

“Deplatforming”, l’attacco a Capitol Hill e la nuova sfera pubblica privatizzata

Philip Di Salvo, PhD

Istituto di media e giornalismo (IMeG)

Università della Svizzera italiana (USI)

philip.di.salvo@usi.ch

Abstract:

Il 2021 si è aperto con l’attacco a Capitol Hill, perpetrato da una folla di sostenitori di Trump convinta, anche dalle dichiarazioni dell’ex Presidente Usa, del fatto che le elezioni fossero state truccate in favore di Biden. In seguito all’assalto, le principali piattaforme di social networking al mondo hanno rimosso gli account di Trump, accusato di usare questi spazi digitali per istigare la violenza. L’episodio, senza precedenti nella storia degli Usa, ha anche fatto ri-esplosione il dibattito attorno alla censura sui social media e al tema del “deplatforming”, ovvero la rimozione di profili o contenuti dai social media in relazione alla violazione dei termini di utilizzo di questi ultimi. In questo articolo si discuterà della possibilità che la rimozione di Trump abbia rappresentato effettivamente un caso di censura, contestualizzandolo nel più ampio discorso sulla libertà di espressione in rete e sulla complessiva privatizzazione della sfera pubblica digitale.

Dalla Casa Bianca all’assalto a Capitol Hill

Nel gennaio del 2021, quando il Congresso statunitense fu vittima di un attacco terroristico perpetuato da una galassia eterogenea di sostenitori del Presidente uscente Donald Trump, composta da elettori, estremisti di destra e gruppi armati della galassia dell’*alt-right*, non è avvenuto solo un evento unico dal punto di vista politico, ma anche mediale e tecnologico. Senza dubbio l’assalto a Capitol Hill - durante il quale morirono 5 persone - ha rappresentato il momento di massima tensione di un passaggio storico iniziato con la vittoria di Trump alle Presidenziali del 2016 e conclusosi proprio quel giorno, tra i vetri infranti del palazzo e le dichiarazioni pubbliche di Trump, prossime all’*eversione*. Quei cinque anni hanno stravolto radicalmente la politica statunitense e, insieme ad essa, anche l’ecosistema mediale del paese e il suo rapporto con il web. In particolare, quel periodo ha visto sorgere un dibattito attorno al concetto di “deplatforming”, con cui si intende la “rimozione di un account da un social media a causa della violazione delle regole della piattaforma” (Rogers, 2020) e sui poteri delle aziende della Silicon Valley. L’oscuramento degli account di Trump, in particolare, ha portato la questione in superficie con inedita centralità, inserendosi anche nei discorsi pubblici attorno alla censura e a quelli riguardanti l’*ascesa* delle forze politiche della destra estrema e populista e dell’*alt-right* negli Usa come altrove, temi mai così intrecciati.

Lo stesso profilo Twitter di Trump ha svolto un ruolo centrale in queste trasformazioni., Canale prioritario di comunicazione dell’ex Presidente e megafono e amplificatore di tutte le spinte più reazionarie e violente della galassia della nuova destra razzista statunitense, il profilo personale di Trump è sempre stato un osservato speciale in termini di *hate speech* e uso sconsiderato del potere: negli anni, ad esempio, Trump lo ha utilizzato per minacciare azioni militari contro i principali nemici degli Usa, per amplificare teorie del complotto di varia natura, per attaccare la

stampa, le minoranze oltre che per bypassare, in ottica di disintermediazione, la stampa mainstream.

Secondo diversi osservatori, a Donald Trump è stato concesso impunemente di violare platealmente le “regole” di Twitter, postando contenuti che, qualora fossero stati messi online da utenti comuni o comunque senza potere, avrebbero portato a sanzioni immediate da parte della piattaforma (Conger, Isaac e Frenkel 2021). Solo i fatti di Capitol Hill - senza precedenti per la storia degli Stati Uniti come per la maggior parte delle democrazie contemporanee - hanno smosso le piattaforme, facendole intervenire. Dopo una serie di blocchi di breve durata, Twitter ha sospeso permanentemente il profilo @realDonaldTrump per “incitazione alla violenza” l’8 gennaio 2021¹, decisione cui ha fatto seguito anche quella di Facebook, che prima ha imposto un silenziamento esteso fino all’insediamento di Biden, trasformato poi in un blocco definitivo a maggio 2021, per decisione dell’Oversight Board dell’azienda di Menlo Park². Il blocco riguarda anche Instagram. Da subito, non sono mancate le reazioni polarizzate alle decisioni delle piattaforme (Noor 2021). Se da più parti si è parlato di una legittima (per quanto tardiva) decisione da parte delle aziende responsabili dei social media, altre voci hanno parlato apertamente di censura e di volontà di silenziare Trump. Al centro della diatriba, a detta di chi scrive, vi è una riflessione più ampia da tenere in considerazione: ovvero se sia concretamente possibile censurare l’essere umano più potente della terra, togliendogli voce sulle piattaforme social, specialmente in un passaggio storico/cronachistico di massima esposizione mediatica come quello in cui sono avvenuti i fatti di Capitol Hill. In parallelo, è comunque fondamentale non sottovalutare il ruolo e il potere delle piattaforme tecnologiche in questo contesto, senza però cadere in facili panici morali. In primis occorre chiarire un punto: le piattaforme social hanno certamente un problema di censura, ma il caso Trump non ne fa parte.

La “vera” censura aziendale nella sfera pubblica privatizzata

Con ottimo tempismo, l’editore Verso ha dato alle stampe il nuovo libro di Jillian C. York, Director for International Freedom of Expression della Electronic Frontier Foundation (EFF), *Silicon Values. The Future of Free Speech under Surveillance Capitalism* (2021). York è una delle massime esperte internazionali di libertà di espressione online e il suo testo è destinato a diventare uno dei punti di riferimento per questi temi, in particolare se messi in relazione con gli elementi costitutivi dell’economia e del potere digitale sui cui si basa la forza delle grandi piattaforme. Il libro di York contiene numerosi esempi di censura effettiva perpetrata ai danni di gruppi minoritari da parte delle piattaforme, complice i rapporti stretti che queste intrattengono con i governi e i loro meccanismi di moderazione dei contenuti, in parte automatizzati o connessi ai meccanismi di segnalazione/*flagging* affidati agli utenti.

York fa riferimento, ad esempio, citando la definizione data dalla ricercatrice Evelyn Douek (2020), ai “content cartel” (“cartelli dei contenuti”, nda), un termine che fa riferimento a tutti i rapporti informali intercorsi tra governi e piattaforme il cui obiettivo è la moderazione dei

¹ La comunicazione ufficiale di Twitter è disponibile qui:

https://blog.twitter.com/en_us/topics/company/2020/suspension.html

² La decisione dell’Oversight Board di Facebook è disponibile qui:

<https://about.fb.com/news/2021/05/facebook-oversight-board-decision-trump/>

contenuti sui social media per finalità politiche e di censura, dinamiche da cui la società civile è normalmente esclusa, con forti ripercussioni in termini di trasparenza. Tra i vari esempi citati da York vi è, ad esempio, il rapporto stretto e formalizzato che intercorre tra la sede di Facebook a Tel Aviv e il governo israeliano, una collaborazione che ha come oggetto sia i contenuti degli utenti israeliani che quelli dei palestinesi e che è marcata da un palese *bias* geopolitico che interferisce direttamente con la libertà di espressione degli utenti palestinesi. Il nucleo del problema, come scrive York, è che le dinamiche di forza (e di censura) del mondo fisico trovano così uno specchio digitale che le ripropone online, con la mediazione di aziende private che gestiscono i flussi comunicativi di miliardi di utenti su scala mondiale. In seguito al riaccendersi degli scontri nella Striscia di Gaza a maggio 2021 si è tornato a parlare dei numerosi esempi di censura segnalati da utenti palestinesi che si sono visti limitare senza ragione o per errore i propri profili e contenuti su varie piattaforme, compresi hashtag di protesta, video livestream o che hanno subito la rimozione diretta di vari contenuti (Ingram, 2021). Il gruppo attivista palestinese 7amleh, the Arab Center for Social Media Advancement, ha pubblicato di recente il suo report annuale che raccoglie e mappa i vari episodi censori occorsi alle voci palestinesi online³.

Questi episodi, però, non riguardano certamente solo la Palestina. In alcuni casi, anche pratiche di moderazione dei contenuti che sulla carta potrebbero sembrare benefiche, come quelle che portano alla rimozione - in modo sostanzialmente automatizzato - di contenuti giudicati come "terroristici" possono essere controverse nei loro effetti. Almeno dal 2017, ad esempio, Facebook ha introdotto sistemi di moderazione gestiti da sistemi di *machine learning* al fine di individuare ed eliminare i contenuti video e foto di sostegno ai gruppi estremisti e terroristici (ISIS in particolare), una pratica poi replicata anche da altre piattaforme, YouTube in primis (York 2021, 115-123). Per quanto efficiente nell'individuare ed eliminare, ad esempio, i video propagandistici delle decapitazioni, questo approccio automatizzato ha però anche degli importanti risvolti censori in quanto, in molti casi, va a eliminare contenuti che potrebbero rappresentare testimonianze storiche o prove di crimini di guerra. L'algoritmo, infatti, non è necessariamente in grado di distinguere un post di istigazione alla violenza a uno che testimonia un episodio di questa natura. Gli stessi algoritmi, scrive ancora York, hanno eliminato dai social media video che testimoniavano le violenze del conflitto in Siria, ad esempio, e in alcuni casi anche prove che potrebbero essere utilizzate nel contesto di processi per crimini di guerra. Una denuncia simile, ad esempio, è pervenuta da Syrian Archive⁴, un'organizzazione che si occupa di raccogliere e archiviare contenuti social che documentano potenziali violazioni delle leggi internazionali e che ha denunciato come l'algoritmo di YouTube abbia rimosso contenuti fondamentali in termini legali e storici anche senza che questi fossero materiale di propaganda estremista (Asher-Schapiro, 2017).

A complicare ulteriormente lo scenario è il fatto che gli algoritmi di moderazione dei contenuti delle grandi piattaforme operano spesso in modo opaco e/o senza che i criteri utilizzati siano

³ Il report è disponibile qui, in lingua inglese: [https://7amleh.org/storage/Research%20and%20Position%20Papers/Hashtag_Palestine_English_2020_9May%20\(1\).pdf](https://7amleh.org/storage/Research%20and%20Position%20Papers/Hashtag_Palestine_English_2020_9May%20(1).pdf).

⁴ Il sito dell'organizzazione è disponibile qui: <https://syrianarchive.org/>.

accessibili o pubblicamente verificabili. Questo avviene in particolare perché questi algoritmi sono di fatto tecnologie proprietarie gelosamente custodite dalle aziende che le utilizzano, un elemento che non aiuta la verifica della presenza di eventuali bias razziali o culturali inclusi in fase di programmazione. Inoltre, l'apporto della moderazione automatizzata si affianca al lavoro dei moderatori umani, le cui condizioni di lavoro spesso estreme in termini psicologici sono state efficacemente trattate anche dal documentario "The Cleaners" di Hans Block e Moritz Rieseewieck (2018). Nessun algoritmo, infatti, è in grado di sostituirsi completamente a un essere umano in queste mansioni o anche di risolvere o correggere i problemi connessi al lavoro umano in questo senso. Come scrive ancora Jillian C. York (2021, 176-177), sia la "formazione" degli algoritmi che le risorse impiegate per i moderatori "umani" non sono equamente distribuite e molti idiomi non hanno una copertura adeguata o la giusta attenzione, per cominciare. York cita a questo proposito il caso della Lingua luganda, l'idioma parlato da 8 milioni di persone in Uganda: senza alcun supporto specifico per questa lingua, scrive l'attivista della EFF, è inevitabile che la moderazione dei contenuti in lingua luganda porti ad abusi, errori e alla rimozione di elementi che non dovrebbero scomparire dalla rete. Problemi di competenza e discernimento simili, però, possono emergere anche nel contesto di altre lingue più diffuse.

L'apporto della moderazione automatizzata, però, è in crescita e la pandemia ha solo accelerato la sua applicazione. A fine 2020 Facebook ha dichiarato come i suoi sistemi di intelligenza artificiale siano ora in grado di eliminare il 94.7% dei contenuti definibili come *hate speech* (Sam Shead, CNBC», 19 novembre 2020), ma esiste ben poca trasparenza su quali contenuti e per quali ragioni vengono inseriti nelle liste di elementi potenzialmente problematici. Questo elemento fa della moderazione algoritmica dei contenuti una "black box" (una "scatola nera", ndr) per usare la metafora utilizzata dal sociologo Frank Pasquale nel suo libro *The Black Box Society*. In questa accezione, le "black box" sono una tecnologia proprietaria i cui meccanismi e dettagli di programmazione sono per lo più inaccessibili e oscuri, nonostante le profonde implicazioni sociali del loro utilizzo (Pasquale 2015). Gli esempi qui citati sono solo alcuni di quelli che vanno a fornire una tassonomia piuttosto ampia della ramificazione della censura applicata dalle aziende della Silicon Valley sui social media. Le categorie e le aree geografiche colpite sono molto più numerose e toccano anche e più direttamente la comunità LGBTQ, le minoranze etniche e linguistiche e i gruppi già più marginalizzati (Lux 2017). Di nuovo, le forze in gioco nel mondo fisico si ripresentano come dinamiche di potere fortemente sbilanciate anche in quello online. Tornando al caso del blocco di Donald Trump sulle principali piattaforme è davvero possibile inserire quell'episodio nella medesima tassonomia? O siamo di fronte a qualcosa di diverso?

Da tempo si discute della possibilità che le piattaforme social abbiano fin qui inteso la libertà di espressione con un sistema a doppio binario, ha scritto Arianna Ciccone su *Valigia Blu* (2021), difendendo la scelta delle piattaforme di eliminare i profili dell'ex Presidente Usa. Se, da un lato, servizi come Facebook o Twitter, come abbiamo visto qui in precedenza, silenziano per compiacenza, errore o per fiducia cieca nei confronti dei loro algoritmi migliaia di contenuti legittimi provenienti da voci marginalizzate, lo stesso non avviene con la medesima attenzione e sofisticatezza nei confronti delle personalità politiche (Schleifer, 2019). Al contrario, fino ai fatti di Capitol Hill, la sofisticatezza si era notata solo negli artifici retorici e nelle dichiarazioni a

mezzo stampa usate per giustificare il fatto che proprio alle personalità politiche più influenti fosse sostanzialmente consentito usare le piattaforme per dire e fare alcunché, come se per le voci dei più potenti della terra non valessero le regole che valgono comunemente per gli utenti comuni, cui è chiaramente vietato usare i social media per incitare la violenza, il terrorismo, sottoporre altre persone ad abusi o ad odio razziale e di genere⁵. Come scrive Jillian C. York, ad esempio, per anni le piattaforme hanno cercato di non intervenire nei confronti degli abusi della politica con i loro strumenti, mascherando la loro inazione dietro concetti piuttosto vaghi. Dal 2016, ad esempio, Facebook ha fatto riferimento alla *newsworthiness exemption* - una sorta di “eccezione di notiziabilità”, ndr - secondo la quale, quanto detto dai politici sarebbe sempre da intendere come “notiziabile” e quindi non condizionato alle regole in vigore per tutti. Nel 2019, Nick Clegg - già vice primo ministro britannico poi divenuto responsabile degli affari globali di Facebook e capo della sua comunicazione - aveva ribadito come quella “eccezione” sarebbe stata estesa a qualsiasi voce politica, senza distinzione⁶. Questo crea esattamente quel “doppio binario” di considerazione delle libertà degli utenti che mette chi è già in una posizione di privilegio, potere e visibilità in una condizione di minore scrutinio sui propri contenuti mentre un utente comune potrebbe - come accade di frequente - vedersi cancellare i propri profili per aver parlato di ISIS in un contesto giornalistico o per aver postato un’immagine di nudo artistico.

Quanto avvenuto con Trump dopo Capitol Hill rappresenta invece la prima occasione in cui le regole comuni sulle piattaforme sono state applicate in modo trasparente e diretto anche nei confronti del Presidente degli Stati Uniti. La domanda che andrebbe posta in questo contesto è pertanto perché un intervento di questo tipo sia avvenuto solo nel contesto di un attacco terroristico armato e con un Parlamento sotto assedio, atti che Trump ha appoggiato se non incitato - usando anche i social come megafono di viralità. La sospensione di Trump dalle maggiori piattaforme è stata poi spesso definita come “senza precedenti” o come un evento destinato a creare un precedente pericoloso. Se da un lato non ci sono dubbi che la sospensione di un Presidente eletto in un paese democratico dai maggiori servizi di comunicazione digitali rappresenti certamente un *unicum*, è altrettanto vero che altre figure politiche controverse, provenienti però da altre latitudini, abbiano in passato vissuto situazioni simili. Ad esempio, nel 2018 Facebook ha rimosso i profili di Min Aung Hlaing, comandante in capo dell’esercito birmano poi responsabile del colpo di stato del 2021 per *hate speech* (Spencer e Thein 2018); Hezbollah - che siede nel parlamento libanese - è stata rimossa da Facebook e Twitter nel 2018 (Sharkov, 2018); nel 2020, un post del Presidente brasiliano Bolsonaro fu rimosso in quanto contenente disinformazione sulla COVID-19 (Constine, 2020); nel 2020 Facebook ha rimosso il profilo del politico indiano T. Raja Singh, di nuovo per *hate speech* (Ajmal, 2020); nel 2017 furono chiusi i profili Facebook e Instagram del leader ceceno Ramzan Kadyrov (Bennets, 2017). Nel 2019, invece, per restare negli USA, Facebook rimosse i profili di diverse personalità statunitensi di estrema destra come Alex Jones di *Infowars*, Milo Yiannopoulos, Paul Joseph Watson, Laura Loomer, e Paul Nehlen (Lorenz, 2019).

⁵ Qui, gli “Standard della community” di Facebook: <https://www.facebook.com/communitystandards/>. Qui, le regole simili di Twitter: <https://help.twitter.com/en/rules-and-policies/twitter-rules>

⁶ Qui la dichiarazione ufficiale di Nick Clegg: <https://about.fb.com/news/2019/09/elections-and-political-speech/>

Verso una società regolamentata come un'azienda?

Al centro del problema vi sono quindi due questioni: la prima riguarda l'equità e la trasparenza nell'applicazione delle stesse norme che le piattaforme si sono date. La seconda, più ampia e di contesto, è connessa invece al ruolo che queste piattaforme private hanno assunto nella sfera pubblica contemporanea (Floridi, 2021). Per la prima questione da anni si susseguono le richieste di maggiore trasparenza nei riguardi di come Facebook, Twitter e gli altri giganti della Silicon Valley approcciano al materia della moderazione dei contenuti. Per quanto vi siano stati dei miglioramenti importanti negli ultimi anni, o comunque si siano visti tentativi che cercassero di rendere il tema più accessibile alle opinioni della società civile, il filtraggio dei contenuti sulle piattaforme è ancora, per lo più, una "black box" tecnologica. Non stupisce, quindi, che alcuni dettagli sul funzionamento della macchina censoria di Facebook, ad esempio, siano emersi grazie al contributo di alcuni whistleblower che hanno collaborato con la stampa al fine di gettare un po' di luce su queste pratiche oscure e potenzialmente molto impattanti da un punto di vista sociale (Fiveash, 2017; Madrigal 2018; Di Salvo, 2019). Al contrario, come ha ricordato anche la Electronic Frontier Foundation (EFF) a fine 2020, i numeri pubblicati dalle piattaforme con i loro report annuali sulla rimozione dei contenuti dicono poco oltre alle cifre sterili che riportano su dove e quanto le piattaforme moderano o censurano.

A essere fondamentale, qui, è il contesto in cui queste decisioni vengono prese, specialmente con la crescita della moderazione automatizzata. Queste decisioni, e la loro trasparenza, ad ogni modo, non sono mai solo questioni tecnologiche. In molti casi sono intrinsecamente geopolitiche, come dicevamo già in precedenza. Il livello di dibattito attorno alle azioni delle piattaforme nei confronti del blocco imposto a Trump, rispetto alla scarsa attenzione nei confronti di situazioni analoghe in altri contesti geografici è la prova dell'esistenza di un pregiudizio occidentale nei confronti di questi temi e dell'attenzione delle piattaforme che, pur dichiarandosi globali, operano spesso replicando gli schemi geopolitici vigenti. Questo è emerso chiaramente di recente con le rivelazioni della ex dipendente di Facebook Sophie Zhang, divenuta whistleblower, che ha denunciato i doppi standard di intervento di Facebook nei confronti delle situazioni di *weaponization* della sua piattaforma da parte di governi o forze politiche che utilizzano le piattaforme digitali come cannoni di disinformazione e propaganda (Wong 2021). Come è emerso grazie ai dettagli forniti da Zhang alla stampa, Facebook non sembra prestare la medesima attenzione a questi fenomeni quando questi avvengono, ad esempio, in Honduras e non negli Stati Uniti o in Europa. Una conferma, se ancora ve ne fosse bisogno, di come l'idea di un cyberspazio libero da condizionamenti politici ed economici e senza confini, come immaginato ad esempio nella "Dichiarazione d'Indipendenza del Cyberspazio"⁷ di John Perry Barlow nel 1996, sia rimasta una utopia soffocata dalla *realpolitik* e dagli interessi economici dei "custodi di Internet".

"Custodi di Internet" non è una espressione casuale, ma un riferimento esplicito a *Custodians of the Internet. Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*, libro fondamentale di Tarleton Gillespie, Principal Research per Microsoft Research

⁷ Il testo è disponibile qui: <https://internethalloffame.org/news/in-their-own-words/declaration-independence-cyberspace>.

(2018). Come scrive Gillespie, le piattaforme della Silicon Valley sono cresciute sulle basi del caos del web originario, fornendo spazi di comunicazione organizzati e controllati, facili da utilizzare e progettati per crescere esponenzialmente, sia in termini di utenza che di monetizzazione. Progressivamente, però, quel caos è tornato indietro però come un boomerang mentre quella crescita avveniva in modo irrefrenabile (2018, 5). Che queste piattaforme potessero essere “neutrali”, come sostenuto dalle più utopiche e deterministe narrative tecno-entusiaste, è una fantasia, scrive sempre Gillespie: le piattaforme hanno assunto progressivamente un ruolo centrale nella società contemporanea e hanno progressivamente anche accumulato un potere davvero senza precedenti. Sono diventate infrastrutture sociali a tutti gli effetti, la cui centralità è destinata irreversibilmente ad aumentare. Questo è dovuto a dinamiche ben più ampie di quelle puramente tecnologiche e che toccano schemi geopolitici interni anche al capitalismo contemporaneo. Di certo, però, questo posizionamento delle piattaforme è avvenuto in assenza di un reale dibattito sulle responsabilità sociali delle piattaforme (Gillespie 2018, 206) e si è dato per scontato che strumenti progettati per la monetizzazione potessero anche, di default, farsi carico di ruoli democratici e di essere attori di responsabilità in ambiti molto sensibili come la libertà di espressione e i diritti. La moderazione dei contenuti, però, è l'essenza delle piattaforme stesse, ricorda ancora Gillespie (2018, 207) o, di sicuro, è l'essenza di quello che le piattaforme sono diventate oggi, mentre si presentavano come qualcosa che non potevano essere: luoghi neutrali o spazi pubblici di agency politica.

I fatti di Capitol Hill, quanto meno, hanno fatto esplodere il problema con evidenza senza precedenti, mandando in soffitta - si spera definitivamente - alcuni falsi miti e aspettative che hanno inevitabilmente inquinato la comprensione di questi fenomeni fin qui. È il momento, ora di decidere come dare forma alla rete del futuro perché, inevitabilmente, da queste decisioni dipenderà anche la forma della democrazia.

Riferimenti bibliografici:

Ajmal, A. (2020) Facebook bans T Raja Singh, tags him as ‘dangerous individual’, in “Times of India”,
http://timesofindia.indiatimes.com/articleshow/77907922.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst.

Asher-Schapiro, A. (2017) YouTube and Facebook Are Removing Evidence of Atrocities, Jeopardizing Cases Against War Criminals, in “The Intercept”,
<https://theintercept.com/2017/11/02/war-crimes-youtube-facebook-syria-rohingya/>.

Bennets, M. (2017) Russia calls for answers after Chechen leader's Instagram is blocked, in "The Guardian", <https://www.theguardian.com/world/2017/dec/26/chechnya-ramzan-kadyrov-social-media-russia-instagram-facebook>.

Ciccone, A. (2021) "Deplatforming" Trump: la giusta decisione di Facebook e Twitter di bloccare gli account del presidente uscente, in "Valigia Blu", <https://www.valigiablu.it/deplatforming-trump-facebook-twitter/>.

Conger, K., Isaac, M. e Frenkel, S. (2021) Twitter and Facebook Lock Trump's Accounts After Violence on Capitol Hill, in "The New York Times", <https://www.nytimes.com/2021/01/06/technology/capitol-twitter-facebook-trump.html>.

Constine, J. (2020) Facebook deletes Brazil president's coronavirus misinfo post, in "TechCrunch", <https://techcrunch.com/2020/03/30/facebook-removes-bolsonaro-video/>.

Di Salvo, P. (2019) Verità, controllo e manipolazione nella società dei dati, in "La nostra città futura", Fondazione Giangiacomo Feltrinelli, <https://fondazionefeltrinelli.it/verita-controllo-e-manipolazione-nella-societa-dei-dati/>.

Douek, E. (2020) The Rise of Content Cartels, Knight First Amendment Institute, <https://knightcolumbia.org/content/the-rise-of-content-cartels>.

Electronic Frontier Foundation. (2020) Thank You For Your Transparency Report, Here's Everything That's Missing, <https://www.eff.org/deeplinks/2020/10/thank-you-your-transparency-report-heres-everything-thats-missing>.

Fiveash, K. (2017) Facebook content moderation guidelines leaked, in "ArsTechnica", <https://arstechnica.com/information-technology/2017/05/facebook-moderator-guidelines-leaked-tory-internet-regulation/>.

Floridi, L. (2021) Trump, Parler, and Regulating the Infosphere as Our Commons, in "Philosophy & Technology", 34, <https://link.springer.com/article/10.1007/s13347-021-00446-7>.

Gillespie, T. (2018) *Custodians of the Internet. Platforms, content moderation, and the hidden decisions that shape social media*, New Heaven, Yale University Press.

Ingram, M. (2021) Social networks accused of censoring Palestinian content, in "The Columbia Journalism Review", https://www.cjr.org/the_media_today/social-networks-accused-of-censoring-palestinian-content.php.

Lorenz, T. (2019) Instagram and Facebook Ban Far-Right Extremists, in "The Atlantic", <https://www.theatlantic.com/technology/archive/2019/05/instagram-and-facebook-ban-far-right-extremists/588607/>.

Lux, C. (2017) Facebook's Hate Speech Policies Censor Marginalized Users, in "Wired", <https://www.wired.com/story/facebooks-hate-speech-policies-censor-marginalized-users/>.

Madrigal, A. (2018) Silicon Valley Sieve: A Timeline of Tech-Industry Leaks, in "The Atlantic", <https://www.theatlantic.com/technology/archive/2018/10/timeline-tech-industry-leaks/572593/>.

Noor, P. (2021) Should we celebrate Trump's Twitter ban? Five free speech experts weigh in, in "The Guardian", <https://www.theguardian.com/us-news/2021/jan/17/trump-twitter-ban-five-free-speech-experts-weigh-in>.

Pasquale, F. (2015) *The Black Box Society. The Secret Algorithms That Control Money and Information*, Cambridge, Harvard University Press.

Rogers, R. (2020) Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media, "The European Journal of Communication", 35(3), pp. 213–229.

Schleifer, T. (2019) Facebook and YouTube will keep letting politicians say what they want if it's "newsworthy", in "Vox", <https://www.vox.com/recode/2019/9/26/20885783/facebook-twitter-youtube-policies-political-content>.

Sharkov, D. (2018). Facebook and Twitter for Hezbollah Abruptly Deactivated, Group Says, in "Newsweek", <https://www.newsweek.com/facebook-and-twitter-hezbollah-deactivated-group-claims-992666>.

Spencer, O. e Thein, Y. Y. (2018) Has Facebook censored Myanmar's commander-in-chief?, in "Frontier Myanmar", <https://www.frontiermyanmar.net/en/has-facebook-censored-myanmars-commander-in-chief/>.

The Cleaners, regia di Hans Block e Moritz Riesewieck (2018).

Wong, J. C. (2021) How Facebook let fake engagement distort global politics: a whistleblower's account, in "The Guardian", <https://www.theguardian.com/technology/2021/apr/12/facebook-fake-engagement-whistleblower-sophie-zhang>.

York, J. C. (2021) *Silicon Values. The Future of Free Speech under Surveillance Capitalism*, Londra, Verso.