterno in formato *.txt*, permettendo così al suo gestore di cambiare i dati per l'accesso senza dover intervenire sul codice sorgente (operazione comunque non eccessivamente complicata) ma, semplicemente, modificando un file di testo (consultato poi in automatico dal programma).

[58] In merito alla maggiore rilevanza dell'indirizzo IP rispetto al semplice *nickname* di un profilo *online* si vedano le considerazioni della nostra giurisprudenza espresse in Cass. Pen., sez. V, 22 novembre 2017, n. 5352, e in Cass. Pen., sez. V, 05 febbraio 2018, n. 5352.

[59] Il *POFMA* stesso fornisce alla Sezione 2 un esempio di quali siano i «servizi di intermediazione di Internet» che identificano i soggetti che devono, su richiesta dell'Autorità Competente, restringere l'uso di un *account*. Esso contempla: servizi di *social networking*, servizi di ricerca *online*, servizi di aggregazione di contenuti, servizi di messaggistica basati su Internet e servizi di condivisione di video. Ne deriva, quindi, che le società che gestiscono piattaforme come, rispettivamente, Facebook, Google, Flipboard, WhatsApp e YouTube possono essere i destinatari di provvedimenti di restrizione dell'uso dei profili dei loro utenti.

[60] Dall'interpretazione delle definizioni contenute alla Sezione 2 – in particolare, della definizione di *Internet intermediary service* – è, quindi, possibile escludere dal novero degli *Internet intermediary* società quali, ad esempio, la AT&T e la Amazon Web Services.

in base alla Sezione 40 (1) del *POFMA* all'*Internet intermediary* si limitano a colpire l'accessibilità di un *account* ad alcune funzionalità di Internet e l'accessibilità a detto profilo per interagire con altri, lasciando così spazio alla creazione di nuovi profili *online* con cui continuare l'opera di disinformazione.

Rimane, comunque, da notare come la chiusura – *rectius* la limitazione delle attività in rete – di un *account* che fosse effettivamente gestito da un *bot* avrebbe il vantaggio di rallentare l'attività di diffusione di contenuti da questo tipo di programma *software*, dato che gli potrebbe poi essere assegnato solamente un profilo privo di quella rete di contatti sociali (il *social network*, appunto) che rendeva efficace la diffusione automatizzata di contenuti *online* e che sarebbe, quindi, necessario parecchio tempo per ricostituirla o per crearne un'altra delle medesime dimensioni.

Infine, pare qui opportuno sottolineare come dalla lettura delle condizioni elencate alla Sezione 40 (2) necessarie perché l'Autorità Competente possa adottare un provvedimento di restrizione dell'uso di un *account* emerga che il legislatore di Singapore non abbia considerato illecito il semplice controllo di un profilo *online* da parte di un *bot* ma, invece, che perché l'Autorità abbia il potere di intervenire sia necessario il soddisfacimento congiunto di molteplici condizioni (tra cui l'effettivo compimento di un illecito tramite esso).

Come visto, oltre alle numerose critiche di facilità d'impiego della normativa della Repubblica di Singapore qui brevemente analizzata come strumento del governo per ridurre al silenzio il dissenso e le critiche[61], sono riscontrabili anche diverse problematiche relative allo specifico ambito della regolamentazione dell'uso di programmi *software* autonomi per la diffusione automatizzata e potenziata di contenuti *online*. Infatti, sebbene dalla lettura del *POFMA* traspaia la comprensione della natura dei *bot* come neutra, strumentale e di elevatori alla potenza della diffusività di contenuti sul *web*, le soluzioni accolte lasciano, comunque, aperte delle "porte sul retro" (per riprendere il termine "*backdoor*", utilizzato in ambito informatico per indicare dei metodi e degli accorgimenti utili ad aggirare le misure di sicurezza di un sistema o di un'infrastruttura) che rischiano di inficiare l'efficacia della sua applicazione. In ultima analisi, anche se la normativa della Città dei Leoni contiene diverse criticità nel suo tentativo di arginare l'utilizzo dei *bot* come strumento di propagazione potenziata di contenuti mistificatori e dannosi ed evidenzi l'orientamento «autoritario e paternalistico» del legislatore singaporiano[62], appaiono, comunque, calzanti le parole di Jones, per la quale «il diritto alla libertà di espressione non implica che le tecniche di manipolazione dell'attenzione, come l'uso di *bot* e *troll*, debbano essere prive di restrizioni»[63].

[61] Si vedano le critiche *supra* menzionate alle note 26, 27, 28 e 29.

[62] A. Carson – L. Fallon, *Fighting fake news: a study of online misinformation regulation in the Asia Pacific*, cit., 15.

[63] K. Jones, *Online Disinformation and Political Discourse: Applying a Human Rights Framework*, in *chathamhouse.org*, novembre 2019, 45 e 56.

2.2. La *Online Advertising and Social Media (Transparency) Bill* 2017 della Repubblica d'Irlanda

Nell'ambito normativo europeo stiamo assistendo da alcuni anni – soprattutto dopo le interferenze nel dibattito *online* sull'evento referendario britannico del 2016[64] – ad una crescente attenzione ai fenomeni dei comportamenti coordinati ed inautentici sul *web* e, soprattutto, sui *social media*. L'interesse dei legislatori del Vecchio Continente si è, così, tradotto in diverse iniziative di riforma delle normative nazionali e comunitarie[65]. Tra le prime pare qui interessante prendere in considerazione quella avanzata al Dáil Éireann (la camera bassa del Parlamento) della Repubblica d'Irlanda, ossia la *Online Advertising and Social Media (Transparency) Bill 2017*[66].
Anche la proposta legislativa attualmente dibattuta alla camera bassa irlandese ha il fine di affrontare il più ampio problema dell'opacità nelle pratiche di pubblicità politica *online* e, come la normativa della Repubblica di Singapore appena analizzata, contiene solo alcune disposizioni contro l'impiego malevolo dei *bot*. A differenza, però, della normativa dell'Isola dei Leoni, la *OASM Bill* ritiene illecito l'uso di questi agenti *software* solo quando sia caratterizzato da finalità politiche.

(A) Definizione di *bot* – Alla Sezione 2 (1) della *OASM Bill* viene data la definizione di «*bot*» come di un: «Qualsiasi elemento di un *software* che è programmato per eseguire compiti automatizzati su una piattaforma *online*».
Viene poi specificato che nella definizione di «piattaforma *online*» rilevante per la legge viene ricompreso: «Qualsiasi sito *web*, applicazione *web* o applicazione digitale (che include un *social network* o un motore di ricerca) che ha 10.000 o più visitatori unici visitatori o utenti unici mensili nello Stato in almeno sei dei dodici mesi precedenti; ovvero, quando la piattaforma *online* è rivolta a un pubblico locale o regionale, che ha 1.000 o più visitatori o utenti unici mensili nello Stato in almeno sei dei precedenti dodici mesi»[67].
Preliminarmente, si può notare come una delineazione in questi termini dei «*bot*» riecheggi le nozioni di "*bot*" e di "*internet bot*" (detti anche "*web bot*") avanzate da Tsvetkova, Garcia-Gavilanes, Floridi e Yasseri[68], riprendendo, per la prima, le analisi compiute

[64] Si veda quanto detto *supra* al paragrafo 1 e, in particolare, alle note 5, 7 e 8.

[65] In merito alla normativa dell'Unione si veda la recente proposta di regolamento *relativo a un mercato unico dei servizi digitali* (COM(2020) 825), nella quale, ai Paragrafi 57 e 68, vengono espressamente presi in considerazione i *bot* come strumenti «che possono condurre alla rapida e ampia diffusione di informazioni che costituiscono contenuti illegali» (Paragrafo 57). Si veda anche lo *Strengthened Code of Practice on Disinformation 2022*, un codice di autoregolamentazione per i gestori delle piattaforme online promosso dalla Commissione europea.

[66] Il testo della proposta di legge presentata il 06 dicembre 2017 al Dáil Éireann è liberamente consultabile sul sito istituzionale dell'Oireachtas (il parlamento irlandese) in *Online Advertising and Social Media (Transparency) Bill 2017*, in *oireachtas.ie*, 6 dicembre 2017.

[67] Sezione 2 (1) della *OASM Bill*.

[68] M. Tsvetkova – R. García-Gavilanes – L. Floridi – T. Yasseri, *Even good bots fight: The case of Wikipedia*, in *PLoS ONE*, 12(2), 2017, spec. 1-2.

da Franklin e Graesser[69] e, per la seconda, anche quelle di Berners-Lee e Cailliau[70], Koster[71] e Cheong[72]. Infatti, parafrasando le parole di Tsvetkova, Garcia-Gavilanes, Floridi e Yasseri, è possibile descrivere un *internet bot* come un programma *software* che viene eseguito in continuazione e che è in grado – senza alcun intervento umano – di attivarsi al verificarsi di determinate condizioni e di individuare le azioni da compiere per portare a termini i compiti affidatigli, avendo consapevolezza dell'ambiente informatico *online* in cui opera ed adattandosi autonomamente ai cambiamenti di questo.
Per quanto riguarda la definizione di «piattaforma *online*» contenuta sempre nella Sezione 2 (1) si può notare come il legislatore irlandese abbia ripreso ampiamente la definizione dei medesimi termini contenuti nel Paragrafo 17940 del *California Business and Professions Code* (introdotto nel settembre del 2018 con la promulgazione del *Bolstering Online Transparency Act of 2018* (S.B. 1001))[73]. Infatti, la normativa d'oltreoceano identifica come «piattaforma *online*» un: «Qualsiasi sito *web*, applicazione *web* o applicazione digitale rivolta al pubblico, inclusi un *social network* o una pubblicazione, che abbia 10.000.000 o più di visitatori o utenti mensili unici negli Stati Uniti per la maggior parte dei mesi durante i 12 mesi precedenti»[74].
Inoltre, è possibile notare come la proposta normativa irlandese si differenzi da quella singaporiana per quanto riguarda l'identificazione delle strutture entro cui sono in grado di operare i *bot*, pur finendo con l'ottenere un risultato estremamente simile. Infatti, mentre la normativa dell'Isola dei Leoni vieta espressamente l'uso di SMS e MMS per l'illecita diffusione di «una falsa dichiarazione su di un fatto»,[75] quella dell'Isola di Smeraldo accoglie una definizione di «piattaforma *online*» che, grazie anche all'inclusione di «qualsiasi […] applicazione digitale», risulta sufficientemente generica ed aperta da annoverare non solo le applicazioni basate sull'uso delle reti GSM e UMTS (le prime a permettere l'invio di SMS e MMS) ma anche altre tecnologie digitali tra quelle illecitamente utilizzabili per far sì che un profilo gestito da un «*bot*» appaia come quello di un individuo e diffonda messaggi elettorali o politici.

(B) Reato di utilizzo di un *bot* per realizzare presenze multiple *online* con il fine politico di presentarsi come un profilo individuale – È la Sezione 6 della *OASM Bill* a determinare le condizioni in base alle quali l'utilizzo di un *bot* integri un illecito (una «*offence*»). In essa, intitolata «reato di utilizzo di un *bot* per realizzare presenze multiple *online* con il fine politico di presentarsi come un *account* o un profilo individuale su di una piattaforma *online*», viene stabilito che: «Chiunque utilizzi consapevolmente

[69] 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.

[70] T. Berners-Lee – R. Cailliau, *The world wide web project*, in *World Wide Web*, 1991.

[71] M. Koster, *Guidelines for robot writers*, in *info.webcrawler.com*, 1993.

[72] F.C. Cheong, *Internet agents: spiders, wanderers, brokers, and bots*, Thousand Oaks (USA), 1993.

[73] Per un'approfondita disamina della normativa statale di contrasto delle attività dei *bot* per finalità politiche della California si veda A. Tedeschi Toschi – G. Berni Ferretti, *Il contrasto legislativo ai socialbot e le soluzioni avanzate negli Stati Uniti d'America a livello federale e statale*, cit., spec. paragrafo 4.4.

[74] Paragrafo 17940 (c) del *California Business and Professions Code.*

[75] Sezione 3 (2)(b) del *POFMA.*

un *bot*, o lo faccia utilizzare, in modo tale da far sì che più presenze *online* con un fine politico si presentino a un utente di una piattaforma *online* come un *account* o un profilo individuale su qualsiasi piattaforma *online* è colpevole di un reato»[76].

Preliminarmente, è necessario rivolgere l'attenzione al significato che viene dato ai termini «fine politico» (in originale «*towards a political end*») all'interno della normativa in oggetto. Questo perché esso riveste un'importanza cruciale nel determinare quando l'uso di un *internet bot* non sia considerato illecito e quando, invece, lo sia, dato che anche nel caso di questo legislatore pare – almeno a chi scrive – essere stata fatta propria la considerazione che i *bot* non sono intrinsecamente atti ad offendere (come implicitamente riconosciuto pure da quello di Singapore)[77] e che possono, quindi, essere utilizzati per diffondere sia contenuti leciti che illeciti. Infatti, viene poi precisato che una «questione» (in originale «*matter*») è considerata come «avente un fine politico» (in originale «*directed towards a political end*») solo: «Se comunica un messaggio che promuove uno o più candidati a cariche pubbliche in un'elezione all'interno dello Stato o un partito politico registrato nel registro dei partiti politici, o se promuove un messaggio su una questione di interesse o importanza politica o una questione che, al momento in cui la pubblicità *online* è posta o promossa, è in discussione o che si intende presentare a una delle Camere dell'Oireachtas o dell'Assemblea dell'Irlanda del Nord o del Parlamento europeo o di un'autorità locale all'interno dello Stato o che sia oggetto di un referendum o che ha una qualsiasi relazione con una controversia industriale in corso nello Stato»[78].

Dalle parole della Sezione 6 si può, quindi, ricavare che la *OASM Bill*, al pari del *BOT Act* californiano e del *POFMA* singaporiano, non istituirebbe un divieto *tout-court* dell'uso di *bot* ma si limiterebbe, invece, a vietare determinate modalità di impiego di essi. Rispetto alla normativa della piccola repubblica del sud-est asiatico, poi, la proposta di legge presentata al Dáil Éireann prevede un ambito di illiceità dell'utilizzo di questi agenti *software* ancora più ristretto, dato che quest'ultima stabilisce la rilevanza penale del loro impiego solo quando sia dissimulato, avvicinandola di più all'ordinamento della Stato d'Oro (che espressamente dichiara illegale fuorviare un'altra persona sulla propria identità artificiale)[79].

In aggiunta, questa proposta di legge risulta avere un ambito di applicazione ancora più ristretto di quello della sua (ormai approvata) controparte di Singapore, in quanto riconosce come reato non il semplice utilizzo di un *internet bot* per la gestione di un profilo ma, invece, ritiene penalmente rilevante il suo utilizzo solo per realizzare «più presenze *online*» (in originale «*multiple online presences*»), le quali vengono indicate nella Sezione 2 (1) in un numero minimo pari a 25 *account* o profili[80]. In altre parole, l'uti-

[76] Il testo in lingua originale della Sezione 6 (1) della *OASM Bill* recita «*Any person who knowingly uses a bot, or causes a bot to be used, in such a way as to cause multiple online presences that are directed towards a political end to present to a user of an online platform as an individual account or profile on any online platform shall be guilty of an offence*».

[77] Si veda in merito quanto detto *supra* al paragrafo 2.1.(C) e, in particolare, quanto riportato alla nota 46.

[78] Sezione 2 (2) della *OASM Bill*.

[79] Paragrafo 17941 (a) del *California Business and Professions Code*.

[80] Come rimarcato anche dal promotore della *OASM Bill*, il deputato James Lawless, nel corso del

lizzo dissimulato di un programma *software* autonomo verrebbe qui reso un illecito (e, quindi, rappresentante una minaccia per la società irlandese) solamente quando esso renda l'impressione che un intero gruppo di persone condivida un messaggio politico «per disinformare o ingannare il pubblico in modo malevolo»[81]. La costituzione della normativa in questi termini mostra come il legislatore irlandese abbia ben compreso come sia facile fornire ad un *bot* le credenziali di accesso ad uno o più profili *online*[82] e farglieli gestire in autonomia e come essi possano rappresentare una seria minaccia non singolarmente ma quando, invece, operano in modo coordinato e mascherato nella gestione di ampi gruppi di *account*[83].

Sempre nel paragone con la legge dell'Isola dei Leoni, poi, vi è da notare come in merito ai contenuti la cui comunicazione costituisce un delitto la *OASM Bill* sia caratterizzata da una maggior ampiezza. Mentre, infatti, la prima ritiene antigiuridica la diffusione di «false affermazioni su di un fatto» che si sappiano o si suppongano poter essere di pregiudizio per l'ordine pubblico di Singapore e per i suoi cittadini[84], per la seconda è sufficiente che essa promuova «uno o più candidati a una carica pubblica in un'elezione […] o un partito politico registrato»[85] oppure «un messaggio su una questione di interesse o importanza politica»[86]. In altre parole, perché si possa considerare violata la Sezione 6 della proposta di legge irlandese è sufficiente usare un *internet bot* per costituire una "presenza *online* con un fine politico", ossia impiegarlo per fargli gestire in forma dissimulata venticinque o più profili *online* con cui diffondere contenuti propagandistici in favore di determinati soggetti o posizioni politiche (dando, quindi, la falsa impressione che questi siano più popolari di quanto non siano in realtà)[87]. La

dibattito del 13 dicembre 2017, disponibile in *Dáil Éireann debate*, in *oireacthtais.ie*, 13 dicembre 2017.

81 *Ibid.*

82 Si veda a questo riguardo quanto descritto *supra* alla nota 57.

83 Questa considerazione viene anche espressa nella Sezione 2 del *Bot Disclosure and Accountability Act of 2018* (S. 3127), presentato al Senato degli Stati Uniti d'America e in cui viene affermato che i *bot* vengono impiegati «per creare un effetto carrozzone, per costruire false tendenze sui *social media* diffondendo automaticamente gli *hashtag*, e anche per sopprimere le opinioni dell'opposizione». Per un'analisi del testo di questa proposta di legge federale americana si veda A. Tedeschi Toschi – G. Berni Ferretti, *Il contrasto legislativo ai socialbot e le soluzioni avanzate negli Stati Uniti d'America a livello federale e statale*, cit.

84 Sezione 7 (1)(b), da (i) a (vi), del *Protection from Online Falsehoods and Manipulation Act 2019*, la quale ha comunque suscitato critiche in merito al fatto che la tutela dell'ordine pubblico di Singapore potrebbe facilmente essere usata come scusa dal governo per ridurre al silenzio il dissenso e le critiche alle autorità (si vedano le critiche *supra* menzionate alle note 26, 27, 28 e 29).

85 Sezione 2 (2) della *OASM Bill.*

86 *Ibid.*

87 In merito alle analisi condotte sull'impiego dei *bot* per distorcere le discussioni *online* su determinati temi, stimolare i seguaci di ideologie o personaggi pubblici e per generare false impressioni di popolarità si vedano E. Mustafaraj – P.T. Metaxas, *From obscurity to prominence in minutes: Political speech and real-time search*, in *Proceedings of the WebSci10*, 2010; J. Ratkiewicz – M. Conover – M. Meiss – B. Gonçalves – A. Flammini – F. Menczer, *Detecting and tracking political abuse in social media*, cit.; Z. Chu – S. Gianvecchio – H. Wang – S. Jajodia, *Detecting automation of Twitter accounts: Are you a human, bot, or cyborg?*, in *IEEE Transactions on Dependable and Secure Computing*, 9(6), 2012, 811 ss.; C. Wagner – S. Mitter – C. Körner – M. Strohmaier, *When social bots attack: Modeling susceptibility of users in online social networks*, in *Proceedings of the WWW '12 Workshop on 'Making sense of microposts'*, 2012, 41 ss.; A. Bessi – E. Ferrara, *Social bots distort the 2016 US Presidential election online discussion*, in *First Monday*, 21(11), 2016; P.N. Howard – B. Kollanyi, *Bots, #Strongerin, and #Brexit: Computational Propaganda During the UK-EU Referendum*, cit.; S.C. Woolley,

strutturazione della normativa dell'Isola di Smeraldo secondo questi termini, invero abbastanza simile a quella dello Stato della California[88], pone, però, tre problemi di puntualizzazione: quando si possa parlare (i.) di comunicazione di un messaggio, (ii.) di contenuti che effettivamente promuovono un candidato od un partito e (iii.) di questione di interesse o di importanza politica.

(i.) Modalità di comunicazione di un messaggio – Per quanto riguarda il primo dei problemi di puntualizzazione congeniti nella *OASM Bill*, bisogna partire dalla considerazione che su Internet il concetto di "comunicazione" da applicare alla definizione di cui alla Sezione 2 (2) può includere una varietà più o meno ampia di azioni, a seconda delle conoscenze e dell'opinione di chi si potrebbe trovare a dover interpretare questa disposizione. Infatti, se da un lato è evidente che azioni compiute da un *bot* come l'invio di messaggi su di un'applicazione di messaggistica *online* ad un determinato destinatario[89], l'invio di una *e-mail* o la pubblicazione di un commento ai contenuti condivisi da altri utenti su di un *social medium* costituiscano delle comunicazioni – in particolare la prima – o comunque delle interazioni con una persona rilevanti per la Sezione 6 della proposta di legge irlandese, altre tipologie di azioni, come la pubblicazione di contenuti sulle piattaforme *social* o la ri-condivisione di contenuti di altri possono far sorgere dei problemi interpretativi[90]. Sembra, però, possibile attribuire anche a queste ultime condotte il concetto di "comunicazione", dato che i contenuti pubblicati all'interno dello "spazio personale" *online* di un profilo *social* (come, ad esempio, il "*diario personale*" di Facebook) vengono solitamente mostrati in automatico dal sistema della piattaforma ai profili con i quali sono stati stretti dei collegamenti o che più semplicemente lo "seguono", permettendo, quindi, una forma di comunicazione con un numero indeterminato di destinatari[91].

Automating power: Social bot interference in global politics, cit.; M.T. Bastos – D. Mercea, *The Brexit botnet and user-generated hyper-partisan news*, in *Social Science Computer Review*, 37(1), 2019, 38 ss.; J.C.M. Serrano – M. Shahrezaye – O. Papakyriakopoulos – S. Hegelich, *The Rise of Germany's AfD: A Social Media Analysis*, in *Proceedings of the 10th International Conference on Social Media and Society*, 2019, 214 ss.; F. Giglietto – N. Righetti – G. Marino, *Understanding coordinated and inauthentic link sharing behavior on Facebook in the run-up to 2018 general election and 2019 European election in Italy*, cit. La considerazione della pericolosità dell'impiego di questi agenti *software* per incrementare artificiosamente la popolarità di contenuti *online* è stata esposta anche dal deputato James Lawless nel corso di un dibattito contestuale alla presentazione della *OASM Bill*, avvenuto il 6 dicembre 2017. La trascrizione del dibattito è disponibile in *Dáil Éireann debate*, cit.

88 Si veda al riguardo quanto esposto in A. Tedeschi Toschi – G. Berni Ferretti, *Il contrasto legislativo ai socialbot e le soluzioni avanzate negli Stati Uniti d'America a livello federale e statale*, cit., spec. 472-478.

89 Si veda in merito a questo tipo di interazioni compiute da un *bot* il recente caso esposto in T. Rasmussen, *There was a tinder election bot fanning the fire of the youth vote*, in *i-d.vice.com*, 15 giugno 2017, in cui è stato riportato che un *bot* venne usato per spingere gli utenti dell'applicazione di incontri Tinder a registrarsi per poter votare alle elezioni per il rinnovo della Camera dei Comuni del Regno Unito del 2017. Stime approssimative riportarono che il Tinder Election Bot abbia inviato tra i 30.000 ed i 40.000 messaggi a persone di età compresa tra i 18 ed i 25 anni. Durante questa votazione si assistette alla più alta percentuale di voti espressi dai giovani durante le elezioni in Inghilterra dal 1992: il 66,4% degli elettori della fascia d'età considerata (non vi fu, però, modo di individuare alcun tipo di correlazione tra l'uso di questo *bot* per scambiare messaggi con gli utenti dell'applicazione e il numero di elettori giovani presentatisi ai seggi).

90 In merito alle critiche circa la vaghezza della terminologia impiegata nella normativa californiana si vedano R. Diresta, *A New Law Makes Bots Identify Themselves–That's the Problem*, in *wired.com*, 24 luglio 2019; J. Horder, *Online Free Speech and the Suppression of False Political Claims*, cit., 15-16.

91 A favore di questo orientamento interpretativo di inclusione della maggior parte delle condotte

L'interpretazione della norma rischia, invece, di risultare più complessa per quanto riguarda l'inquadramento dell'apposizione delle cosiddette "reazioni" ai contenuti pubblicati da altri (ad esempio i «Mi piace» di Facebook, gli «*Upvote*» di Reddit, o i «Consiglia» di LinkedIn). Riguardo a queste forme di azione nei confronti dei contenuti pubblicati da altri profili, è opinione di chi scrive che esse debbano rientrare nel novero delle comunicazioni rilevanti per la normativa in oggetto, dato che con esse si è in grado di comunicare (in forma più o meno ampia, a seconda della strutturazione della piattaforma su cui esse vengono impiegate) un'opinione riguardo a quanto detto e a farla sapere non solo al suo creatore ma anche alla platea di persone che abbia modo di visionarla, soprattutto perché – come già detto riguardo ai contenuti pubblicati all'interno dello "spazio personale" *online* di un profilo *social* – i sistemi dei *social media* solitamente mostrano ai profili di terzi che sono ad essi collegati o che li seguono i contenuti pubblicati od oggetto di "reazioni"[92].

(ii.) Contenuti che promuovono un candidato o un partito – Ulteriore nodo venutosi a creare con la strutturazione della proposta di legge irlandese nei suoi termini attuali riguarda la definizione di quali contenuti possano essere ritenuti di promozione per un candidato, un partito o un'opinione su di una questione di interesse o importanza politica tramite la loro diffusione in rete potenziata dai *web bot*.

Innanzitutto, è importante evidenziare come la normativa sia strutturata in modo da sanzionare come illecito la mera comunicazione di contenuti promuoventi soggetti, organizzazioni politiche registrate od opinioni, senza che abbia rilevanza l'effettiva realizzazione di una qualche forma di ascendenza sui destinatari. Di conseguenza, nemmeno le forme più puerili di supporto ad un certo candidato o ad un certo partito politico possono dirsi estranee all'ambito di applicazione della *OASM Bill*, in quanto, a prescindere dalla loro reale efficacia, se messe in atto da un profilo gestito da un *bot*, esse integrano una violazione delle sue disposizioni, dato che suggeriscono comunque l'espressione di una determinata preferenza. Siamo, quindi, anche in questo caso[93], di fronte ad un'antigiuridicità integrata tramite una mera condotta e non a seguito della realizzazione di un evento dannoso.

La decisione di considerare alla Sezione 6 della proposta di legge irlandese un reato il semplice mantenimento di un comportamento anziché la realizzazione di un evento dannoso può avere una duplice motivazione. La prima ragione che pare trasparire dalla lettera della normativa è la scarsa chiarezza circa il bene giuridico da tutelare imponendo delle limitazioni all'uso di *internet bot* per la diffusione di contenuti *online*. Infatti,

attuabili su Internet nel concetto di "comunicare" vi è anche 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., in cui l'autore opera un'approfondita disamina del testo della riforma legislativa approvata nello Stato della California (che ha diversi punti di contatto con quella qui analizzata) e, riferendosi ai precedenti giurisprudenziali *Mirkin v. Wasserman*, 858 P.2d 568, 578 (Cal. 1993) e *Comm. on Children's Television, Inc. v. Gen. Foods Corp.*, 673 P.2d 660, 674 (Cal. 1983), ha evidenziato come questo genere di pubblicazione di contenuti ben possa qualificarsi come una "comunicazione indiretta" e, quindi, violare le disposizioni del Paragrafo 17941 (a) del *California Business and Professions Code*.

[92] Questa posizione è sostenuta anche in B. Stricke, *People v. Robots: A Roadmap for Enforcing California's New Online Bot Disclosure Act*, cit., 853.

[93] Si veda quanto riferito *supra* al paragrafo 2.1.(C) di questo articolo in merito alle condizioni per la violazione della Sezione 8 (1) del *POFMA* di Singapore.

mentre dal *POFMA* trapela chiaramente la preoccupazione del legislatore singaporiano di proteggere beni giuridici di primaria importanza per lo Stato, quali l'ordine pubblico, la sicurezza e la salute dei suoi cittadini e il corretto svolgimento delle elezioni[94], le indicazioni della *OASM Bill* risultano alquanto laconiche nel definire l'obiettivo della norma (il fine indicato nel preambolo è, infatti, semplicemente di «garantire la trasparenza nella divulgazione delle informazioni nella pubblicità politica online; e per provvedere a questioni correlate»). Sembra, comunque, possibile immaginare che il proposito del legislatore irlandese sia quello di tutelare la libera formazione delle opinioni politiche dei suoi cittadini, evitando che le loro cerchie virtuali di connessioni vengano infiltrate in maniera dissimulata da agenti *software* e sottoposte ad un'intensa opera di diffusione di contenuti secondo modalità per le quali vengano percepiti dagli utenti umani come espressione di un genuino e condivisibile sentimento popolare[95].
La seconda delle ragioni intuibili dal senso della proposta di legge – conseguente alla prima – è la difficoltà quasi insormontabile di provare l'effettiva realizzazione di un evento dannoso, quale il compimento di un'indebita influenza sui convincimenti dei destinatari delle comunicazioni di «un messaggio che promuove uno o più candidati […] o un partito politico registrato nel registro dei partiti politici, o se promuove un messaggio su una questione di interesse o importanza politica»[96], tramite strumenti di falsificazione della reale rinomanza di un contenuto (quali sono gli *internet bot*). Soprattutto, è quasi insuperabile la problematicità di dimostrare che sia stata sufficiente questa forma di manipolazione della celebrità di un'opinione a convincerli della validità di quei candidati, partiti od opinioni[97].
Nonostante, come appena visto, non sia rilevante ai fini dell'attribuzione del carattere delittuoso delle condotte la realizzazione di un evento dannoso, permane il problema di definire quando i contenuti del messaggio comunicato possano dirsi di promozione di candidati, partiti od opinioni. Infatti, se, come detto, anche le forme più puerili di invito all'espressione di una determinata preferenza elettorale sono considerate come forme di promozione che il combinato disposto delle Sezioni 2 e 6 sanziona come illecite, riportare acriticamente solo alcuni dei caratteri, delle affermazioni o degli episodi della vita di un candidato o di un partito o solo alcuni punti del loro programma elettorale risulta già una condotta non immediatamente inquadrabile come esaltatoria o meno. Questo perché un'indicazione parziale di caratteri, azioni o parole di un personaggio politico (o dei membri di un partito), anche quando veritiera e non celebrativa, ben può essere in concreto idonea a fornire al destinatario di una comunicazione *online*

[94] Si veda la Sezione 7 (1)(b), da (i) a (vi), del *POFMA*, analizzata *supra* al paragrafo 2.1.(D) di questo articolo.

[95] Si vedano al riguardo i contributi riportati *supra* alle note 17 e 54.

[96] Sezione 2 (2) della *OASM Bill*.

[97] Nonostante siano stati condotti numerosi studi in merito al fatto che le persone tendono a prestare attenzione a ciò che sembra popolare e a fidarsi delle informazioni che circolano all'interno del loro contesto sociale e l'importante ruolo giocato dalle cerchie virtuali di connessioni nella formazione delle convinzioni personali e delle opinioni politiche, sembra difficile che citare dette ricerche sarebbe sufficiente a provare in sede giurisdizionale la sussistenza di un nesso di causalità tra l'accrescimento artefatto della diffusione di determinati contenuti e l'espressione da parte dei destinatari di dette opere propagandistiche dell'opinione da essi veicolata. Per gli studi condotti in quest'ambito, si veda quanto riportato *supra* alle note 16, 17, 18, 54 e 87.

una rappresentazione distorta del loro valore in un senso per questi favorevole e finire, così, col suscitare o rafforzare in colui che riceve il messaggio il convincimento ad esprimere una preferenza in ambito elettorale che altrimenti non avrebbe compiuto. In altre parole, un resoconto veritiero ed obiettivo ma parziale delle qualità di un candidato o di un partito ben potrebbe essere considerato come una forma di supporto che, ai sensi della *OASM Bill*, non può essere compiuta attraverso l'utilizzo di un *bot*.

Sebbene un'interpretazione della normativa in oggetto in termini così restrittivi possa apparire eccessiva, è necessario prendere in considerazione che questi agenti *software* vengono anche impiegati in gruppi coordinati e composti da un alto numero di unità per la diffusione su ampia scala di contenuti che seppelliscano opinioni ed interpretazioni diverse da quelle dei loro "padroni". Questa attività, definita *astroturfing*[98], può facilmente essere compiuta anche in ambito elettorale per nascondere i contenuti negativi verso un candidato o un partito dietro una quantità esorbitante di altri che riportano – anche in modo veritiero e spassionato – solo aspetti di questi visti come positivi dalla collettività a cui appartengono, impedendo, così, ai loro elettori di compiere delle ricerche complete su di loro e riuscendo, di fatto, a supportarli.

Inoltre, il nodo relativo alla determinazione di quali messaggi abbiano una connotazione promozionale di candidati, partiti od opinioni interessa anche l'apposizione di "reazioni" ai contenuti di terzi tramite un *bot*. Infatti, data l'interpretazione estensiva sopra esposta al sottoparagrafo 2.2.(B)(I) di questo articolo di quali azioni *online* – e in particolare sui *social media* – possano essere considerate modalità di comunicazione di un messaggio, far cliccare ad un *socialbot*, ad esempio, il tasto «Mi piace» sotto ad un *post* su Facebook o quello «Consiglia» su LinkedIn risulta pacificamente integrante una violazione del combinato disposto delle Sezioni 2 (2) e 6 (1) della *OASM Bill*. Questo perché azioni di questo tipo sono in grado di trasmettere al creatore di un *post* e a coloro che possono visionarlo un giudizio di merito sul suo contenuto e, quindi, eventualmente anche di supportare le persone o le opinioni politiche che sono menzionati in esso[99]. L'importanza di questa considerazione deriva anche dal fatto che un alto numero di "reazioni" ad un determinato contenuto e, quindi, ai soggetti o alle opinioni in esso descritti, potrebbe rendere una falsa impressione di popolarità di esso e dei suoi soggetti e indurre delle persone a prestargli attenzione o a ritenerlo attendibile come non avrebbero fatto in assenza della sua celebrità artificiale[100].

---

[98] J. Ratkiewicz – M. Conover – M. Meiss – B. Gonçalves – A. Flammini – F. Menczer, *Detecting and tracking political abuse in social media*, cit.; Ratkiewicz – M. Conover – M. Meiss – B. Gonçalves – S. Patil – A. Flammini – F. Menczer, *Truthy: Mapping the spread of astroturf in microblog streams*, cit.; 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.; N. Abokhodair – D. Yoo – D.W. Mcdonald, *Dissecting a Social Botnet: Growth, Content and Influence in Twitter*, cit.; S.C. Woolley, *Automating power: Social bot interference in global politics*, cit.; 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 – 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.

[99] Questa interpretazione è avanzata anche in B. Stricke, *People v. Robots: A Roadmap for Enforcing California's New Online Bot Disclosure Act*, cit., 853, dove l'autore definisce le reazioni ai contenuti di terzi come "funzioni di sostegno" (in originale "*Endorsement features*").

[100] In merito alla tendenza delle persone a prestare attenzione a ciò che sembra popolare e a fidarsi delle

Rimane da ultimo da notare come la Sezione 2 (2), nel definire le finalità politiche che vengono vietate dalla Sezione 6 nell'impiego dei *bot*, parli solo di promozione di soggetti, organizzazioni ed opinioni politici, tralasciando l'ipotesi che detti agenti *software* vengano, invece, utilizzati per avversarli. Sebbene sia facile immaginare che un tale sfruttamento di questi strumenti digitali autonomi possa rientrare nell'ambito delle condotte vietate dalle norme di tutela della reputazione di persone od organizzazioni, di fatto colmando quella che a prima vista appare come una notevole lacuna, la *OASM Bill* rischia, comunque, di sembrare priva della necessaria considerazione per la diffusione potenziata *online* di due particolari tipologie di contenuti: quelli che riportano solo fatti veri ma considerati riprovevoli dalla collettività o dalla legge e quelli che non promuovono un messaggio su una questione di interesse o importanza politica ma che si limitano, invece, a criticare quelli espressi da altri.

In merito alla prima tipologia di contenuti pare possibile riportare le medesime considerazioni esposte *supra* in merito alla presentazione di solo alcune caratteristiche o qualità di un candidato o di un partito politici, ossia che, pur essendo veritiera, una descrizione parziale è comunque capace di riportare una rappresentazione distorta del suo oggetto. Sebbene tale tipologia di condotta non possa ovviamente essere definita come illecita, data la continenza formale dei messaggi comunicati, essa può avere, comunque, degli effetti nefasti sull'andamento del dibattito pubblico. Infatti, come visto, gli *internet bot* possono essere impiegati anche in attività di *astroturfing* che, in questo caso, sarebbero finalizzate a seppellire i contenuti riguardanti aspetti positivi sotto una soverchiante quantità di altri che riportano solo i caratteri socialmente riprovevoli di candidati o partiti politici esecrati dai "padroni" di questi agenti *software*, finendo, così, con lo sviare artificiosamente le convinzioni personali e le opinioni politiche degli elettori dal loro naturale percorso di formazione.

Per quanto riguarda, invece, la propagazione tramite *web bot* di contenuti contrari ai messaggi espressi da altri in merito a una «questione di interesse o importanza politica», pare qui possibile argomentare che qualsiasi critica – che non scada nell'ingiuria o nella diffamazione – alle opinioni espresse da altri in merito ad una questione sia comunque una forma implicita di promozione della tesi opposta e ricada, quindi, nell'ambito del divieto sancito dalla Sezione 6 della proposta di legge introdotta al Dáil Éireann. Tuttavia, permane una certa ambiguità riguardo alla possibilità di includere queste particolari condotte.

(iii.) Questioni di interesse o di importanza politica – Oltre ai dubbi interpretativi visti in merito al concetto di comunicazione di un contenuto e della determinazione del suo carattere promozionale o meno, la scarna definizione del reato di utilizzo di un *bot* con finalità politiche lascia anche un fosco alone di incertezza in merito a cosa rientri nel concetto di «questione di interesse o importanza politica». Questo anche perché oggi addirittura argomenti come l'uso di materiale medico-sanitario[101] o di cannucce

informazioni che circolano all'interno delle loro cerchie virtuali di connessioni si veda quanto riportato *supra* alle note 16, 17, 18, 54 e 87.

[101] Si vedano, ad esempio, L.H. Kahane, *Politicizing the Mask: Political, Economic and Demographic Factors Affecting Mask Wearing Behavior in the USA*, in *Eastern Economic Journal*, 47(2), 2021, 163 ss.; L. Aratani, *How did face masks become a political issue in America?,* in *theguardian-com*, 29 giugno 2020.

di plastica[102] possono assumere connotazioni violentemente politiche. Soprattutto di fronte alle successive previsioni contenute nella Sezione 2 (2) di quali altri argomenti non possano essere oggetto di promozione tramite l'uso di un *bot* camuffato da essere umano, ossia quelli oggetto di un dibattito parlamentare, di una proposta di legge o di un referendum (quindi, in qualche modo identificabili facendo riferimento a precisi indicatori esterni, come il calendario dei lavori parlamentari), appare ancora più arduo identificare i confini dell'ambito di applicazione di questa indicazione di stampo sostanzialmente residuale, rendendo, così, facilmente plausibile che ai sensi della Sezione 6 (1) nessun argomento possa essere oggetto di contenuti diffusi in forma potenziata dall'impiego di programmi *software* automatizzati[103].

Deve, poi, essere evidenziato come la Sezione 6 (1) della *OASM Bill* vieti di usare i *bot* con l'obiettivo «di presentarsi come un *account* o un profilo individuale», ossia che essa definisca come illecito l'impiego di agenti *software* per gestire «più *account* falsi mascherati da entità individuali»[104]. In altre parole, è la direzione compiuta da dietro le quinte e tramite l'utilizzo di un *bot* di almeno 25 profili, creati in modo da sembrare appartenere ad altrettante persone vere, perché questi supportino un candidato, un politico o un'opinione su di un tema politico (facendoli apparire più popolari di quanto non siano in realtà tra la gente) ad essere decretata come un crimine. Di conseguenza, pare ammissibile – o quantomeno non immediatamente riconoscibile come vietato – un utilizzo "in chiaro" di questi programmi *software* autonomi, ossia il loro impiego non dissimulato per la realizzazione di «più presenze *online*». Questa ipotesi, invero qui non apertamente contemplata, è invece espressamente prevista al Paragrafo 17941 (a) del *Bolstering Online Transparency Act of 2018* della California (con cui la *OASM Bill* ha diversi punti di contatto) e avrebbe i suoi meriti per quanto riguarda le necessità di buon funzionamento del Potere giurisdizionale. Allo stesso tempo, però, lascerebbe aperta – forse anche ingigantendola – la problematica relativa all'apposizione di "reazioni" ai contenuti di terzi. Infatti, mentre un messaggio pubblicato su Internet da un profilo chiaramente marcato come gestito da un *bot* renderebbe consci i suoi lettori della natura artificiale del suo comunicatore e gli permetterebbe di soppesare con maggior precisione e consapevolezza il suo valore, nell'apposizione di reazioni ai contenuti di terzi, l'identità robotica di coloro che hanno compiuto quest'azione si perde nella moltitudine degli identificativi sommariamente riportati tramite un semplice numero, finendo di fatto col perpetrare quell'indebita influenza sulla formazione delle convinzioni personali di coloro ai quali detti contenuti vengono mostrati[105].

---

[102] Si vedano, ad esempio, E. Bradner, *Plastic straws are the subject of the latest 2020 culture war*, in *edition.cnn.com*, 5 settembre 2019; L. O'Neil, *How plastic straws became the latest battleground in the US culture wars*, in *theguardian.com*, 3 agosto 2018.

[103] La vaghezza di questa disposizione è stata messa in evidenza anche nel corso dei dibattiti parlamentari che hanno riguardato la sua approvazione. Si veda, ad esempio, l'intervento compiuto dal parlamentare Brian Stanley nel corso del dibattito del 13 dicembre 2017, disponibile in *Dáil Éireann debate*, cit., durante il quale egli ha espressamente chiesto ai suoi colleghi con fare retorico «cosa costituisce una "questione di interesse politico"?».

[104] Intervento del deputato Lawless durante la presentazione della *OASM Bill* del 06 dicembre 2017.

[105] In merito alla tendenza delle persone a prestare attenzione a ciò che sembra popolare si veda quanto riportato *supra* alle note 15 e 16.

Infine, per quanto la proposta di legge avanzata dal deputato Lawless davanti al Dáil Éireann abbia il merito di aver riconosciuto la necessità ormai impellente di un intervento di regolamentazione di un aspetto nuovo della realtà sociale o politica per poter meglio tutelare i cittadini da indebite influenze – perpetrate attraverso delle vere e proprie frodi in merito ad aspetti ritenuti, anche inconsciamente, come degni di attenzione – sul processo di formazione delle loro opinioni ed orientamenti politici, essa non è esente da difetti, sia in merito alle definizioni proposte che riguardo all'identificazione delle condotte da bollare come criminose. Per di più, l'assenza di qualsiasi potere di adottare dei provvedimenti di restrizione dell'uso degli *account* gestiti da un *internet bot* che abbia violato le disposizioni della legge stessa (come, invece, previsto dalla normativa della Repubblica di Singapore)[106] sembra mettere in luce una scarsa comprensione da parte del legislatore irlandese della capacità di questi agenti *software* di portare avanti la propria opera di propaganda e disinformazione rimanendo al di fuori della giurisdizione delle Autorità deputate a far rispettare la legge dello Stato[107].

3. Conclusioni

l'avvento di Internet e dei *social media* ha permesso uno scambio di contenuti con dei volumi ed una velocità in precedenza impensabili, portando ad un progresso della conoscenza e ad un allargamento dei dibattiti che non sarebbero stati altrimenti possibili negli stessi termini. Tuttavia, le architetture alla base di questo sistema di libero scambio di idee hanno tralasciato alcuni aspetti di sicurezza ed affidabilità che oggi costituiscono delle vere e proprie falle di sistema che permettono a chiunque (umano e non) di mostrarsi con le caratteristiche che più gli aggradano e permettendo, così, di non far trasparire su Internet la propria effettiva natura e di distorcere la realtà dei fatti tramite strumenti relativamente semplici da usare, quali i *socialbot*.

Lo sfruttamento delle fragilità delle piattaforme *social* per la malevola diffusione di contenuti propagandistici e disinformativi è ormai giunto ad un livello tale da essere identificato come uno dei «rischi sistemici» per la tenuta delle istituzioni democratiche di un paese[108]. Di fronte a questi pericoli, alcune nazioni hanno deciso di adottare delle normative che – agli occhi dei loro legislatori – dovrebbero formare un solido bastione di difesa della stabilità dei loro ordinamenti e della fiducia dei loro cittadini.

La Repubblica di Singapore ha voluto rispondere con fermezza a questi nuovi fenomeni approvando una legge che è stata fortemente criticata per la durezza delle soluzioni adottate e per la facilità con cui essa potrebbe essere trasformata da strumento di tutela

[106] Si veda quanto scritto *supra* al paragrafo 2.1.(E) di questo articolo in merito alla Sezione 40 del *POFMA*.

[107] Come già perspicacemente rilevato anche dal deputato Seán Kyne nel corso del dibattito del 13 dicembre 2017, il quale ha dichiarato 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 di là della portata del disegno di legge».

[108] Si veda quanto riportato al considerando 57 della 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)*, detto anche *Digital Service Act*.

dell'opinione pubblica a mezzo di repressione del dissenso (in particolare per via del fatto che ha tra i propri capisaldi una valutazione del contenuto delle opinioni espresse dalle persone)[109]. Inoltre, la normativa dell'isola-stato del sud-est asiatico presenta numerose criticità nella definizione del proprio oggetto di regolamentazione e delle condotte sanzionabili. Tuttavia, tutti questi fattori non possono mettere in secondo piano come il *POFMA* contenga anche delle previsioni interessanti per quanto riguarda il contrasto ai *socialbot*.

La Repubblica d'Irlanda sta cercando a sua volta di arginare un fenomeno a cui ha potuto assistere da vicino e che viene visto dai suoi legislatori come una minaccia per il suo ordinamento. Nel tentativo di edificare una propria difesa legislativa, l'Isola di Smeraldo ha tratto evidente ispirazione dalla normativa statale della California statunitense, cercando, però, anche di superare le problematiche in essa contenute[110]. Al momento, come visto, la proposta di legge in discussione al Dáil Éireann non è priva di criticità – sia per quanto riguarda le definizioni date che in merito alle condotte da vietare – ed ha già subito una battuta d'arresto[111]. Ad ogni modo, essa attualmente non è stata abbandonata e non è nemmeno arrivata alla camera alta del Parlamento irlandese. Rimane, quindi, ancora possibile che venga modificata in futuro in termini più appropriati al suo scopo e poi approvata.

L'analisi qui compiuta delle normative delle due repubbliche (che condividono un modello di ordinamento giuridico di *common law* di origine britannica) ha permesso – almeno agli occhi di chi scrive – di porre in evidenza alcune problematiche relative alla regolamentazione delle nuove tecnologie digitali e dei fenomeni ad esse collegate.

Innanzitutto, il fatto che una produzione legislativa relativa a innovazioni recenti e con profondi riflessi sulla società civile debba per forza appoggiarsi alle esperienze e alle competenze di soggetti che abbiano una piena e profondissima comprensione delle questioni considerate. In caso contrario, il rischio di adottare riforme inefficaci o addirittura controproducenti è estremamente alto.

In secondo luogo – e in diretta conseguenza della considerazione appena esposta – la comparazione qui svolta ha fatto emergere come delle leggi nazionali possano facilmente risultare inadatte ad impedire o anche solo arginare delle condotte nefaste che possono essere compiute al di fuori dei loro ambiti territoriali di applicazione[112]. L'insegnamento che pare potersi trarre da questa comparazione è, quindi, che sia necessario pensare ad una soluzione di regolazione sovranazionale che permetta di superare i limiti degli interventi dei singoli Stati (in questa direzione la Commissione europea ha già mosso i primi passi tramite la promozione dello *Strengthened Code of Practice on*

[109] Si veda al riguardo quanto riportato *supra* al paragrafo 2.1. di questo articolo e, in particolare, alle note 26, 27, 28 e 29.

[110] Per un'analisi delle problematiche della normativa di contrasto ai *socialbot* della California statunitense, si veda A. Tedeschi Toschi – G. Berni Ferretti, *Il contrasto legislativo ai socialbot e le soluzioni avanzate negli Stati Uniti d'America a livello federale e statale*, cit.

[111] Si veda in merito M. O'Halloran, *Government defeated on online advertising and social media Bill*, in *irishtimes.com*, 14 dicembre 2017. Successivamente al voto contrario della Camera bassa, la proposta di legge è tornata alla commissione parlamentare dove dal 31 marzo 2021 è ancora in fase di modifica.

[112] Si vedano a questo proposito le perplessità esposte dal deputato irlandese Seán Kyne e riportate *supra* alla nota 103.

Disinformation 2022, un codice di autoregolamentazione per i *Social Network Providers*)[113]. In ultima analisi, le normative qui considerate sono tutto fuorché adeguate, sia a causa delle loro criticità strutturali sia perché cercano di regolare in forma locale un fenomeno transnazionale che esula dagli ambiti di esercizio della sovranità dei singoli Stati, ma costituiscono comunque un primo tentativo da parte dei legislatori nazionali di garantire una maggior tutela alla libera formazione delle opinioni dei loro cittadini.

[113] La stessa Commissione europea, però, aveva riconosciuto come la prima versione di questo codice di autoregolamentazione avesse delle criticità significative, tra le quali un'applicazione incoerente e incompleta della sua disciplina tra le diverse piattaforme e gli Stati membri, delle limitazioni intrinseche alla sua natura di strumento di autoregolamentazione e diverse lacune (SWD(2020) 180 final, *Assessment of the Code of Practice on Disinformation–Achievements and areas for further improvement*).