



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

UNIVERSITÀ DEGLI STUDI DI BOLOGNA

Dipartimento di Scienze Politiche e Sociali - SPS

Laurea Magistrale in Governance e Politiche dell’Innovazione Digitale - GePID

**DUBITARE DI CIÒ CHE VEDIAMO:
DISINFORMAZIONE TRA *FAKE NEWS* E *DEEPFAKE*.**

Laura Tonsi

Matricola 1164043

Politica Comparata nell’Era Digitale (B0327)

A.A. 2024/25

<i>Introduzione</i>	1
<i>Fake news e social media : come interferiscono nel processo elettorale?</i>	2
<i>I deepfake: la nuova frontiera della disinformazione</i>	4
<i>Conclusioni</i>	6
<i>Bibliografia</i>	7

Introduzione

Nella società odierna, Internet e i social media non sono più semplici mezzi di comunicazione ma una componente imprescindibile della vita quotidiana di ciascuno di noi. Il Digital 2024 Global Overview Report conferma come l'utilizzo dei social sia oggi più diffuso che mai. L'analisi rivela infatti che attualmente esistono 5,22 miliardi di utenti attivi a livello globale. Tra le motivazioni principali che spingono gli utenti a utilizzare queste piattaforme, vi sono la lettura di notizie (34,4%) e la ricerca di contenuti quali articoli o video (30,4%) (Kemp, 2024).¹

Come sottolineato da Maddalena and Gili (2018), il fenomeno della digitalizzazione è interconnesso a quello della globalizzazione. Tuttavia, ciò ha portato anche a un aumento della diffusione di menzogne e conseguentemente del numero d'individui ingannati. La questione delle fake news è ampiamente dibattuta, soprattutto in relazione ai social media che amplificano sia la portata che la modalità di ricezione da parte degli utenti. A oggi non esiste una definizione universale di fake news. Alcuni autori le classificano come una forma di misinformazione (senza intento dannoso), altri di disinformazione (intento di arrecare danno), mentre altri ancora le associano a entrambi i concetti (Aïmeur et al., 2023; Wardle & Derakhshan, 2017). Tuttavia, il Collins Dictionary le definisce come “informazioni false e talvolta sensazionalistiche presentate come fatti, pubblicate e diffuse su Internet” (Collins English Dictionary, 2024).

Il tema della disinformazione è rilevante poiché rappresenta non solo una minaccia per i diritti umani ma compromette anche i processi elettorali, mettendo a rischio la libertà di pensiero, il diritto alla privacy e la partecipazione democratica (Colomina et al., 2021). L'obiettivo del paper è quindi quello di analizzare le modalità di creazione e diffusione delle fake news e dei deepfake attraverso la revisione di casi recenti, come quello della fabbricazione di fake news nelle Filippine e della diffusione di deepfake durante le elezioni statunitensi nel 2024.

¹ Dati di Ottobre 2024.

Fake news e social media : come interferiscono nel processo elettorale?

Quello delle *fake news* non è un fenomeno moderno: la loro origine risale infatti all'invenzione della stampa (Soll, 2016). L'avvento della tecnologia e in particolare dei social network ha tuttavia amplificato questo fenomeno, soprattutto in termini di velocità e facilità di diffusione (Shu et al., 2017). Visto il numero di utenti attivi sulle piattaforme social, le *fake news* possono avere un impatto significativo sui lettori: negli Stati Uniti e nel Regno Unito, ad esempio, i social sono diventati la fonte d'informazione primaria per i giovani (Wakefield, 2016). La ricerca di Vosoughi et al. (2018) ha dimostrato che le *fake news* si diffondono online sei volte più rapidamente rispetto ai contenuti veritieri e che il 70% degli utenti non è più in grado di distinguere il vero dal falso. Questi dati sollevano preoccupazioni sul rischio delle disinformazioni, particolarmente evidente nelle campagne elettorali e nei processi democratici, aree particolarmente colpite dalle *fake news*. Le elezioni presidenziali statunitensi del 2016 rappresentano un esempio concreto di come questo fenomeno non possa essere sottovalutato: il 9,15% delle notizie diffuse digitalmente era fabbricato, rispetto al 7,76% prodotto dai media tradizionali. Questa tendenza evidenzia la crescente erosione dell'accuratezza delle notizie stesse (Silverman, 2016). Il problema, tuttavia, non riguarda solo gli Stati Uniti ma ha una portata globale, come documentano i casi di diffusione di notizie false su diverse piattaforme social durante le elezioni in Brasile, Taiwan, Germania, Francia e Myanmar (Bradshaw & Howard, 2018). La ricerca condotta da Prajapati et al. (2024) ha analizzato la diffusione delle *fake news* in sei Paesi (USA, India, Taiwan, Sudafrica, Brasile e Francia), analizzando i dati e somministrando sondaggi per valutare l'impatto che queste hanno sulla percezione degli individui. Nei Paesi campione, le *fake news* hanno mostrato un picco di diffusione durante i cicli elettorali, seppur in modo diseguale. Questi contenuti si concentrano generalmente su temi di forte rilevanza nazionale, come l'immigrazione in Francia, la corruzione in Brasile e Taiwan e la religione in India. Tuttavia, calcolare l'impatto reale delle *fake news* resta una sfida complessa, poiché molti elettori le incontrano inconsapevolmente. Ciononostante, i risultati della ricerca rivelano che l'esposizione a contenuti falsi altera la percezione degli elettori: il 57% dei partecipanti brasiliani al sondaggio e il 43% di quelli filippini affermano infatti di essere stati influenzati nelle loro scelte di voto.

Le Filippine sono un esempio emblematico di come la manipolazione mediatica possa influire sui processi democratici. Katie Harbath (2018) le ha definite "paziente zero" delle *fake news*, a causa dell'intenso utilizzo da parte dei candidati e delle strategie adottate. La campagna presidenziale del 2016, definita da Fallorina et al. (2023) come la prima "social media election", ha dimostrato come

la manipolazione mediatica possa inoltre trasformarsi in un elemento chiave per il successo. Durante la corsa presidenziale, le *fake news* hanno provocato uno spostamento del 5,5% delle preferenze da Leni Robredo a favore di Rodrigo Duterte (Prajapati et al., [2024](#)). Le ricerche di Ong et al. ([2019](#)) dimostrano che una quota crescente di risorse viene destinata dai candidati non solo al finanziamento delle campagne online, ma anche per sviluppare tattiche volte a mascherare le conseguenze di eventuali rivelazioni sulle loro operazioni di disinformazione. Nelle Filippine, infatti, è stata creata una vera e propria industria della disinformazione, basata sulla micro-segmentazione operata da un team di esperti. Questi utilizzano le proprie competenze per creare assiduamente *meme* e *slogan* allo scopo di “intossicare” Internet (Combinido & Curato, [2021](#)). La portata di questo fenomeno è stata pienamente rivelata dalla vittoria di Ferdinand “Bongbong” Marcos Junior alle elezioni presidenziali del 2022. La campagna di disinformazione non era finalizzata esclusivamente al successo elettorale, mirava bensì a una riscrittura sistematica della storia della famiglia Marcos e dell’intera nazione. “Bongbong” si è impegnato a ripulire l’immagine del padre, Ferdinand Marcos Sr., capo di un governo autoritario durato 21 anni e segnato da gravi violazioni dei diritti umani ([Ucci, 2022](#)). Per attuare questa strategia, il target principale delle *fake news* è stato quello della popolazione sotto i 40 anni, priva di memoria storica diretta e quindi più facile da manipolare. La campagna di disinformazione si è spostata così da social più tradizionali quali Facebook a TikTok, dove i contenuti celebravano l’ex dittatore come il “miglior presidente del mondo”. Non sorprende quindi che oltre il 70% degli elettori di età compresa tra i 18 e i 24 anni, nativi digitali e principali utilizzatori delle piattaforme social, si sia dichiarato sostenitore di Marcos Jr ([Alù, 2024](#)).

I *deepfake*: la nuova frontiera della disinformazione

L'avvento dell'Intelligenza Artificiale (IA) ha reso la disinformazione ancora più insidiosa, consentendo la creazione di immagini, audio e video falsificati noti come *deepfake*. Nati su Reddit nel 2017 come video satirici, questi contenuti hanno sin subito mostrato il loro lato pericoloso con la creazione di video hard. Da allora, grazie soprattutto allo sviluppo delle reti neurali (Deep Neural Network) questi video sono divenuti più facili da creare e maggiormente credibili (Arruzzoli, [2022](#)). I *deepfake* possono essere classificati in base alla loro precisione e agli strumenti utilizzati per realizzarli. Tra le tecniche di creazione più comuni, Smith and Mansted ([2020](#)) elencano il *face swapping* (sovraposizione del volto di una persona su quello di un'altra), il *re-enactment* (sovraposizione del volto di una persona su quello di un utente, consentendo di manipolare i suoi movimenti ed espressioni facciali), il *lip syncing* (sincronizzazione dei movimenti delle labbra), il *motion transfer* (trasferimento dei movimenti da una persona a un'altra), l'*image generation* (creazione di immagini digitali) e l'*audio generation* (creazione di voci sintetizzate a partire da un campione vocale). Questi strumenti consentono di attribuire a chiunque parole e azioni mai accadute, con rischi maggiori per i personaggi pubblici, come i politici, di cui esistono numerose immagini facilmente reperibili (Monti, [2024](#)). Come dichiarato nel *deepfake* creato *ad hoc* dall'attore Jordan Peele in collaborazione con BuzzFeed in cui simula Barack Obama, la sfera politica resta l'ambito più rischioso di applicazione (King, [2023](#)). Secondo Weikmann and Lecheler ([2022](#)), la pericolosità della disinformazione visiva risiede nell'impatto psicologico che questa ha sugli utenti. Le immagini, infatti, possono essere richiamate più facilmente alla mente rispetto ai testi, inoltre i video hanno la capacità di sovrascrivere i ricordi preesistenti. Sebbene alcuni studiosi paventino scenari catastrofici come il “collasso della realtà” (Habgood-Coote, [2023](#)), Vaccari and Chadwick ([2020](#)) sottolineano un rischio più tangibile: l'incertezza generata dai *deepfake* può erodere la fiducia nelle notizie incontrate sui social, favorendo la polarizzazione e rendendo il dibattito insostenibile. Politici disonesti potrebbero sfruttare tale incertezza per respingere qualsiasi accusa, sostenendo che nulla può essere verificato. Un esempio concreto di questo fenomeno, definito da Schiff et al. (2024) come “vantaggio del bugiardo” è il caso del Ministro delle Finanze indiano Thiagarajan, il quale ha negato l'autenticità di due audio scandalo in cui accusava i membri del suo partito di attività illecite. Tuttavia, l'analisi condotta da tre esperti di tecnologia *deepfake* ha confermato che il secondo audio era autentico, mentre il primo potrebbe essere stato manomesso (Christopher & Sengupta, [2023](#)).

Il 2024 è stato un anno di prova per i *deepfake*, poiché segnato da elezioni in vari Paesi. Tra questi ci sono gli Stati Uniti, dove gli esperti hanno temuto che l'uso dell'IA applicata alla generazione di disinformazione potesse addirittura alterare i risultati (Econopoly, 2024). Sebbene ciò non si sia verificato, secondo Verma et al. (2024) l'IA ha eroso ulteriormente la fiducia degli elettori. Sin da gennaio, i *deepfake* a tema politico hanno iniziato a circolare: tra questi sono diventati virali quello in cui Joe Biden esortava gli elettori a non votare alle primarie e quello di Kamala Harris che festeggiava il ritiro di Biden dalla candidatura. Altro caso emblematico è l'accusa di Trump a Harris di aver utilizzato l'IA per creare una folla finta in un aeroporto, quando in realtà quella folla era reale (Frasso, 2024).

Il problema che sorge è quindi se si possa o meno credere a ciò che si vede. Sebbene l'alfabetizzazione sull'argomento possa sembrare la soluzione più immediata al problema, un sondaggio condotto da Ternovski et al. (2022) mostra che informare gli elettori circa l'aumento dei *deepfake* non li rende più capaci di identificarli, bensì mina ulteriormente la loro fiducia. Durante l'esperimento, infatti, le informazioni fornite riguardo ai *deepfake* hanno indotto i partecipanti a non credere nemmeno ai video che in realtà non erano stati manipolati. A livello tech, si stanno sviluppando diversi algoritmi basati sul *deep learning* che cercano di contrastare i *deepfake* partendo proprio da come vengono creati. Tuttavia, questi modelli riconoscono solo tipi di manipolazione già conosciute, rischiando di etichettare come reali i *deepfake* creati con tecniche nuove. Per questo motivo, sono in fase di sviluppo tecnologie sempre più avanzate di identificazione come il riconoscimento biometrico e nuove Intelligenze Artificiali capaci di riconoscere le emozioni degli individui (Verdoliva, 2024).

Conclusioni

L'analisi di alcuni casi reali e di articoli della letteratura riguardante *fake news* e *deepfake* ha confermato che la disinformazione rappresenta una sfida rilevante nella società contemporanea, ponendosi non solo come un problema connesso alla distorsione della percezione della realtà da parte dei cittadini ma anche come una potenziale minaccia per i processi democratici. In particolare, l'evoluzione dei *deepfake* si configura come un ambito particolarmente insidioso: nuove tecniche di creazione emergono quotidianamente, rendendoli difficili da identificare e complicandone il contrasto. Sebbene il timore di un completo “collasso della realtà” (Habgood-Coote, [2023](#)) possa sembrare uno scenario remoto e potenzialmente irrealizzabile grazie anche all'adattabilità degli utenti alle innovazioni, è fondamentale non solo informare e alfabetizzare la popolazione riguardo ai disturbi dell'informazione ma sviluppare anche soluzioni che affrontino il problema alla radice. Come emerge dai dati, infatti, i social media sono ormai parte integrante della vita quotidiana della maggior parte della popolazione globale e necessitano, quindi, di una regolamentazione che ne garantisca trasparenza e affidabilità, evitando così un'ulteriore erosione della fiducia negli utenti. È dunque cruciale ripulire l'informazione consentendo ai cittadini di poterla considerare nuovamente attendibile, evitando che il “vantaggio del bugiardo” (Schiff et al., [2024](#)) si rafforzi e prevenendo l'espansione dei problemi di accountability. Per proteggere i valori fondamentali della democrazia e della libertà di pensiero è necessario quindi adottare un approccio multilivello che integri soluzioni a livello tecnologico, campagne d'informazione e regolamentazione nazionale. Si tratta quindi di sviluppare una risposta efficace che impedisca a chi detiene il potere di sfruttare la vulnerabilità dei destinatari dell'informazione.

Bibliografia

- Aïmeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining*, 13(1). <https://doi.org/10.1007/s13278-023-01028-5>
- Alù, A. (2024, May 22). *Elezioni filippine: il ruolo dei social e il peso della generazione Z nel ritorno al potere dei Marcos*. Agenda Digitale. Retrieved December 15, 2024, from <https://www.agendadigitale.eu/cultura-digitale/elezioni-filippine-e-social-network-il-ruolo-dei-social-e-il-peso-della-generazione-z-nel-ritorno-al-potere-dei-marcos/>
- Arruzzoli, F. (2022, January). Deepfake & Cyber Intelligence - ICT Security Magazine. *ICT Security Magazine*. Retrieved December 16, 2024, from <https://www.ictsecuritymagazine.com/pubblicazioni/deepfake-cyber-intelligence/>
- Bradshaw, S., & Howard, P. N. (2018). Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation. *The Computational Propaganda Project*, 28(4), 18–34. <https://demtech.ox.ac.uk/wp-content/uploads/sites/12/2018/07/ct2018.pdf>
- Christopher, N., & Sengupta, D. M. (2023, July 6). An Indian politician says scandalous audio clips are AI deepfakes. We had them tested. *Rest of World*. <https://restofworld.org/2023/indian-politician-leaked-audio-ai-deepfake/>
- Colomina, C., Sánchez Margalef, H., & Youngs, R. (2021). The impact of disinformation on democratic processes and human rights in the world. *European Parliament*, QA-02-21-559-EN-N (pdf). <https://doi.org/10.2861/59161>
- Combinido, P., & Curato, N. (2021). Curing “Patient Zero”: Reclaiming the Digital Public Sphere in the Philippines. In A. Sinpeng & R. Tapsell (Eds.). In *From Grassroots Activism to Disinformation* (pp. 19–42). ISEAS–Yusof Ishak Institute. <https://www.cambridge.org/core/books/abs/from-grassroots-activism-to-disinformation/curing-patient-zero-reclaiming-the-digital-public-sphere-in-the-philippines/8441197737AA86A8C47A8806BB1F4132>
- Econopoly. (2024, June 7). *Cyber attacchi e deepfake, quanto sono a rischio le elezioni?* Econopoly. https://www.econopoly.ilsole24ore.com/2024/06/07/cyber-attacchi-deepfake-elezioni/?refresh_ce=1
- FAKE NEWS definition and meaning | Collins English Dictionary. (2024). In *Collins Dictionaries*. <https://www.collinsdictionary.com/dictionary/english/fake-news>

Fallorina, R., Hall Lanuz, J. M., Felix, J. G., Sanchez II, F., Ong, J. C., & Curato, N. (2023). From Disinformation to Influence Operations: The Evolution of Disinformation in Three Electoral Cycles. In *Internews*. Retrieved December 9, 2024, from <https://internews.org/resource/from-disinformation-to-influence-operations-the-evolution-of-disinformation-in-three-electoral-cycles/>

Frasso, E. (2024, August 14). *Donald Trump: “La folla che ha accolto Kamala Harris creata con AI”. Ma non è vero.* AI News. <https://ainews.it/donald-trump-la-folla-che-ha-accolto-kamala-harris-creata-con-ai-ma-non-e-vero/>

Habgood-Coote, J. (2023). Deepfakes and the epistemic apocalypse. *Synthese*, 201(3). <https://doi.org/10.1007/s11229-023-04097-3>

Kemp, S. (2024, October 23). *The global state of Digital in 2024 — DataReportal – Global Digital Insights*. DataReportal – Global Digital Insights. <https://datareportal.com/reports/digital-2024-october-global-statshot>

King, C. (2023, October 31). *What are deepfakes and how are they impacting society?* Edinburgh Impact | the University of Edinburgh. <https://impact.ed.ac.uk/opinion/what-are-deepfakes-and-how-are-they-impacting-society/>

Maddalena, G., & Gili, G. (2018). Post-verità e fake news: radici, significati attuali, inattesi protagonisti e probabili vittime. *Media Education*, 9(1), 1–16. <https://iris.unimol.it/handle/11695/77750>

Monti, M. (2024, May 27). *Deepfake, i rischi “politici”: per la democrazia e l’informazione online.* Agenda Digitale. Retrieved December 16, 2024, from <https://www.agendadigitale.eu/cultura-digitale/deepfake-i-pericoli-politici-per-la-democrazia-e-linformazione-online/>

Ong, J., Tapsell, R., & Curato, N. (2019). Tracking Digital Disinformation in the 2019 Philippine Midterm Election. *New Mandala*. <https://researchprofiles.anu.edu.au/en/publications/tracking-digital-disinformation-in-the-2019-philippine-midterm-el>

Prajapati, A., Kumar, G., & Srivastava, M. (2024). The role of fake news in political campaigns and elections: A global perspective. *International Journal for Multidisciplinary Research*, 6(6). <https://doi.org/10.36948/ijfmr.2024.v06i06.30907>

Schiff, K. J., Schiff, D. S., & Bueno, N. S. (2024). The Liar’s dividend: Can politicians claim misinformation to evade accountability? *American Political Science Review*, 1–20. <https://doi.org/10.1017/s0003055423001454>

- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media. *ACM SIGKDD Explorations Newsletter*, 19(1), 22–36. <https://doi.org/10.1145/3137597.3137600>
- Silverman, C. (2016, November 16). This analysis shows how viral fake election news stories outperformed real news on Facebook. *BuzzFeed News*. <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook>
- Smith, H., & Mansted, K. (2020). Weaponised deep fakes: national security and democracy. *ASPI*, 28/2020, 6–9. <https://apo.org.au/node/303485>
- Soll, J. (2016, December 18). The Long and Brutal History of Fake News. *POLITICOMAGAZINE*. <https://www.politico.com/magazine/story/2016/12/fake-news-history-long-violent-214535/>
- Ternovski, J., Kalla, J., & Aronow, P. (2022). Negative Consequences of Informing Voters about Deepfakes: Evidence from Two Survey Experiments. *Journal of Online Trust and Safety*, 1(2). <https://doi.org/10.54501/jots.v1i2.28>
- Ucci, M. (2022, May 18). La disinformazione riporta i Marcos al potere. *LUISS*. <https://zetaluiss.it/2022/05/18/elezioni-filippine-disinformazione-marcos/>
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1). <https://doi.org/10.1177/2056305120903408>
- Verdoliva, L. (2024, May 21). *Deepfake, come affrontare attacchi sempre nuovi: evoluzioni e sfide future*. Agenda Digitale. Retrieved December 16, 2024, from <https://www.agendadigitale.eu/cultura-digitale/deepfake-come-affrontare-attacchi-sempre-nuovi-evoluzioni-e-sfide-future/>
- Verma, P., Oremus, W., & Zakrzewski, C. (2024, November 9). AI didn't sway the election, but it deepened the partisan divide. *The Washington Post*. <https://www.washingtonpost.com/technology/2024/11/09/ai-deepfakes-us-election/>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Wakefield, J. (2016, June 14). *Social media “outstrips TV” as news source for young people*. BBC News. Retrieved December 7, 2024, from <https://www.bbc.com/news/uk-36528256>

Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. In *Council of Europe Portal*. Retrieved December 7, 2024, from <https://edoc.coe.int/en/media/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html#>

Weikmann, T., & Lecheler, S. (2022). Visual disinformation in a digital age: A literature synthesis and research agenda. *New Media & Society*, 25(12), 3696–3713. <https://doi.org/10.1177/14614448221141648>