

2/2025 novembre

Rivista di diritto dei media

ISSN 2532-9146

DIRETTORE RESPONSABILE

Oreste Pollicino

DIRETTORI

Marco Bassini

Marina Castellaneta

Carlo Melzi d'Eril

Giulio Enea Vigevani

EDITORE

Carlo Melzi d'Eril

REDAZIONE

Cordinamento editoriale

Rebecca Bovatti (Università di Milano)

Martina Cazzaniga (Università di Milano-Bicocca)

Giulia Napoli (Università di Milano-Bicocca)

Componenti

Flavia Bavetta (Università Bocconi – Milano)

Amalia Mariana Buzura (Avvocato in Milano)

Teresa Catalano (Università di Bari)

Federico Ceci (Università di Bari)

Marco Cecili (Università di Firenze)

Alessia Forte (Università di Milano)

Maria Galbusera (Università di Milano – Bicocca)

Giuseppe Gallo (Università di Messina)

Giovana Lopes (Tilburg University)

Matteo Monti (Università Bocconi – Milano)

Giuseppe Muto (Università Bocconi – Milano)

Federica Paolucci (Università Bocconi – Milano)

Stefania Rutigliano (Università di Bari)

SEDE**CONTACTS**

Studio legale Melzi d'Eril Vigevani

Via San Barnaba 32 - 20122 Milano

e-mail: redazione@rivistadidirittodeimedia.it

ANNO 2025 - CODICE ISSN 2532-9146

Impaginazione a cura di Chiara Pavesi

COMITATO SCIENTIFICO

Ernesto Apa <i>Avvocato in Roma</i>	Alberto Maria Gambino <i>Università Europea</i>	Marina Pietrangelo <i>ITTIG-CNR</i>
Fabio Basile <i>Università di Milano</i>	Gianluca Gardini <i>Università di Ferrara</i>	Giovanni Pitruzzella <i>Corte costituzionale</i>
Mirzia Bianca <i>Sapienza – Università di Roma</i>	Glauco Giostra <i>Sapienza – Università di Roma</i>	Francesco Pizzetti <i>Università di Torino</i>
Raffaele Bifulco <i>LUISS Guido Carli</i>	Ottavio Grandinetti <i>Avvocato in Roma</i>	Andrea Pugiotto <i>Università di Ferrara</i>
Francesca Biondi <i>Università di Milano</i>	Enrico Grosso <i>Università di Torino</i>	Margherita Ramajoli <i>Università di Milano</i>
Carlo Blengino <i>Avvocato in Torino</i>	Antonio Gullo <i>LUISS Guido Carli</i>	Sofia Ranchordás <i>LUISS Guido Carli</i>
Manfredi Bontempelli <i>Università di Milano</i>	Antonio Iannuzzi <i>Università Roma Tre</i>	Giorgio Resta <i>Università Roma Tre</i>
Elda Brogi <i>European University Institute</i>	Simone Lonati <i>Università Bocconi</i>	Giovanni Maria Riccio <i>Università di Salerno</i>
Giuseppe Busia <i>Autorità Nazionale Anticorruzione</i>	Nicola Lupo <i>LUISS Guido Carli</i>	Gianpaolo Maria Ruotolo <i>Università di Foggia</i>
Licia Califano <i>Università di Urbino, già Garante la protezione dei dati personali</i>	Enrico Maria Mancuso <i>Università Cattolica</i>	Lucia Scaffardi <i>Università di Parma</i>
Angelo Marcello Cardani <i>Università Bocconi, già Autorità per le garanzie nelle comunicazioni</i>	Vittorio Manes <i>Università di Bologna</i>	Salvatore Scuto <i>Avvocato in Milano</i>
Marta Cartabia <i>Università Bocconi, Presidente emerito della Corte costituzionale</i>	Michela Manetti <i>Università di Siena</i>	Sergio Seminara <i>Università di Pavia</i>
Corrado Caruso <i>Università di Bologna</i>	Roberto Mastroianni <i>Università di Napoli “Federico II”</i>	Salvatore Sica <i>Università di Salerno</i>
Massimo Ceresa-Gastaldo <i>Università Bocconi</i>	Alessandro Melchionda <i>Università di Trento</i>	Pietro Sirena <i>Università Bocconi</i>
Adolfo Ceretti <i>Università di Milano – Bicocca</i>	Luigi Montuori <i>Garante per la protezione dei dati personali</i>	Bruno Tonoletti <i>Università di Pavia</i>
Marco Cuniberti <i>Università di Milano</i>	Donatella Morana <i>Università di Roma Tor Vergata</i>	Luisa Torchia <i>Università Roma Tre</i>
Marilisa D’Amico <i>Università di Milano</i>	Antonio Nicita <i>LUMSA, già Autorità per le garanzie nelle comunicazioni</i>	Gianluca Varraso <i>Università Cattolica</i>
Filippo Danovi <i>Università di Milano – Bicocca</i>	Marco Orofino <i>Università di Milano</i>	Francesco Viganò <i>Corte costituzionale</i>
Filippo Donati <i>Università di Firenze</i>	Ilja Pasquali Cerioli <i>Università di Milano</i>	Luciano Violante <i>Futuri Probabili</i>
Michele Faioli <i>Università Cattolica</i>	Paolo Passaglia <i>Università di Pisa</i>	Lorenza Violini <i>Università di Milano</i>
Giusella Finocchiaro <i>Università di Bologna</i>	Davide Petrini <i>Università di Torino</i>	Roberto Zaccaria <i>Università di Firenze</i>
Maurizio Fumoglia <i>Suprema Corte di Cassazione</i>	Vincenzo Pezzella <i>Suprema Corte di Cassazione</i>	Nicolò Zanon <i>Università di Milano</i>
		Vincenzo Zeno-Zencovich <i>Università Roma Tre</i>

COMITATO DEGLI ESPERTI PER LA VALUTAZIONE

Maria Romana Allegri <i>Sapienza – Università di Roma</i>	Giovanna De Minico <i>Università di Napoli “Federico II”</i>	Stefania Ninatti <i>Università di Milano – Bicocca</i>
Giulio Allevato <i>IE University</i>	Federica De Simone <i>Università della Campania</i>	Francesco Paolo Micozzi <i>Avvocato in Cagliari</i>
Benedetta Barbisan <i>Università di Macerata</i>	Gabriele Della Morte <i>Università Cattolica</i>	Piergiuseppe Otranto <i>Università di Bari</i>
Marco Bellezza <i>Dipartimento per la trasformazione digitale</i>	Luca Di Majo <i>Università della Campania</i>	Omar Makimov Pallotta <i>Università di Macerata</i>
Marco Biasi <i>Università di Milano</i>	Federica Fabrizzi <i>Sapienza – Università di Roma</i>	Anna Papa <i>Università di Napoli “Parthenope”</i>
Daniela Bifulco <i>Corte di cassazione</i>	Fernanda Faini <i>Università Pegaso</i>	Francesca Paruzzo <i>Università di Torino</i>
Elena Bindi <i>Università di Siena</i>	Fabio Ferrari <i>Università di Verona</i>	Irene Pellizzone <i>Università di Milano</i>
Paolo Bonetti <i>Università di Milano – Bicocca</i>	Roberto Flor <i>Università di Verona</i>	Sabrina Peron <i>Avvocato in Milano</i>
Monica Bonini <i>Università di Milano – Bicocca</i>	Giulia Formici <i>Università di Parma</i>	Federico Gustavo Pizzetti <i>Università di Milano</i>
Fernando Bruno <i>Autorità per le garanzie nelle comunicazioni</i>	Federico Furlan <i>Università di Milano – Bicocca</i>	Benedetto Ponti <i>Università di Perugia</i>
Daniele Butturini <i>Università di Verona</i>	Alessandra Galluccio <i>Università di Milano</i>	Augusto Preta <i>International Institute of Communications</i>
Simone Calzolaio <i>Università di Macerata</i>	Giovanni Battista Gallus <i>Avvocato in Cagliari</i>	Federico Riboldi <i>Avvocato in Milano</i>
Quirino Camerlengo <i>Università di Pavia</i>	Emma Imperato <i>Università di Napoli “L’Orientale”</i>	Francesca Rosa <i>Università di Urbino</i>
Nicola Canzian <i>Università di Milano – Bicocca</i>	Elisabetta Lamarque <i>Università di Milano – Bicocca</i>	Marianna Russo <i>Università della Campania</i>
Marina Caporale <i>Università di Modena e Reggio Emilia</i>	Erik Longo <i>Università di Firenze</i>	Lucia Sciannella <i>Università di Teramo</i>
Andrea Cardone <i>Università di Firenze</i>	Valerio Lubello <i>Università Bocconi</i>	Monica Alessia Senor <i>Autorità garante per la protezione dei dati personali</i>
Bruno Carotti <i>LUISS Guido Carli</i>	Claudia Massa <i>Università di Napoli “Federico II”</i>	Silvio Sonnatì <i>Università Bocconi</i>
Stefano Catalano <i>Università di Verona</i>	Antonio Manganelli <i>Università di Siena</i>	Stefania Stefanelli <i>Università di Perugia</i>
Pier Giorgio Chiara <i>Università di Bologna</i>	Paola Marsocci <i>Sapienza – Università di Roma</i>	Giulia Tiberi <i>Università dell’Insubria</i>
Salvatore Curreri <i>Università di Enna “Kore”</i>	Claudio Martinelli <i>Università di Milano – Bicocca</i>	Luca Vanoni <i>Università di Milano</i>
Monica Delsignore <i>Università di Milano – Bicocca</i>	Alberto Mattiacci <i>Sapienza – Università di Roma</i>	Scilla Vernile <i>Università di Milano – Bicocca</i>
Giovanni De Gregorio <i>Católica Global School of Law</i>	Massimiliano Mezzanotte <i>Università di Teramo</i>	Pietro Villaschi <i>Università di Milano</i>
	Lorenzo Nannipieri <i>CNR-ITTIG</i>	Paolo Zicchittu <i>Università di Milano – Bicocca</i>
	Costanza Nardocci <i>Università di Milano</i>	

INTERNATIONAL ADVISORY BOARD

Shulamit Almog
University of Haifa

Max Baumgart
Universität Luzern

Edoardo Celeste
Dublin City University

Agustina Del Campo
*Center for Studies on Freedom of Expression
and Access to Information*

Elaine Fahey
City Law School, University of London

Irene Kamara
Tilburg University

Uta Kohl
University of Southampton

Krystyna Kowalik-Bańczyk
Tribunale dell'Unione europea

Michale Geist
University of Ottawa

Thomas Margoni
KU Leuven

Christopher Marsden
Monash University

Manuel D. Masseno
Instituto Politécnico de Beja

Dawn C. Nunziato
George Washington University

Miquel Pequera
Universitat Oberta de Catalunya

Marius Dragomir
Central European University

Andrej Savin
Copenhagen Business School

Bilyana Petkova
University of National and World Economy

Thomas Wischmeyer
Universität Bielefeld

MediaLaws - Rivista di diritto dei media è una rivista quadrimestrale telematica, ad accesso libero, che si propone di pubblicare saggi, note e commenti attinenti al diritto dell'informazione italiano, comparato ed europeo.

La rivista nasce per iniziativa di Oreste Pollicino, Giulio Enea Vigevani, Carlo Melzi d'Eril e Marco Bassini e raccoglie le riflessioni di studiosi, italiani e stranieri, di diritto dei media.

I contributi sono scritti e ceduti a titolo gratuito e senza oneri per gli autori. Essi sono attribuiti dagli autori con licenza Creative Commons "Attribuzione – Non commerciale 3.0" Italia (CC BY-NC 3.0 IT). Sono fatte salve, per gli aspetti non espressamente regolati da tale licenza, le garanzie previste dalla disciplina in tema di protezione del diritto d'autore e di altri diritti connessi al suo esercizio (l. 633/1941).

Il lettore può utilizzare i contenuti della rivista con qualsiasi mezzo e formato, per qualsiasi scopo lecito e non commerciale, nei limiti consentiti dalla licenza Creative Commons "Attribuzione – Non commerciale 3.0 Italia" (CC BY-NC 3.0 IT), in particolare menzionando la fonte e, laddove necessario a seconda dell'uso, conservando il logo e il formato grafico originale.

La rivista fa proprio il Code of Conduct and Best Practice Guidelines for Journal Editors elaborato dal COPE (Committee on Publication Ethics).

La qualità e il rigore scientifici dei saggi della Rivista sono garantiti da una procedura di *double-blind peer review* affidata a un comitato di esperti per la valutazione individuato secondo criteri di competenza e rotazione e aggiornato ogni anno.

MediaLaws - Rivista di diritto dei media

Regolamento per la pubblicazione dei contributi

1. “MediaLaws – Rivista di diritto dei media” è una rivista telematica e ad accesso aperto che pubblica con cadenza quadrimestrale contributi attinenti al diritto dell’informazione.
2. Gli organi della rivista sono il Comitato di direzione, il Comitato scientifico e il Comitato degli esperti per la valutazione. L’elenco dei componenti del Comitato di direzione e del Comitato scientifico della rivista è pubblicato sul sito della stessa (rivista.medialaws.eu). Il Comitato degli esperti per la valutazione è sottoposto ad aggiornamento una volta l’anno.
3. La rivista si compone delle seguenti sezioni: “Saggi”, “Note a sentenza” (suddivisa in “Sezione Europa”, “Sezione Italia” e “Sezione comparata”), “Cronache e commenti” e “Recensioni e riletture”. I singoli numeri potranno altresì ospitare, in via d’eccezione, contributi afferenti a sezioni diverse.
4. La sezione “Saggi” ospita contributi che trattano in maniera estesa e approfondita un tema di ricerca, con taglio critico e supporto bibliografico.
5. La sezione “Note a sentenza” ospita commenti alle novità giurisprudenziali provenienti dalle corti italiane, europee e straniere.
6. La sezione “Cronache e commenti” ospita commenti a questioni e novità giuridiche di attualità nella dimensione nazionale, europea e comparata.
7. La sezione “Recensioni e riletture” ospita commenti di opere rispettivamente di recente o più risalente pubblicazione.
8. La richiesta di pubblicazione di un contributo è inviata all’indirizzo di posta elettronica submissions@medialaws.eu, corredata dei dati, della qualifica e dei recapiti dell’autore, nonché della dichiarazione che il contributo sia esclusiva opera dell’autore e, nel caso in cui lo scritto sia già destinato a pubblicazione, l’indicazione della sede editoriale.
9. La direzione effettua un esame preliminare del contributo, verificando l’attinenza con i temi trattati dalla rivista e il rispetto dei requisiti minimi della pubblicazione.
10. In caso di esito positivo, la direzione procede ad assegnare il contributo alla sezione opportuna.
11. I saggi sono inviati alla valutazione, secondo il metodo del doppio cieco, di revisori scelti dall’elenco degli esperti per la valutazione della rivista secondo il criterio della competenza, della conoscenza linguistica e della rotazione. I revisori ricevono una scheda di valutazione, da consegnare compilata alla direzione entro il termine da essa indicato. Nel caso di tardiva o mancata consegna della scheda, la direzione si riserva la facoltà di scegliere un nuovo revisore. La direzione garantisce l’anonimato della valutazione.
12. La direzione comunica all’autore l’esito della valutazione.
Se entrambe sono positive, il contributo è pubblicato.
Se sono positive ma suggeriscono modifiche, il contributo è pubblicato previa revisione dell’autore, in base ai commenti ricevuti, e verifica del loro accoglimento da parte della direzione. La direzione si riserva la facoltà di sottoporre il contributo così come modificato a nuova valutazione, anche interna agli organi della rivista. Se solo una valutazione è positiva, con o senza modifiche, la direzione si riserva la facoltà di trasmettere il contributo a un terzo valutatore. Se entrambe le valutazioni sono negative, il contributo non viene pubblicato.
13. Per pubblicare il contributo, l’Autore deve inviare una versione definitiva corretta secondo le regole editoriali della rivista pubblicate sul sito della stessa, un abstract in lingua italiana e inglese e un elenco di cinque parole chiave. Il mancato rispetto dei criteri editoriali costituisce motivo di rigetto della proposta.
14. Le valutazioni vengono archiviate dalla direzione della rivista per almeno tre anni.
15. A discrezione della direzione, i saggi di autori di particolare autorevolezza o richiesti dalla direzione possono essere pubblicati senza essere sottoposti alla procedura di referaggio a doppio cieco ovvero essere sottoposti a mero referaggio anonimo, previa segnalazione in nota.

Focus

The Venice Commission Doctrine on Digital Technologies and the Rule of Law

a cura di Davide Paris

- 14 Introduction
- 15 The Venice Commission's Digital Doctrine: Harmonising Law, Technology, and Democracy
Rafael Rubio Núñez
- 52 Cybersecurity in the Electoral Sphere: Legal and Technical Considerations on a Hybrid Threat
Tamara Álvarez Robles
- 73 Information Disorders in Elections and Their Impact on the Principles of Democracy
Janine M. Otálora Malassis
- 91 The Role of the Venice Commission in Establishing a Clear Legal Framework on Digital Technologies and Democracy
Edoardo Alberto Rossi
- 106 The Right to Free Elections in the TikTok Era: Towards Digital Expertise of the Venice Commission?
Raphaël Déchaux

Saggi

- 129 Lo *shadowban* dei contenuti politici e i suoi effetti sulle campagne elettorali online
Alberto Di Chiara
- 156 Il *Digital Welfare state in Europa*: tensioni costituzionali e prospettive evolutive
Elena A. Ferioli

177 Better Law-making and Evaluation for the EU Digital Rulebook

Marco Bassini, Mariateresa Maggiolino,
Alexandre de Streel

224 Libertà di manifestazione del pensiero online e content moderation dei nuovi poteri privati: rischi ed opportunità alla luce della nuova co-regolazione euro-unitaria

Francesca Donà

266 Gli *Executive Orders* in materia di *Artificial Intelligence*: amministrazione Biden e Trump a confronto

Elisabetta Gherri

Focus

La sentenza n. 13200/2025 delle Sezioni Unite della Cassazione e la diffamazione nell'ambito della cronaca giudiziaria

297 Le Sezioni Unite civili alle prese con l'erronea attribuzione da parte del giornalista della qualifica di imputato anziché di indagato e/o del reato consumato in luogo di quello tentato. È diffamazione?

Riccardo D'Auro

314 La diffamazione a mezzo stampa nell'ambito della cronaca giudiziaria: un "faro acceso" sulla domanda di informazione

Giorgio Sichera

329 La cronaca giudiziaria e gli errori nella corretta rappresentazione dello status giuridico: quando sussiste la diffamazione

Sabrina Peron

Note

- 341 Incitamento all'odio contro le persone LGBTI, diritto al rispetto della vita privata e libertà di stampa: la pronuncia resa dalla Corte di Strasburgo nel caso *Minasyan et al. c. Armenia***
Federica Falconi
- 367 Diffamazione, *online hate speech* e odio religioso**
Luciana Goisis
- 376 Diffamazione *online*: lo “strano caso” dei video trasmessi su “Tik Tok” tramite la modalità della “diretta”**
Emma Ludovica Breda

Cronache

- 393 La legislazione delegata in materia di intelligenza artificiale: la costruzione di una disciplina organica al confine tra scelte gover-native, controllo parlamentare e vincoli europei**
Oreste Pollicino, Giuseppe Muto
- 415 Dalla colpa al codice. Ripensare la responsabilità nell'era dell'intelligenza artificiale**
Enrico Maestri, Carlotta Mognato
- 430 Dai *Dark Patterns* agli *Hyper-Engaging Dark Patterns*: per un'applicazione moderna del divieto ex art. 25 DSA**
Alessio Scaffidi, Vincenzo Forte
- 443 Federated Learning: A Tangible Solution for GDPR Compliance?**
Matteo Johannes Bürgler

Focus

The Venice Commission Doctrine on Digital Technologies and the Rule of Law

edited by Davide Paris

14 Introduction

15 The Venice Commission's Digital Doctrine: Harmonising Law, Technology, and Democracy

Rafael Rubio Núñez

52 Cybersecurity in the Electoral Sphere: Legal and Technical Considerations on a Hybrid Threat

Tamara Álvarez Robles

73 Information Disorders in Elections and Their Impact on the Principles of Democracy

Janine M. Otálora Malassis

91 The Role of the Venice Commission in Establishing a Clear Legal Framework on Digital Technologies and Democracy

Edoardo Alberto Rossi

106 The Right to Free Elections in the TikTok Era: Towards Digital Expertise of the Venice Commission?

Raphaël Déchaux

Essays

129 The Shadowbanning of Political Content and Its Effects on Online Electoral Campaigns

Alberto Di Chiara

156 The Digital Welfare State in Europe: Constitutional Frictions and Prospects for Evolution

Elena A. Ferioli

177 Better Law-making and Evaluation for the EU Digital Rulebook

Marco Bassini, Mariateresa Maggiolino,
Alexandre de Streel

224 Freedom of Expression Online and Content Moderation by New Private Powers: Risks and Opportunities in Light of the New EU Co-Regulation

Francesca Donà

266 The Executive Orders on Artificial Intelligence: a comparative analysis of the Biden and Trump Administrations

Elisabetta Gherri

Focus

The Italian Supreme Court Ruling no. 13200/2025 and Defamation in the Context of Judicial Reporting

297 When a Journalist Gets It Wrong: The Italian Supreme Court on Misreporting Someone as a Defendant Rather Than a Suspect. Defamation or Mere Error?

Riccardo D'Auro

314 Defamation through the Press in the Context of Judicial Reporting: A 'Spotlight' on the Demand for Information

Giorgio Sichera

329 Legal Reporting and Misrepresentation of the Legal Status: When is There Defamation?

Sabrina Peron

Case notes

341 Hate Speech Targeting LGBTI People, Respect for Private Life and Press Freedom: The ECtHR judgment in Minasyan et al. c. Armenia case

Federica Falconi

367 Defamation, Online Hate Speech and Religious Hatred

Luciana Goisis

376 Online Defamation: The ‘Strange Case’ of TikTok Live Streaming

Emma Ludovica Breda

Comments

393 Delegated Legislation on Artificial Intelligence: A Comprehensive Frame-work at the Intersection of Governmental Choices, Parliamentary Over-sight, and European Constraints

Oreste Pollicino, Giuseppe Muto

415 Rethinking Liability in the Age of Artificial Intelligence

Enrico Maestri, Carlotta Mognato

430 From Dark Patterns to Hyper-Engaging Dark Patterns: towards a modern application of Article 25 of the DSA

Alessio Scaffidi, Vincenzo Forte

443 Federated Learning: A Tangible Solution for GDPR Compliance?

Matteo Johannes Bürgler

Sono stati sottoposti a referaggio anonimo a doppio cieco i saggi di Alberto Di Chiara, Elena Ferioli, Marco Bassini - Mariateresa Maggiolino - Alexandre de Streel, Francesca Donà ed Elisabetta Gherri. Sono stati sottoposti a referaggio anonimo i saggi di Raphaël Déchaux, Janine M. Otálora Malassis, Rafael Rubio Núñez, Tamara Álvarez Robles, Edoardo Alberto Rossi e le note e cronache di Emma Ludovica Breda, Matteo Burgler, Riccardo D'Auro, Federica Falconi, Vincenzo Forte e Alessio Scaffidi, Luciana Goisis, Enrico Maestri, Carlotta Mognato, Giuseppe Muto e Oreste Pollicino, Sabrina Peron, Giorgio Sicchera.

Lo *shadowban* dei contenuti politici e i suoi effetti sulle campagne elettorali *online**

Alberto Di Chiara

Sommario

1. L'introduzione di uno *shadowban* dei contenuti politici da parte del Gruppo Meta – 2. Lo *shadowban* nei *social network*: un problema anzitutto definitorio – 3. *Shadowban*: uno scenario di non sole ombre dal punto di vista delle piattaforme di *social network* – 4. Lo *shadowban* e i suoi effetti sul dibattito pubblico – 5. *Shadowban* e lesione dei diritti fondamentali: spunti a partire da una recente sentenza belga – 6. Il divieto di *shadowban* nel DSA – 7. Quale destino per lo *shadowban*? Verso un (auspicabile) superamento, per salvaguardare il dibattito politico e la partecipazione democratica.

1. L'introduzione di uno *shadowban* dei contenuti politici da parte del Gruppo Meta

Le piattaforme *online* offrono infinite possibilità di accesso a notizie, informazioni politiche e commenti, promuovendo allo stesso tempo nuove forme di espressione individuale e di partecipazione politica¹. I *social network* da almeno un decennio hanno assunto il ruolo di fonte principale d'informazione per decine di milioni di persone: è così per il 42% degli italiani, davanti ai siti di news (28%) e alla carta stampata (16%). La predilezione per i *social network* vale soprattutto per la fascia anagrafica 18-24 anni, che «ha molta più probabilità di usare i social media invece di un sito o di una app d'informazione»².

* Il saggio in oggetto è stato realizzato nell'ambito delle attività del PRIN “Normative and Digital Solutions to Counter Threats during National Election Campaigns (RightNets)”, finanziato dall'Unione Europea - NextGenerationEU, Missione 4, Componente 2, Investimento 1.1 bando PRIN 2022 PNRR, P2022MCYCK, CUP D53D23022340001. L'articolo è stato sottoposto, in conformità al regolamento della Rivista, a referaggio “a doppio cieco”.

¹ Cfr. L. Diamond, *In search of democracy*, London, 2015, 132 ss.

² Istat, *La partecipazione politica in Italia, Anno 2019*. Sul medesimo argomento, si veda: R. Borghi, *Italiani sempre meno interessati alla politica. Si informano soprattutto su social e web, poi tv e radio, poco sui giornali*, 2020; Censis, *Sedicesimo rapporto sulla comunicazione. I media e la costruzione dell'identità*, Franco Angeli, 2020; Pew Research Center, *News Platform Fact Sheet*, 17 September 2024; S. Ellerbeck, *Most people get their news online - but many are switching off altogether. Here's why*, World Economic Forum, 9 September 2022.

Tuttavia, come divenuto evidente in progresso di tempo, la realtà si è rivelata più complessa a causa dell'attività di moderazione dei contenuti, in grado di controllarne l'amplificazione o il silenziamento³. Ciascuna di tali attività viene regolata sulla base di algoritmi che organizzano le informazioni e modellano le interazioni tra gli utenti⁴: alcuni studi hanno evidenziato come le preferenze politiche di questi ultimi possono essere influenzate dalla strutturazione dei *feed* nei *social network* che viene effettuata dagli algoritmi⁵, rimuovendo o privilegiando alcuni contenuti⁶. Infatti, è noto come i *news feed* non mostrino i contenuti in ordine cronologico, ma secondo complesse tecniche algoritmiche, selezionando e dando priorità ai *post* in base alle preferenze degli utenti, con la finalità ultima di incrementare i propri introiti pubblicitari⁷.

Sin dai primi mesi del 2021, molti utenti avevano segnalato, poco dopo l'assalto a Capitol Hill, che i propri contenuti su temi politico/sociali ricevevano un numero di visualizzazioni e di interazioni ridotto rispetto alla media abituale dei contenuti su altre tematiche⁸. Critiche analoghe venivano mosse da alcuni attivisti palestinesi sin dal 2021, al punto che alcuni tra questi hanno modificato i titoli dei *post* nella speranza di evitare i meccanismi dell'algoritmo (c.d. "*algospeak*")⁹.

Nel febbraio del 2024, Meta ha dichiarato che avrebbe smesso di consigliare in modo proattivo su Instagram e Threads i *post* di natura politica di *account* non seguiti, dovendosi intendere come tali «i contenuti legati a temi quali le leggi, le elezioni o argomenti di natura sociale»¹⁰: pertanto, essi continueranno ad essere visibili agli utenti che seguono tali profili o pagine, ma per tutti gli altri non compariranno più nei *feed*, nelle

³ J. Donovan-D. Boyd, *Stop the presses? Moving from strategic silence to strategic amplification in a networked media ecosystem*, in *American Behavioral Scientist*, (65)2, 2021, 333–350. Z. Duan-J. Li-J. Lukito-K.C. Yang-F. Chen-D.V. Shah-S. Yang, *Algorithmic agents in the hybrid media system: Social bots, selective amplification, and partisan news about covid-19*, in *Human Communication Research*, (48)3, 2022, 516–542.

⁴ C. Wagner-M. Strohmaier-A. Olteanu-E. Kicman-N. Contractor-T. Eliassi-Rad, *Measuring algorithmically infused societies*, in *Nature*, (7866)595, 2021, 197 ss.

⁵ A.M. Guess-N. Malhotra-J. Pan-P. Barberá-H. Allcott-T. Brown-A. Crespo-Tenorio-D. Dimmery-D. Freelon-M. Gentzkow, et al., *How do social media feed algorithms affect attitudes and behavior in an election campaign?*, in *Science*, (6656)381, 2023, 398; A. Manganeli, *Piattaforme digitali e "social network": fra pluralità degli ordinamenti, pluralismo informativo e potere di mercato*, in *Giur. cost.*, 2, 2023, 883.

⁶ J.E. Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism*, Oxford, 2010; T. Gillespie, *The politics of 'platforms'*, in *New media & society*, (12)3, 2010, 347–364; P. Popiel, *Regulating datafication and platformization: Policy silos and tradeoffs in international platform inquiries*, in *Policy & Internet*, 6 March 2022.

⁷ P. Leersen, *The Soap Box as a Black Box: Regulating Transparency in Social Media Recommender Systems*, in *European Journal of Law and Technology*, (11)2, 2020.

⁸ V. Giannoli, *Meta oscura i contenuti politici, senza dare spiegazioni. L'esperto: "Crollo tra il 70 e l'80% della loro visibilità". Ma la Commissione europea vuole vederci chiaro*, 9 marzo 2024.

⁹ R. Antonucci, *"I miei post sulla Palestina? Invisibili": lo spettro dello 'shadowban' su Facebook, Instagram e TikTok. Le risposte dei colossi social*, 17 ottobre 2023; Human Rights Watch, *Israel/Palestine: Facebook Censors Discussion of Rights Issues*, 8 October 2021.

¹⁰ Instagram, *Rafforziamo il nostro approccio ai contenuti di natura politica su Instagram e Threads*, 9 febbraio 2024.

homepage e, in generale, tra i contenuti consigliati. Tale opzione può essere disattivata dagli utenti, modificando le impostazioni del proprio *account*. Gli *account* per professionisti su Instagram possono consultare la voce “stato dell’*account*” per verificare la propria idoneità a essere consigliati, che varia in base alla pubblicazione recente di contenuti di natura politica. Per essere consigliati di nuovo, sempre da “stato dell’*account*” è possibile modificare o rimuovere *post* recenti, richiedere un controllo se non si concorda con la decisione della piattaforma, oppure smettere di pubblicare questo tipo di contenuti per un periodo di tempo. Va precisato sin d’ora che in questa sede non ci si occuperà del tema – pur rilevante nell’ambito delle campagne elettorali *online* – della procedura recentemente introdotta da Meta per la sponsorizzazione delle inserzioni di carattere politico¹¹.

La decisione di introdurre lo *shadowban* dei contenuti politici è stata motivata da Meta sulla base di due argomenti: da un lato essa sembra costituire una risposta alle accuse rivolte all’azienda negli ultimi anni, secondo cui il gruppo sarebbe stato quanto meno corresponsabile di una serie di disordini a livello mondiale, dallo scandalo Cambridge Analytica all’amplificazione della disinformazione ai tempi della pandemia fino all’assalto a Capitol Hill, estromettendo dunque quei contenuti che potrebbero diventare problematici se diffusi su larga scala; dall’altro lato, intendeva contribuire a rassicurare i legislatori intenzionati a regolamentare le grandi aziende tecnologiche, ritraendosi dal dibattito politico.

In questa sede si cercherà di dimostrare come tale tecnica di moderazione dei contenuti produca effetti rilevanti sulla partecipazione democratica dal punto di vista costituzionalistico, nella prospettiva del disinteresse dei cittadini per il dibattito politico, il cui esito ultimo si manifesta con un astensionismo elettorale preoccupante¹².

A tal fine, il contributo analizza anzitutto le possibili definizioni attribuite al fenomeno in esame e le tipologie in cui esso si manifesta; verranno poi considerate le ragioni addotte dalle piattaforme in ordine alla sua applicazione, evidenziando gli effetti che tale pratica produce sia sul dibattito pubblico, sia sotto il profilo del rispetto dei diritti fondamentali. Infine, ci si concentrerà sulla disciplina introdotta dal DSA, sottolineando i profili di contrasto dello *shadowban* rispetto alla disciplina europea.

Si tratta di un tema la cui indagine presenta alcune peculiarità che ne rendono complesso lo studio. Anzitutto, è una forma di moderazione dei contenuti nei *social network* che rimane occulta, nel senso che le piattaforme non ne informano gli utenti. La mancanza di informazione preventiva (e, conseguentemente, di prove successive) complica il lavoro degli studiosi, in quanto impedisce di ottenere dati comprovati anche attraverso la ripetibilità degli esperimenti, in grado di dimostrare il collegamento tra comportamenti individuali e applicazione dell’algoritmo¹³. In secondo luogo, occorre rilevare che molti sono i fattori esterni capaci di influenzare i risultati di una ricerca sull’argomento, data la presenza di *account* falsi e

¹¹ Gruppo Meta, *Inserzioni su temi sociali, elezioni o politica*.

¹² Cfr. S. Lieto, *Astensionismo populista e rappresentanza democratica*, in *Costituzionalismo.it*, 2, 2024, 8 ss.

¹³ G. Nicholas, *Shedding Light on Shadowbanning*, in *Center for Democracy & Technology*, 6 April 2022, 6.

bot. In terzo luogo, lo studio sui dati dei *social media* è sempre più difficile da trattare a causa delle restrizioni imposte dalla regolamentazione della privacy. Infine, anche il continuo aggiornamento delle piattaforme e dei *social media* rende sempre più complicata la raccolta dei dati attraverso le tecniche di *web scraping*, postulando un continuo aggiornamento di questi ultimi sulla base dell'evoluzione delle piattaforme medesime¹⁴.

Il 30 aprile 2024 la Commissione europea ha annunciato l'apertura di una procedura formale contro Meta per verificare l'eventuale violazione del regolamento europeo sui servizi digitali, o *Digital Services Act* (d'ora in avanti DSA), che vieta lo *shadowban*, definendolo una «retrocessione nel posizionamento o nei sistemi di raccomandazione, oppure nella restrizione dell'accessibilità da parte di uno o più destinatari del servizio o nell'esclusione dell'utente da una comunità online senza che quest'ultimo ne sia consapevole»¹⁵.

Tra i vari profili, di particolare interesse in una prospettiva costituzionalista risulta l'analisi sulla politica adottata dall'azienda sul declassamento dei contenuti politici nei sistemi di diffusione di Instagram e di Facebook. L'indagine della Commissione – i cui esiti ad oggi non risultano essere stati resi noti – pare si stia concentrando in particolare sulla compatibilità di tale politica con gli obblighi di trasparenza, di tutela degli utenti, e di mitigazione dei rischi sul dibattito pubblico e sui procedimenti elettorali.

In Italia, l'imposizione dello *shadowban* da parte del Gruppo Meta è stata oggetto nel maggio 2024 di un esposto all'Agcom, sottoscritto in maniera bipartisan da 43 parlamentari¹⁶.

Anche alcuni membri del Parlamento europeo hanno presentato delle interrogazioni sul tema: la prima, del 2 aprile 2024, muoveva dalla considerazione secondo cui i contenuti di politici riconosciuti vengono sistematicamente retrocessi senza spiegazione, mentre i contenuti di *influencer* o semplici cittadini, anche se contenenti *hate speech* e *fake news*, continuano a risultare suggeriti agli utenti, rendendoli più visibili¹⁷. Viene

¹⁴ K. Jaidka, *Shadowbans on political expression*, in *SSRN Electronic Journal*, January 2023, 5 ss.

¹⁵ European Commission, *Commission opens formal proceedings against Facebook and Instagram under the Digital Services Act*, 30 April 2024.

¹⁶ P.L. Pisa, *L'esposto dei partiti ad Agcom: "Meta limita i contenuti politici"*, 17 maggio 2024.

¹⁷ In tema di contrasto al fenomeno delle *fake news online*, cfr. F. Sciacchitano-A. Panza, *"Fake news" e disinformazione "online": misure internazionali*, in questa Rivista, 1, 2020, 102 ss.; G. Papa, *Libertà di informazione e rilievi critici per il contrasto della diffusione di "fake news"*, in *dirittifondamentali.it*, 2, 2023, 489 ss.; M. Mezzanotte, *"Fake news", "deepfake" e sovranità digitale nei periodi bellici*, in *federalismi.it*, 33, 2022, 44 ss.; S. Flaminio, *Lotta alle "fake news": dallo stato dell'arte a una prospettiva di regolamentazione per il vivere digitale a margine del "Digital Services Act"*, in *Riv. it. inf. dir.*, 2, 2022, 75 ss.; C. Colapietro, *Libera manifestazione del pensiero, fake news e privacy, oggi*, in *dirittifondamentali.it*, 2, 2022, 422 ss.; L. Califano, *La libertà di manifestazione del pensiero ... in rete: nuove frontiere di esercizio di un diritto antico. "Fake news", "hate speech" e profili di responsabilità dei "social network"*, in *federalismi.it*, 26, 2021, 1 ss.; C. Valditara, in *Il Diritto dell'informazione e dell'informatica*, 2, 2021, 257 ss.; A. Sciortino, *"Fake news" e "post"-verità nella società dell'algoritmo*, in *dirittifondamentali.it*, 2, 2021, 422 ss.; M. Cavino, *Il triceratopo di Spielberg. "Fake news", diritto e politica*, in *federalismi.it*, 11, 2020, 32 ss.; F. Fracchia, *Le "fake news" come luogo di osservazione dei fenomeni della tecnologia e delle reti nella prospettiva del diritto*, ivi, 43 ss.; P. Passaglia, *"Fake news" e "fake democracy": una convergenza da scongiurare*, ivi, 126 ss.; M. Monti, *La disinformazione "online", la crisi del rapporto pubblico-esperti e il rischio della privatizzazione della censura nelle azioni dell'Unione Europea ("Code of practice on*

poi ricordato come da un lato il DSA vieti lo *shadowban* e, dall'altro, il codice di condotta sulla disinformazione e le linee guida della Commissione del marzo 2024 contengano raccomandazioni chiare per impedire la diffusione di tale tipologia di contenuti. Secondo il parlamentare interrogante, sarebbe possibile dedurre che Meta, al fine di evitare le responsabilità derivanti dal DSA, nasconda espressamente contenuti politici perché giudicati potenzialmente divisivi a prescindere, mentre persista nella promozione di contenuti che generano *engagement*, anche qualora veicolino messaggi pericolosi. Tale misura sarebbe risultata ancor più grave in ragione dell'approssimarsi delle elezioni europee, in quanto sarebbe stata colpita la possibilità per i candidati di far legittimamente conoscere il proprio messaggio. A tal fine, l'interrogazione richiede alla Commissione di rendere nota la risposta del gruppo Meta e quali azione intendesse intraprendere rispetto al fenomeno in questione¹⁸. Da parte sua, la Commissione si è soltanto limitata a confermare di aver inviato il 1° marzo 2024 a Meta una richiesta di informazioni riguardante, tra i vari aspetti, lo *shadowban* su Facebook e Instagram, cui ha fatto seguito l'avvio di un procedimento formale nei confronti della società, anche in relazione all'approccio verso i contenuti politici¹⁹.

All'inizio della nuova legislatura è stata presentata una seconda interrogazione parlamentare alla Commissione sul medesimo argomento, chiedendo di analizzare l'impatto che lo *shadowban* ha prodotto sulle elezioni europee del 2024, se l'applicazione di tale pratica abbia colpito determinati partiti o candidati, nonché quali misure siano state assunte dalla Commissione europea al fine di assicurare la trasparenza e la correttezza della competizione elettorale²⁰. Anche in tale occasione, la Commissione si è limitata a ribadire di aver avviato un procedimento di indagine su Meta avente ad oggetto, tra i vari aspetti, anche lo *shadowban*²¹.

Nel gennaio 2025 il gruppo Meta sembra essere parzialmente tornato sui propri passi, ammettendo che la riduzione dei contenuti politici attuata a partire dal 2021 è avvenuta «in modo piuttosto brusco», e aprendo alla possibilità di una loro graduale reintroduzione con un approccio personalizzato sulla base delle preferenze degli utenti, in modo che chi desidera vederne in misura maggiore nel proprio *feed* potrà farlo²², pur non essendo specificato in che modo ciò dovrebbe avvenire.

disinformation"), ivi, 282 ss.; M. Malvicini, *Parlamento e "fake-news": spunti per un dibattito*, ivi, 339 ss.; A. Candido, *Libertà di informazione e democrazia ai tempi delle "fake news"*, in *federalismi. it*, 20, 2020, 105 ss.; S. Poli, *Le misure Ue di contrasto alle attività di disinformazione russe, alla prova della Carta europea dei diritti fondamentali*, in *Quad. cost.*, 3, 2022, 626 ss.

¹⁸ Interrogazione parlamentare - E-000987/2024 del 2 aprile 2024, a firma dell'on. Benifei.

¹⁹ Risposta di Thierry Breton a nome della Commissione europea - Interrogazione parlamentare - E-000987/2024(ASW) del 23 maggio 2024.

²⁰ Interrogazione parlamentare - E-001816/2024 del 25 settembre 2024, a firma dell'on. Tom Vandendriessche.

²¹ Risposta della Commissione europea all'interrogazione E-001816/2024(ASW), datata 15 novembre 2024.

²² Meta, *More Speech and Fewer Mistakes*, 7 gennaio 2025.

2. Lo *shadowban* nei *social network*: un problema anzitutto definitorio

La moderazione dei contenuti viene tradizionalmente definita come il processo in forza del quale le piattaforme stabiliscono quali contenuti sono consentiti, decidendo di conseguenza sulle azioni da intraprendere per ammetterli, rimuoverli o limitarne la diffusione²³. La moderazione si articola poi in diverse forme, tra cui è possibile individuare il c.d. *shadowban*, termine che indica l'adozione non trasparente da parte degli algoritmi di sanzioni temporanee a carico degli utenti per isolarne i contenuti dalle piattaforme senza che questi ne abbiano consapevolezza²⁴.

Il concetto di *shadowban* ha subito un'evoluzione nel corso del tempo²⁵: in origine sembrava riferirsi ai partecipanti ai *forum online* i cui *post* venivano oscurati agli altri utenti, senza che gli autori ne avessero consapevolezza²⁶, accezione che pare aver trovato riscontro anche nel DSA. Nel senso più recente del suo utilizzo, lo *shadowban* pare riferirsi ad una serie di tecniche di moderazione dei contenuti che interferiscono con la visibilità di questi ultimi, come il *delisting* e il *downranking*²⁷. Tali rimedi non eliminano del tutto l'accesso ad un contenuto, ma ne riducono la visibilità nei risultati di ricerca e nei suggerimenti.

Più dettagliatamente, sono quattro le tipologie di *shadowban* che è possibile individuare: la prima e più semplice è quella di escludere un utente o i suoi contenuti dal motore di ricerca della piattaforma. La seconda è la rimozione dalla barra dei suggerimenti – funzione che completa le parole all'inserimento delle prime lettere in base alla popolarità dei termini – il cui rilievo è più difficile da accertare, in quanto potrebbe dipendere dalla scarsa popolarità del termine inserito, e non da una decisione deliberata della piattaforma. La terza è il c.d. *ghost ban*, in forza del quale la risposta o il commento di un determinato utente (per esempio nella piattaforma Twitter) non risulta visibile agli altri utenti. L'ultima forma, che si pone come una variante della precedente, consiste nel collocare le risposte di un utente ad un certo contenuto talmente in basso nei *feed* da farle risultare invisibili²⁸.

Nonostante l'ampiezza e la vaghezza dello *shadowban*, è comunque possibile identificare un nucleo comune a tutte le definizioni: tale pratica indica quel fenomeno per cui il contenuto dell'utente di una piattaforma *social*

²³ M. Paolanti, *LA e moderazione dei contenuti sui social media: il principio del "Human in the loop" nel campo del diritto all'informazione e alla comunicazione*, in questa Rivista, 1, 2024, 85 ss.

²⁴ K. Jaidka-S. Mukerje-Y. Lelkes, *Silenced on social media: the gatekeeping functions of shadowbans in the American Twitterverse*, in *Journal of Communication*, (73)2, 2023, 163–178; S. Myers West, *Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms*, in *New Media and Society*, (20)11, 2018, 4366 ss.

²⁵ P. Leerseen, *An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation*, in *Computer Law & Security Review*, 48, 2023, 2 ss.

²⁶ C. Radsch, *Shadowban/shadow banning*, in *Glossary of Platform Law and Policy Terms*, L. Belli-N. Zingales-Y. Curzi (eds.), *Internet Governance Forum*, 2021.

²⁷ K. Cotter, *Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram*, in *New Media & Society*, (21)4, 2018.

²⁸ K. Jaidka-S. Mukerje-Y. Lelkes, *Silenced on social media*, cit., 165.

network viene nascosto senza che l'autore ne abbia alcuna consapevolezza. Questa almeno era la definizione classica della pratica in esame, in uso a partire dai primi anni 2000. Da allora i *social media* sono cresciuti sia in termini di dimensioni che nell'uso di algoritmi nella gestione della propria attività, circostanza che rende la precedente definizione non in grado di cogliere appieno la portata del fenomeno (si pensi, ad esempio, al caso della rimozione di un utente o di un contenuto dalla barra delle ricerche, oppure al mancato invio delle notifiche o, ancora, ai *post* che non vengono inseriti nei suggerimenti).

Una definizione estremamente ampia del fenomeno in questione – con tutti i limiti di utilità che può incontrare tale ipotesi – è stata individuata nella legislazione dello stato americano del Wisconsin: «*“Shadow ban” means to limit or eliminate the exposure of a user, or content or material posted by a user, to other users of the social media Internet site through any means, regardless of whether the action is determined by an individual or an algorithm, and regardless of whether the action is readily apparent to a user*»²⁹. In dottrina, la definizione di shadowban più condivisa qualifica tale pratica come «*a service provider “deliberately [makes] someone’s content undiscoverable to everyone except the person who posted it, unbeknownst to the original poster”*»³⁰.

3. Shadowban: uno scenario di non sole ombre dal punto di vista delle piattaforme di *social network*

Se, da un lato, risulta evidente sin da una prima analisi il pericolo che lo *shadowban* assuma le vesti di una vera e propria forma di censura, con le conseguenze che ne derivano sotto il profilo della libertà di manifestazione del pensiero e del diritto all'informazione³¹. D'altro canto, meno indagati sono i suoi possibili effetti positivi: tale tecnica di moderazione dei contenuti è stata spesso utilizzata in risposta alla violazione delle regole delle piattaforme e al contrasto dei fenomeni d'odio che, secondo un impiego rigido dei meccanismi di rimozione, darebbe luogo alla sospensione degli

²⁹ Wisconsin Senate Bill 582, 2021. G. Nicholas, *Shedding Light on Shadowbanning*, 10.

³⁰ V. Gadde-V., K. Beykpour, *Setting the record straight on shadow banning*, 26 July 2018.

³¹ Cfr. M.E. Bucalo, *La libertà di espressione nell'era dei “social network” fra “content moderation” e necessità di una regolazione flessibile*, in *Dir. pubb. comp. eur.*, 1, 2023, 143 ss.; A. Cremone, *Libertà di espressione e nuove tecnologie: “a complicated relationship”*, 1, 2025, 441 ss.; L. Fabiano, *“Content moderation” e “free speech clause”: il controverso rapporto fra libertà e responsabilità delle piattaforme digitali nella più recente giurisprudenza della Corte suprema federale USA*, in *federalismi.it*, 27, 2024, 88 ss.; M. Tomasi, *La Corte EDU torna sui caratteri del discorso politico online: una diluizione della libera manifestazione del pensiero?*, in *DPCE online*, 1, 2024, 741 ss.; G.E. Vigevari, *Dal “caso Casapound” del 2019 alla “sentenza Casapound” del 2022: piattaforme digitali, libertà d'espressione e odio “on line” nella giurisprudenza italiana*, in questa *Rivista*, 2, 2023, 142 ss.; U. Ruffolo, *Piattaforme e content moderation nella dialettica tra libertà di espressione ed autonomia privata*, in *Eur. J. Priv. L. & Tech.*, 1, 2023, 9 ss.; R. Niro, *Piattaforme digitali e libertà di espressione fra autoregolamentazione e coregolamentazione: note ricostruttive*, in *Oss. fon.*, 3, 2021, 1369 ss.; G.F. Ferrari, *L’ “executive order” sulla prevenzione della censura online: quali effetti sull'autonomia dei “social network”?*, in *DPCE online*, 2, 2020, 1145 ss.; M. Betzu, *Libertà di espressione e poteri privati nel cyberspazio*, in *Dir. cost.*, 1, 2020, 117 ss.

account o all'espulsione di tali utenti³². Infatti, è noto come le tecniche di moderazione dei contenuti rappresentino il primo strumento dei *social media* per difendere gli utenti da quei fenomeni che spaziano dalla disinformazione al c.d. *hate speech*, minacce che, in considerazione della loro espansione, sono a loro volta in grado di ledere un ampio ventaglio di diritti fondamentali, e che hanno indotto le piattaforme *online* a rafforzare tutti quei sistemi di rimozione o riduzione della visibilità di tale categoria di contenuti³³.

Una delle (poche) motivazioni che le piattaforme hanno addotto a giustificazione dell'uso dello *shadowban* riguarda la prevenzione dei c.d. *sockpuppet accounts*: nella prassi è estremamente frequente che, qualora un profilo venga bloccato per la diffusione di contenuti in contrasto con le politiche della piattaforma, ne venga creato immediatamente dopo uno nuovo sotto altro nome. L'applicazione dello *shadowban* – che, com'è noto, avviene senza consapevolezza da parte dell'utente – permette di raggiungere l'obiettivo di ridurre la visibilità di tali contenuti senza che il proprietario ne abbia consapevolezza, evitando che la misura venga aggirata dalla creazione di nuovi *account*. Un ragionamento analogo sorregge l'applicazione di tale tecnica di moderazione dei contenuti nel caso di insulti razziali e, più in generale, di termini in contrasto con le politiche delle piattaforme: se ogni volta che ricorre una delle ipotesi appena menzionate i *social media* bannassero i contenuti in questione, gli utenti potrebbero facilmente identificare i termini "proibiti" e aggirarli utilizzandone altri; applicando lo *shadowban*, invece, tali contenuti hanno una visibilità ridotta senza alcuna comunicazione agli autori, che in questo modo non sono indotti ad aggirarne l'uso³⁴.

Alcune piattaforme di *social media* applicano poi lo *shadowban* ad alcuni specifici *hashtag*: se, in alcune ipotesi (come nel caso di contenuti che sponsorizzano i disturbi alimentari o incentivano fenomeni violenti) tale misura può risultare utile, in altri essa può comportare all'emarginazione di comunità già fortemente discriminate, come avviene nel caso di alcuni *hashtag* usati dalla comunità LGBTQIA+ in Paesi come Russia, Arabia Saudita, Bosnia e Indonesia, rappresentando un'ulteriore lesione dei diritti fondamentali dei cittadini appartenenti a tali comunità³⁵.

Un altro ambito di applicazione dello *shadowban* riguarda quelle ipotesi in cui le piattaforme potrebbero avvertire l'esigenza di stemperare una situazione che, se esacerbata, potrebbe istigare la commissione di atti violenti, pur trattandosi di contenuti che in sé non rappresentano una violazione delle *policy* dei *social network*. Tale potrebbe risultare, ad esempio, un dibattito particolarmente acceso attorno ad una notizia, oppure attorno ad un contenuto di carattere politico: come si avrà modo di approfondire, vale la

³² *Twitter Terms of service*.

³³ G. Nicholas, *Shedding Light on Shadowbanning*, cit., 6. Cfr. P. Dunn, *Moderazione automatizzata e discriminazione algoritmica: il caso dell'"hate speech"*, in *Riv. it. inf. dir.*, 1, 2022, 133 ss.; M. Nino, *The freedom of expression and hate speech in cyberspace*, in *Com. int.*, 1, 2023, 33 ss.

³⁴ Ivi, 16 ss.

³⁵ Ivi, 12 ss. F. Ryan-A. Fritz-D. Impiombato, *TikTok and WeChat: Curating and controlling global information flows*, in *Australian Strategic Policy Institute*, 2020.

pena sottolineare che tale pratica potrebbe porsi in contrasto con il diritto all'informazione dei cittadini elettori, limitando la possibilità per questi ultimi di formare una propria opinione.

Lo *shadowban* può essere poi utilizzato anche nel caso di *account* sospettati di agire come *bot*, come testimonia l'uso che ne aveva fatto la piattaforma Twitter verso quegli *account* sospetti di essere tali³⁶, caratteristica che può essere desunta dall'invio di numerosi inviti o di *post* molto frequenti³⁷.

Nonostante i casi sopra riportati, ad oggi l'unica motivazione che sembra essere stata addotta dalle piattaforme di *social media* per negare apertamente l'applicazione dello *shadowban* sembra da individuare nella volontà di evitare che alcuni governi autoritari chiedano alle piattaforme di farne applicazione nei rispettivi Paesi. Non ammettendo di ricorrere a tale tecnica di moderazione dei contenuti, le aziende potevano continuare a sostenere di non avere le necessarie competenze tecniche per soddisfare tale richiesta. Tuttavia, occorre precisare che tali informazioni risultano di difficile verificabilità, in quanto le comunicazioni tra governi e piattaforme avvengono generalmente in via riservata: i rapporti sulla trasparenza pubblicati dalle società sulle richieste dei governi di moderare contenuti specifici spesso forniscono solo informazioni generiche, per lo più in termini di numero di richieste di rimozione e di risposte delle piattaforme³⁸. Quel che è certo è che i governi di tutto il mondo sono a conoscenza del concetto di *shadowban*, anche attraverso le denunce in cui lamentano di esserne stati a loro volta oggetto³⁹. Se, da un lato, società civile e studiosi in molte parti del mondo manifestano preoccupazione sul fatto che i loro governi possano chiedere o imporre ai fornitori di servizi di applicare tale tecnica⁴⁰, dall'altro lato una maggiore trasparenza da parte dei *provider* sul rapporto con i governi consentirebbe a media tradizionali, ricercatori e cittadini di valutare tali informazioni⁴¹.

4. Lo *shadowban* e i suoi effetti sul dibattito pubblico

Assai significativi sono gli effetti che tale tecnica di moderazione dei contenuti è in grado di produrre sulle campagne elettorali *online* e, più in generale, sui diritti di partecipazione politica che ne risultano colpiti sia dal punto di vista del prodromico diritto all'informazione e alla formazione

³⁶ Y.-S. Chen-T. Zaman, *Shaping opinions in social networks with shadow banning*, in *PLoS ONE*, (19)3, 2020, 2 ss.

³⁷ M. Horten, *Algorithms patrolling content: where's the harm?*, in *International Review of Law, Computers & Technology*, (38)1, 2024, 45.

³⁸ C. Vogus-E. Llansó, *Making Transparency Meaningful: A Framework for Policymakers*, in *Center for Democracy & Technology*, 2021.

³⁹ P. Kaushika, *New allegation against Twitter – Modi supporters say PM is “shadow-banned”*, in *ThePrint*, 16 February 2019; G. Szakacs, *Hungary mulls sanctions against social media giants*.

⁴⁰ G. Lakier, *Informal Government Coercion and The Problem of “Jawboning”*, in *Lawfare*, 26 July 2021; P. Ravi, *Shadow Bans, Criminal Cases, Takedowns: Inside India's Expanding Digital Crackdown*, in *Article 14*, 26 July 2021.

⁴¹ G. Nicholas, *Shedding Light on Shadowbanning*, cit., 20.

di un'opinione libera e consapevole in capo ai cittadini, sia per quanto riguarda la possibilità per partiti, movimenti e singoli candidati di far conoscere il proprio messaggio concorrendo su un piano paritario⁴².

In apparenza, può apparire condivisibile la tesi di Zuckerberg, secondo cui tale pratica è nata dall'esigenza di stemperare la polarizzazione del dibattito pubblico sui *social network* dopo l'assalto a Capitol Hill. È stato a partire da tale momento che Meta – sulla base di analisi mai rese pubbliche – ha iniziato a sostenere che la politica su Facebook sarebbe risultata divisiva, da cui l'intento di evitare che essa “prendesse il sopravvento” (anche se non si è compreso su cosa, se non sui contenuti a carattere puramente commerciale). Dopo lo scandalo Cambridge Analytica, le teorie del complotto su QAnon e le numerose critiche sulla disinformazione circolante sui *social* ai tempi della pandemia, sembra quindi che il gruppo di Zuckerberg abbia deciso di imprimere ai *social network* una virata lontana dal mondo dell'informazione – e dell'informazione politica in particolare – per ricondurli ad una funzione di mero intrattenimento. La ritirata strategica dalle piattaforme potrebbe lasciare nuovamente spazio ai media tradizionali, che tornerebbero a produrre un'informazione accurata, verificata, e priva della ricerca di *engagement* che caratterizza i contenuti sui *social network*.

Tuttavia, non si può fare a meno di sottolineare i pericoli che tale decisione rischia di produrre sull'informazione e, di riflesso, sul dibattito pubblico in generale.

Prima di analizzare gli effetti derivanti da tale forma di moderazione dei contenuti, occorre considerare che l'informazione veicolata dai *social* è tipicamente di tipo passivo, in quanto gli utenti visualizzano ciò che viene

⁴² In tema di comunicazione politica e regolazione delle campagne elettorali *online*, cfr. E. Caterina, *La comunicazione elettorale sui “social media” tra autoregolazione e profili di diritto costituzionale*, in *Oss. fon.*, 1, 2021, 1393 ss.; G. Menegus, *La disciplina della comunicazione istituzionale durante le campagne elettorali online*, in *federalismi.it*, 19, 2025, 190 ss.; M. Monti, *Elezioni regionali, Co.Re.Com. e propaganda politico-elettorale “online”: nodi irrisolti e prospettive future fra il livello regionale, nazionale ed europeo*, in *Ist. fed.*, 1, 2023, 127 ss.; G. Trenta, *Quale aspettativa futura per una nuova disciplina in materia di comunicazione elettorale sulle piattaforme online*, in *ambientediritto.it*, 1, 2022, 515 ss.; R. Borrello, *Comunicazione politica, comunicazione istituzionale e propaganda tra area della politica-libertà e area della politica-funzione*, in *Giur. cost.*, 2, 2016, 720 ss.; G. Stegher, *Da cittadini elettori a cittadini consumatori: osservazioni sull'importanza di regolare le campagne elettorali sui “social media”*, in *Nomos*, 1, 2023, 9 ss.; M. Ladu, *La costruzione del consenso politico-elettorale e l'utilizzo dei social media nell'epoca della “disintermediazione democratica”*, in *federalismi.it*, 23, 2022, 170 ss.; G. Donato, *Il potere senza responsabilità dei “social media” nelle campagne elettorali*, in questa *Rivista*, 2, 2020, 360 ss.; C. Giunta, *L'espressione del pensiero politico: comunicazione, propaganda, pubblicità*, in *federalismi.it*, 14, 2025, 21 ss.; G. d'Ippolito, *Comunicazione politica “online”: dal messaggio politico commercializzato alle sponsorizzazioni sui “social network”*, in questa *Rivista*, 1, 2020, 159 ss.; M. Monti, *Le “Internet platforms”, il discorso pubblico e la democrazia*, in *Quad. cost.*, 4, 2019, 811 ss.; G. Trenta, *Quale aspettativa futura per una nuova disciplina in materia di comunicazione elettorale sulle piattaforme online*, in *ambientediritto.it*, 1, 2022, 515 ss.; F. Falconi, *Contrasto allo “hate speech” e responsabilità nella comunicazione politica online: note alla sentenza della Corte Europea dei diritti umani Sanchez C. Francia*, in *Com. internaz.*, 1, 2024, 115 ss.; G.E. Vigevari - M. Cazzaniga, *Comunicazione politica e Costituzione tra televisione e piattaforme digitali*, in questa *Rivista*, 1, 2024, 14 ss.; G. Muto, *The use of AI in electoral campaigns: key issues and remedies*, in questa *Rivista*, 3, 2024, 262 ss.; F. Fabrizzi, *Campagne elettorali, formazione del consenso, astensionismo: nel silenzio assordante sui referendum, ha senso oggi il silenzio elettorale?*, in *federalismi.it*, 16, 2022, 4 ss.; F. Donati, *“Internet” e campagne elettorali*, in *federalismi.it*, 16, 2019, 4 ss.

proposto loro nelle *homepage* e nei *feed*, prevalentemente sulla base delle preferenze registrate dall'algoritmo: l'esclusione dei contenuti di natura politica colpirebbe coloro che già erano interessati a tali argomenti (ma che magari non seguivano specifici pagine o profili), ma anche tutti gli altri utenti, che non avrebbero modo di informarsi e magari, come nel caso dei più giovani, iniziare a seguire il dibattito pubblico. Un esempio può essere utile per chiarire la portata del fenomeno: un noto quotidiano nazionale riporta che tra febbraio e aprile 2024, i propri *post* sono stati distribuiti per il 94% ad utenti che ancora non seguivano la testata e che hanno avuto modo di scoprirla attraverso le inserzioni oggi precluse. Di questi *non-follower*, oltre mezzo milione ha deciso di interagire con i *post*, dimostrando interesse per gli argomenti trattati⁴³.

In primo luogo, va sottolineato che una definizione vaga e imprecisa di contenuto politico lascia un margine di discrezionalità enorme alle piattaforme, da cui potrebbe derivare un indebito vantaggio (e una patente violazione del principio di eguaglianza e della parità di trattamento) per alcuni candidati, partiti, movimenti o gruppi i cui contenuti – a causa della vaghezza della definizione sopra evidenziata – potrebbero continuare a comparire nei *feed* e nelle *homepage* degli utenti, anche se lo *shadowban* sembra applicarsi soprattutto a profili non verificati più che a quelli ufficiali di esponenti politici. Rimane peraltro da appurare il rapporto con le inserzioni a contenuto politico sponsorizzate, alle quali al momento non si comprende se si applichi lo *shadowban* generale: come si avrà modo di approfondire nell'analisi della pronuncia belga, sembra che anche tale categoria di contenuti possa essere in realtà soggetta a *shadowban*, ponendo quindi il tema della risarcibilità dell'eventuale danno arrecato al candidato cui tale misura sia stata applicata.

In secondo luogo, va tenuta presente l'influenza che i *social* hanno avuto nel plasmare l'informazione e il modo in cui le notizie vengono riportate: dopo gli sforzi e gli investimenti economici messi in campo dai media tradizionali per adattarsi ai *social network* (con risultati spesso criticabili sotto il profilo della qualità e dell'accuratezza dell'informazione), ora si troverebbero nella difficile situazione di tornare allo stato precedente, perdendo nel frattempo una considerevole fonte di entrate economiche, con il rischio di accelerare ancora di più la desertificazione del panorama dei media.

Uno degli effetti più immediati che lo *shadowban* potrebbe produrre è che molti utenti si sentirebbero oggetto di una forma di censura che ne limita la libertà di manifestazione del pensiero. Proprio con riferimento ai contenuti oggetto di *shadowban*, è sottile il confine tra moderazione e censura, soprattutto qualora si aderisse all'opinione che li inquadra nell'ambito del *borderline speech*⁴⁴, ossia «not covered by the barriers to freedom of expression but is considered inappropriate in the public debate»⁴⁵.

In effetti, la de-prioritizzazione dei contenuti di un utente da parte delle

⁴³ L. Ferrante, *Bavaglio social, Instagram ha deciso di oscurare i post che trattano di politica e temi sociali. E l'Ue indaga sulle violazioni*, 2 maggio 2024.

⁴⁴ K. Jaidka, *Shadowbans on political expression*, cit., 3.

⁴⁵ A. Heldt, *Borderline speech: Caught in a free speech limbo?*, in *Internet Policy Review*, 15 October 2020.

piattaforme può assumere i tratti di una forma di censura, da cui deriva la lesione della libertà di manifestazione del pensiero, nonché del diritto all'informazione⁴⁶. Alcune forme di de-prioritizzazione equivalgono di fatto ad una rimozione del contenuto, ad esempio nell'ipotesi in cui finisca talmente nascosto nei *feed* degli utenti da risultare invisibile. Allo stesso tempo – e qui sta la difficoltà nell'orientarsi in una pratica ambigua come quella in esame – il punto rilevante è che non tutte le forme di de-prioritizzazione equivalgono ad uno *shadowban*: come si è già avuto modo di evidenziare, il funzionamento delle piattaforme si basa sulla raccomandazione di contenuti, e quelli meno recenti o che riscuotono meno *engagement* verranno inevitabilmente collocati in posizioni secondarie all'interno dei *feed* o saranno oggetto di un minor numero di suggerimenti. Tuttavia, sia per l'utente che abbia postato un contenuto, sia per gli altri utenti, è impossibile stabilire se quel risultato derivi dall'applicazione dello *shadowban* o dal normale funzionamento della piattaforma. Tale convinzione, unita alla scarsa trasparenza sul funzionamento delle piattaforme *online* che non consente di distinguere quando ci si trovi effettivamente dinanzi ad un'ipotesi di *shadowban*, genera negli utenti un senso di isolamento e frustrazione⁴⁷, rendendo al contempo più difficoltoso il contatto con comunità *online* di cui i singoli potrebbero voler condividere i contenuti. Ciò contribuisce ad un'ulteriore radicalizzazione delle opinioni individuali e alla polarizzazione del dibattito pubblico⁴⁸, sfociando in teorie del complotto che sembrano trovare terreno fertile nella galassia del mondo conservatore⁴⁹. D'altro canto, l'accusa secondo cui lo *shadowban* colpirebbe in misura più significativa contenuti di stampo conservatore è rimasta indimostrata⁵⁰, mentre sembrano dotate di qualche fondamento le critiche secondo cui i *social network* colpirebbero in particolare persone di colore, donne o soggetti appartenenti alla comunità *queer*⁵¹. Occorre poi notare come alcuni studi compiuti sia in Francia⁵² che negli USA⁵³ evidenziano come gli *account* verificati e quelli appartenenti ad esponenti politici risultino meno colpiti da tale fenomeno rispetto alla

⁴⁶ K. Cotter, "Shadowbanning is not a thing": Black box gaslighting and the power to independently know and credibly critique algorithms, in *Information, Communication & Society*, 2021, 1 ss.; O. Papakyriakopoulos-J.C. Medina Serrano-S. Hegelich, *Political communication on social media: A tale of hyperactive users and bias in recommender systems*, in *Online Social Networks and Media*, 15, 2020, 1 ss..

⁴⁷ G. Nicholas, *Shedding Light on Shadowbanning*, cit., 33.

⁴⁸ Ivi, 10.

⁴⁹ Ivi, 35 ss.; L. Stack, *What Is a 'Shadow Ban,' and Is Twitter Doing It to Republican Accounts?*, in *The New York Times*, 26 July 2018; V. Gadde-K. Beykpour, *Setting the record straight on shadow banning*, cit.; K. Jaidka, *Shadowbans on political expression*, cit., 5.

⁵⁰ Ibidem.

⁵¹ J. Cook, *Instagram's shadow ban on vaguely 'inappropriate' content is plainly sexist*, in *Huffington Post*, 2 December 2022; C. Joseph, *Instagram's murky 'shadow bans' just serve to censor marginalised communities*, in *The Guardian*, 8 November 2019.

⁵² E. Le Merrer-B. Morgan-G. Trédan, *Setting the Record Straight on Shadow Banning*, in *INFOCOM 2021- IEEE International Conference on Computer Communications*, May 2021, 1 ss.

⁵³ K. Jaidka-S. Mukerje-Y. Lelkes, *Silenced on social media*, cit.

media degli utenti comuni, confermando la tendenza secondo cui le élite politiche godrebbero di un trattamento differenziato nell'ambito dei *social network*⁵⁴.

Guardando più da vicino agli effetti che l'applicazione dello *shadowban* è in grado di produrre sul dibattito pubblico, un primo aspetto che merita di essere evidenziato è che la minore visibilità dei contenuti politici influenza la possibilità per gli utenti di accedere alle informazioni e alla capacità di elaborare una propria opinione, impoverendo il panorama dell'informazione nel suo complesso⁵⁵: l'aspetto che si vuole sottolineare è che fino a oggi il dibattito dottrinale si è concentrato prevalentemente sull'influenza che i meccanismi di moderazione esercitano sulla formazione dell'opinione degli utenti attraverso la targetizzazione di contenuti, nel caso dello *shadowban* il problema riguarda a monte l'eliminazione di questi ultimi e l'impoverimento del dibattito pubblico e politico⁵⁶.

Secondo le ricerche più recenti, lo *shadowban* è in grado di influenzare l'opinione degli utenti nei *social network*: se, da un lato, esso è risultato utile nel contrasto alla polarizzazione e alla disinformazione – stemperando la visibilità di determinati contenuti che sarebbero divenuti altrimenti “virali” – dall'altro lato può sortire l'effetto opposto, polarizzando il dibattito *online* nascondendo alcuni contenuti e indirizzando l'attenzione degli utenti su altri, con l'ulteriore aggravante che tutto ciò avviene dietro un'apparente neutralità (anche politica) e senza alcuna consapevolezza da parte dell'utente⁵⁷.

In termini generali, è stato dimostrato come i profili più recenti e quelli non verificati siano più facilmente soggetti a *shadowban*⁵⁸. Ad oggi non sembra tuttavia che attraverso tale pratica di moderazione dei contenuti si sia prodotto l'effetto di favorire una determinata ideologia, poiché

⁵⁴ J. Horwitz, *Facebook says its rules apply to all, company documents reveal a secret elite that's exempt*, 13 September 2021.

⁵⁵ J. Kulshrestha-M. Eslami-J. Messias-M.B. Zafar-S. Ghosh-K.P. Gummadi-K. Karahalios, *Quantifying search bias: Investigating sources of bias for political searches in social media*, in *CSCW '17: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 417 ss.

⁵⁶ Cfr. L. Cianci, *Il diritto ad essere informati alla prova delle strategie di “microtargeting” per la comunicazione politica*, in *Nomos*, 1, 2023, 8 ss.; C. Coniglione, *L'utilizzo dei “big data” in ambito politico-elettorale e il loro impatto sulla democrazia rappresentativa*, ivi, 15 ss.

⁵⁷ Y.-S. Chen-T. Zaman, *Shaping opinions in social networks with shadow banning*, cit., 22 ss.; A. Sirbu-D. Pedreschi-F. Giannotti-J. Kertész, *Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model*, in *PloS one*, (14)3, 2019; A.F. Peralta-M. Neri-J. Kertész-G. Iniguez, *Effect of algorithmic bias and network structure on coexistence, consensus, and polarization of opinions*, in *Physical Review E*, (104)4, 2021; N. Perra-L.E. Rocha, *Modelling opinion dynamics in the age of algorithmic personalisation*, in *Nature – Scientific reports*, (9)1, 2019, 1 ss.; C. Blex-T. Yasseri, *Positive algorithmic bias cannot stop fragmentation in homophilic networks*, in *The Journal of Mathematical Sociology*, (46)1, 2022, 80 ss.; G. Iannelli-G. De Marzo-C. Castellano, *Filter bubble effect in the multistate voter model*, in *Chaos: An Interdisciplinary Journal of Nonlinear Science*, (32)4, 2022; F. Cinus-M. Minici-C. Monti-F. Bonchi, *The effect of people recommenders on echo chambers and polarization*, in *Proceedings of the International AAAI Conference on Web and Social Media*, 16, 2022, 90 ss.; A.L.M. Vilela-L.F.C. Pereira-L. Dias-H.E. Stanley-L.R. Da Silva, *Majority-vote model with limited visibility: An investigation into filter bubble*, in *Physica A: Statistical Mechanics and its Applications*, 563, 2021.

⁵⁸ K. Jaidka-S. Mukerje-Y. Lelkes, *Silenced on social media*, cit., 174.

tutti i contenuti politici possono esserne soggetti in maniera indistinta, suggerendo la conclusione per cui Twitter non sarebbe particolarmente ostile a contenuti di matrice conservatrice e riconducibili all'area politica della destra⁵⁹.

Secondo studi recenti, più chiari – e maggiormente interessanti ai fini della presente analisi – sembrano gli effetti che lo *shadowban* è in grado di esercitare sulla nascita di *echo chambers*⁶⁰, con riflessi sia sulla partecipazione politica offline⁶¹ che sullo scarso interesse degli elettori per il momento elettorale⁶².

Come risulta evidente già ad una primissima lettura, lo *shadowban* si pone al centro del bilanciamento tra esigenze di moderazione dei contenuti, libertà di manifestazione del pensiero e diritto di informazione degli utenti nelle piattaforme *online*, costituendo quindi una prospettiva dalla quale è possibile scorgere le sfide che interessano oggi i *social network*⁶³. Il tema della capacità manipolativa che tali piattaforme sono in grado di esercitare sulle opinioni degli utenti è al centro della riflessione dottrinale anche per il costituzionalismo: come si è avuto modo di verificare, le piattaforme possono delineare la moderazione dei contenuti in modo da manipolare deliberatamente le opinioni degli utenti attraverso una distribuzione arbitraria della visibilità dei contenuti, anche se (per il momento) non è dimostrato che tale tecnica abbia avvantaggiato una determinata ideologia. Se è vero che lo *shadowban* può risultare utile nella limitazione della diffusione di contenuti pericolosi e nel contrasto alla disinformazione *online*, non si possono sottacere i rischi derivanti da un utilizzo di tale tecnica nel senso di indirizzare (o sviare) l'attenzione degli utenti a favore o contro un determinato argomento, soprattutto quando questo avvenga nell'ambito di una campagna elettorale e, per di più sotto le mentite spoglie di una moderazione dei contenuti neutrale e senza che gli utenti ne siano informati.

In via del tutto interlocutoria, è possibile delineare alcune (provvisorie) conclusioni: in termini generali, sembra trovare condivisione in dottrina la considerazione per cui la strada da seguire per le piattaforme debba muovere nel senso di una maggior trasparenza delle *policies* sulla moderazione dei

⁵⁹ Ivi, 172.

⁶⁰ S. Bhadani-S. Yamaya-A. Flammini-F. Menczer-G.L. Ciampaglia-B. Nyhan, *Political audience diversity and news reliability in algorithmic ranking*, in *Nature Human Behaviour*, 6, 2022; S. Majo-Vazquez-M. Congosto-T. Nicholls-R.K. Nielsen, *The role of suspended accounts in political discussion on social media: Analysis of the 2017 French, UK and German elections*, in *Social Media Society*, (7)3, 2021; S. Zannettou, "I won the election!": *An empirical analysis of soft moderation interventions on twitter*, in *Proceedings of the International AAAI Conference on Web and Social Media*, (15)2021, 865 ss.

⁶¹ J. Matthes-J. Knoll-S. Valenzuela-D.N. Hopmann-C. Von Sikorski, *A meta-analysis of the effects of cross-cutting exposure on political participation*, in *Political Communication*, (36)4, 2019, 523 ss.

⁶² D. Cosley-S.K. Lam-I. Albert-J.A. Konstan-J. Riedl, *Is seeing believing? how recommender system interfaces affect users' opinions*, in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2003, 585 ss.

⁶³ Y.-S. Chen-T. Zaman, *Shaping opinions in social networks with shadow banning*, 2 ss.

contenuti – anche se ciò potrebbe facilitarne l’aggiramento⁶⁴ - e di una maggiore collaborazione tra *big-tech* e mondo accademico.

In ogni caso, sembra di poter ragionevolmente affermare che sia necessaria una maggiore documentazione sull’uso dei dati e sui diritti degli utenti, in modo che questi ultimi possano comprendere i loro diritti e impugnare le decisioni degli algoritmi quando necessario⁶⁵, opportunità che andrebbero soprattutto a vantaggio delle voci altrimenti emarginate o prive di potere. Molti sono i commentatori che hanno invocato l’adozione di modelli alternativi, che spaziano dall’introduzione di un modello di moderazione dei contenuti incentrato più sugli aspetti educativi nei confronti degli utenti piuttosto che su una funzione meramente repressiva verso le piattaforme⁶⁶, per arrivare a sistemi caratterizzati da una trasparenza più significativa ai quali applicare il concetto di giusto processo, in modo che gli utenti ricevano spiegazioni chiare sulle decisioni che li riguardano, comprese le regole che hanno violato, il contenuto specifico in contrasto con esse e il procedimento con cui è stato identificato⁶⁷.

5. *Shadowban* e lesione dei diritti fondamentali: spunti a partire da una recente sentenza belga

Dopo aver approfondito l’impatto che lo *shadowban* può determinare sul dibattito pubblico, sia nella prospettiva degli effetti sul singolo utente che sulla capacità di orientare la formazione dell’opinione pubblica, si intende ora spostare il piano dell’analisi sulla lesione dei diritti fondamentali derivante da tale pratica, analizzando dapprima i profili procedurali delle decisioni e poi le conseguenze sul piano sostanziale.

Muovendo dal primo aspetto, è da notare che, rispetto alle sanzioni immotivate, quelle comminate segretamente dalle piattaforme sono più difficili da rilevare e da contestare, compromettendo il diritto di impugnarle nell’ambito di un giusto processo. Inoltre, come rilevato da recenti studi, l’applicazione dello *shadowban* è apparsa arbitraria, imposta unilateralmente dalle piattaforme senza alcun preavviso e in assenza della violazione delle *policy* interne⁶⁸. Nelle comunicazioni via e-mail intercorse tra alcuni utenti e le piattaforme è stato confermato che l’applicazione delle restrizioni non era dipesa dalla illiceità dei contenuti ma dalla quantità e della frequenza con cui un utente pubblicava⁶⁹: il fatto che tale circostanza costituisca un

⁶⁴ K. Jaidka-S. Mukerje-Y. Lelkes, *Silenced on social media*, cit., 174.

⁶⁵ A. Heldt, *Borderline speech: Caught in a free speech limbo?*, in *Internet Policy Review*, 15 October 2020.

⁶⁶ S. Myers West, *Censored, suspended, shadowbanned: User interpretations of content moderation on social media platforms*, cit.

⁶⁷ N. Suzor-S. Myers West, *What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation*, in *International Journal of Communication*, (13)2019, 1526 ss., United Nations, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, A/73/348, 2018.

⁶⁸ M. Horten, *Algorithms patrolling content: where’s the harm?*, in *International Review of Law, Computers & Technology*, (38)1, 2024, 56.

⁶⁹ *Ibid.*

parametro ai fini dell'applicazione dello *shadowban* rappresenta un problema rilevante per una campagna elettorale condotta nelle piattaforme *online*, poiché la capacità diffusiva del contenuto è coesistente alla natura e allo scopo della propaganda politica.

Un ulteriore profilo da considerare è la necessaria prevedibilità delle conseguenze delle proprie azioni da parte degli utenti. A tal fine, le regole applicabili dalle piattaforme devono essere chiare, precise e disponibili al pubblico, e l'azione intrapresa deve essere proporzionata rispetto all'obiettivo da raggiungere⁷⁰. L'applicabilità degli standard di rispetto dei diritti umani è stata confermata nel 2021 dal Facebook Oversight Board (2021), organo istituito dalla piattaforma per giudicare le azioni di moderazione dei contenuti: nella decisione sul caso numero 2020-005-FB-UA, il Board ha basato la sua valutazione sull'art. 19 del Patto internazionale sui diritti civili e politici (ICCPR), richiamando i principi di legalità, necessità e proporzionalità, e ricordando che le azioni di moderazione non devono essere percepite dall'utente come arbitrarie. Inoltre, ha invitato Facebook a fornire una notifica che citi la regola specifica (Community Standard) utilizzata come base per la rimozione⁷¹. Gli standard comunitari di Facebook sono scritti in modo da lasciare alla piattaforma un ampio margine di manovra, ma non sono facilmente comprensibili dagli utenti. La mancanza di informazioni pubbliche ha fatto sì che costoro non potessero prevedere come le varie restrizioni sul comportamento li avrebbero riguardati, soprattutto quando esse si basavano su elementi quantitativi, come il numero dei *post* o dei commenti.

La posizione dell'amministratore della pagina in questi casi è stata aggravata dalla mancanza di possibilità di appello o di ottenere un risarcimento. Facebook afferma che agli utenti viene offerta l'opportunità di appellarsi alla rimozione dei contenuti ma, a differenza della rimozione, non è chiaro come questo si applichi quando il motivo dell'applicazione dello *shadowban* è un comportamento, e non uno specifico contenuto. Alcuni amministratori di pagine colpite da tale misura hanno provato a ricorrere al processo di appello di Facebook, ma non sempre è stata offerta loro tale opportunità, da cui deriva un limitato numero di ricorsi secondo i dati sulla trasparenza riportati dalla piattaforma⁷².

Spostando l'analisi sul piano sostanziale, uno degli effetti più immediati ed evidenti derivanti dallo *shadowban* è la limitazione della libertà di manifestazione del pensiero a causa della retrocessione del contenuto colpito da tale pratica di moderazione, al punto da ipotizzare la configurazione di una forma di censura vera e propria. Merita poi di essere evidenziato che lo *shadowban* colpisce anche tutti coloro che seguivano le pagine o i profili: nel diritto alla libertà di espressione e di informazione è infatti rinvenibile anche un versante passivo, come diritto a ricevere informazioni o contenuti da una determinata fonte, la cui portata viene lesa dall'applicazione della pratica di moderazione in assenza di qualsiasi

⁷⁰ G. Smith, *Online Harms and the Legality Principle*, 20 June 2020.

⁷¹ M. Horten, *Algorithms patrolling content: where's the harm?*, cit., 56.

⁷² *Ibid.*

forma di trasparenza⁷³. Risultato ancor più paradossale se si considera che si trattava di contenuti leciti, oscurati dall'applicazione di algoritmi che attribuiscono un peso determinante al volume dei *likes* e alla frequenza dei *post*, in luogo della liceità del contenuto in questione⁷⁴.

Tale pratica di moderazione dei contenuti lede quindi il diritto all'informazione sotto ogni profilo, impoverendo il dibattito pubblico nei *social network* sui contenuti politici, attribuendo visibilità ad altro tipo di contenuti – o, come appare più preoccupante, a determinati contenuti politici, sulla base di criteri non individuati né individuabili – compromettendo la possibilità per gli utenti di formare una propria opinione politica in modo consapevole, con effetti dimostrati sulla nascita di *echo chambers* e sulla polarizzazione del dibattito pubblico, ponendosi dunque in contrasto con la finalità di stemperare tali fenomeni talora addotta come giustificazione per l'adozione di tale pratica di moderazione dei contenuti.

Occorre poi sottolineare come le élite – anche politiche – risultino più difficilmente soggette a *shadowban*, vuoi perché meno oggetto di segnalazioni da parte degli utenti, vuoi perché più propense ad utilizzare un linguaggio meno censurabile, confermando quella tendenza per cui i *social media* sono divenuti qualcosa di molto distante dall'idea originaria di spazio di democrazia deliberativa e partecipativa, per assumere i tratti di un sistema molto più simile ai media tradizionali dal punto di vista del ruolo che l'intermediazione riveste nella moderazione dei contenuti. Ciò pone, tra gli altri, anche un problema di parità di trattamento tra gli utenti, nel momento in cui, pur condividendo i medesimi contenuti politici, sono soggetti o meno all'applicazione di una misura restrittiva della visibilità dei medesimi⁷⁵.

Alcune delle lesioni dei diritti fondamentali appena illustrate sono state riconosciute da una sentenza di un tribunale belga, che ha condannato al risarcimento dei danni la piattaforma Facebook per l'applicazione dello *shadowban* alla pagina di un politico nazionale⁷⁶: è opportuno precisare sin dalla premessa che la condanna si è basata sul regolamento generale sulla protezione dei dati personali, e non sul DSA al tempo dei fatti non ancora vigente.

Muovendo dai fatti di causa, META aveva ritenuto che i *post* del politico in questione violassero i Termini di servizio e gli Standard della community di Facebook, citando come motivi incitamento all'odio, sostegno a organizzazioni con tale scopo, bullismo e intimidazione. Per questo motivo, all'inizio del 2021, sono state imposte restrizioni temporanee alle pagine Facebook del politico. Tra le varie misure, è stato applicato anche uno *shadowban* che ha ridotto la visibilità della pagina. Nonostante META abbia affermato che le misure sono state revocate alla fine del

⁷³ *Ibid.* M. Husovec, *(Ir)Responsible Legislature? Speech Risks under the EU's Rules on Delegated Digital Enforcement*, in *SSRN Papers*, 18 February 2021.

⁷⁴ M. Horten, *Algorithms patrolling content: where's the harm?*, cit., 56.

⁷⁵ J. Horwitz, *Facebook says its rules apply to all, company documents reveal a secret elite that's exempt*, in *The Wall Street Journal*, 13 September 2021.

⁷⁶ L. Stolle, *Belgian case regarding payment of damages for shadowbanning a politician by META*, in *IRIS Legal Observations of the European Audiovisual Observatory*.

2021, il ricorrente ha contestato il fatto, fornendo dati che indicavano che la misura è persistita oltre il 2021 e ha continuato a influire sulla visibilità media della pagina. Il ricorrente ha adito la giustizia belga, sostenendo che META avrebbe violato il contratto non rispettando il suo diritto alla libertà di espressione, discriminandolo illegalmente in base alle sue convinzioni politiche, non adempiendo al contratto secondo buona fede e applicando clausole vessatorie nei suoi confronti. Inoltre, il ricorrente ha sostenuto come META lo abbia sottoposto a un *profiling* illegale ai sensi del GDPR sulla base delle sue convinzioni politiche. Il ricorso è stato rigettato in primo grado, con l'argomentazione per cui i tribunali belgi non avevano giurisdizione in materia, decisione ribaltata poi dalla Corte d'appello che ha proceduto all'esame del caso.

La Corte afferma anzitutto che il ricorrente non poteva invocare il diritto alla libertà di espressione contro META, come sancito, tra l'altro, dall'art. 10 della CEDU in quanto tale disposizione non avrebbe efficacia orizzontale. La Corte non ha poi ritenuto necessario pronunciarsi sulla pretesa violazione delle leggi belghe in materia di discriminazione.

Per quanto riguarda la questione della profilazione illegittima ai sensi del GDPR, la Corte osserva che, ai sensi dell'art. 22, par. 1, del regolamento, un provvedimento basato esclusivamente su un processo decisionale automatizzato non è ammessa quando comporti decisioni che possono avere effetti giuridici per gli interessati o incidere significativamente sulla loro persona. A tal fine, Meta non ha provato che vi sia stato un intervento umano sull'imposizione delle restrizioni alle pagine in questione. Per la sua analisi, la Corte distingue tra la cancellazione dei *post* o l'inserimento di avvisi sulla pagina Facebook e lo *shadowban* vero e proprio: solo quest'ultimo comporta effetti giuridici o di analoga rilevanza per il ricorrente, in ragione del suo impegno politico e dell'impatto che la riduzione della visibilità produce.

Sebbene l'art. 22 del GDPR preveda alcune eccezioni alla regola di cui al par. 1 – tra cui l'esecuzione di un contratto (art. 22, par. 2, lett. a) – la Corte osserva, ai sensi del successivo par. 3, che la piattaforma non ha predisposto le misure di protezione adeguate alla tutela dei diritti dell'interessato, non essendo stato previsto il diritto di ottenere l'intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione.

META avrebbe poi dovuto dimostrare di aver fornito informazioni utili sulla logica sottostante al processo decisionale automatizzato, nonché sull'importanza e le conseguenze previste di tale trattamento. Inoltre, l'informativa sulla privacy non rende adeguatamente edotti gli utenti sull'esistenza di qualsiasi forma di processo decisionale automatizzato secondo l'art. 13, par. 2, lett. f) del regolamento. La semplice menzione del "trattamento automatizzato" nell'informativa sulla privacy non è sufficiente se è in atto un processo decisionale di quel tipo. Di conseguenza, la Corte belga afferma che l'uso del processo decisionale automatizzato sia per lo *shadowban* che per le altre sanzioni imposte al profilo è illegittimo per i motivi precedentemente discussi.

Il ricorrente ha sostenuto che, a causa delle sanzioni imposte da META, sono stati sopportati costi pubblicitari aggiuntivi: a tal fine, la Corte

ha ritenuto che la necessità di pagare per gli annunci rivolti ai follower rappresentasse una conseguenza dello *shadowban*, osservando che il ricorrente avrebbe potuto spendere tali risorse per altre spese professionali, tuttavia, la circostanza che le restrizioni fossero continuate oltre il 2021 non è stata considerata provata. Ha quindi calcolato ai fini del risarcimento solo i costi di tali annunci rivolti ai follower per il periodo compreso tra l'inizio e la fine del 2021, condannando Meta al pagamento di oltre 27.000 euro di risarcimento del danno in favore del ricorrente.

La pronuncia in esame risulta di particolare interesse anche rispetto all'ordinamento italiano, in quanto essa si basa integralmente su disposizioni di origine comunitaria (come il GDPR) applicabili in tutto il territorio dell'Unione, anche se, come si avrà modo di approfondire nel paragrafo che segue, tale pratica potrebbe risultare vietata grazie al DSA.

6. Il divieto di *shadowban* nel DSA

La legislazione europea più recente ha riconosciuto per la prima volta l'esistenza dello *shadowban* e il suo utilizzo come pratica restrittiva: nel DSA tale pratica è stata inserita tra gli strumenti di moderazione dei contenuti – art. 3, lett. t) e art. 17, par. 1, lettera a) – al fine di stabilire le regole che le piattaforme *online* devono seguire, con particolare riferimento alle garanzie procedurali da applicare nei casi in cui i *post* degli utenti siano stati erroneamente limitati.

Il DSA dedica particolare attenzione al tema della moderazione dei contenuti, a partire dai Considerando che precedono il testo del regolamento⁷⁷. In particolare, i prestatori di servizi intermediari «dovrebbero includere informazioni su politiche, procedure, misure e strumenti utilizzati ai fini della moderazione dei contenuti, tra cui il processo decisionale algoritmico e la revisione umana, nonché le regole procedurali del loro sistema interno di gestione dei reclami»⁷⁸. Le norme sui meccanismi di segnalazione e azione dei contenuti «dovrebbero essere armonizzate a livello dell'Unione, in modo da consentire il trattamento tempestivo, diligente e non arbitrario

⁷⁷ Sulle innovazioni introdotte dal DSA, cfr. A. Iannotti Della Valle, *Libertà di espressione e valori democratici alla prova dei "social media": il DSA e un nuovo caso TikTok europeo*, in *federalismi.it*, 13, 2025, 84 ss.; N. Pica, *La tutela della libertà di informazione nel "Digital Services Act" tra contrasto alle "manipolazioni algoritmiche" e limiti alla "content moderation"*, in questa Rivista, 1, 2024, 12 ss.; I. de Vivo, *Il potere d'opinione delle piattaforme-online: quale ruolo del "regulatory turn" europeo nell'oligopolio informativo digitale?*, in *federalismi.it*, 2, 2024, 45 ss.; S. Sassi, *L'Unione Europea e la lotta alla disinformazione "online"*, in *federalismi.it*, 15, 2023, 183 ss.; E. Longo, *"Digital Services Act" - Libertà di informazione. Libertà di informazione e lotta alla disinformazione nel "Digital Services Act"*, in *Giorn. dir. amm.*, 6, 2023, 737 ss.; L. D'Agostino, *Disinformazione e obblighi di "compliance" degli operatori del mercato digitale alla luce del nuovo "Digital Services Act"*, in questa Rivista, 2, 2023, 16 ss.; E. Birritteri, *Contrasto alla disinformazione, "Digital Services Act" e attività di "private enforcement": fondamento, contenuti e limiti degli obblighi di "compliance" e dei poteri di autonormazione degli operatori*, *ivi*, 52 ss.; M.E. Bucalo, *La libertà di espressione in rete fra "content moderation" dei "social network" e regolazione dell'Unione Europea*, in *dirittifondamentali.it*, 3, 2022, 270 ss.

⁷⁸ Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali), considerando 45.

delle segnalazioni sulla base di norme uniformi, trasparenti e chiare, che forniscano solide garanzie a tutela dei diritti e degli interessi legittimi di tutte le parti interessate, in particolare dei loro diritti fondamentali garantiti dalla Carta»⁷⁹.

La *shadowban* viene esplicitamente nominato nel considerando 55 del DSA, secondo la seguente definizione: «La restrizione della visibilità può consistere nella retrocessione nel posizionamento o nei sistemi di raccomandazione, come pure nella restrizione dell'accessibilità da parte di uno o più destinatari del servizio o nell'esclusione dell'utente da una comunità online senza che quest'ultimo ne sia consapevole («shadow banning»).

Il considerando 70 riconosce che un elemento essenziale dell'attività di una piattaforma online consiste nel modo in cui le informazioni sono messe in ordine di priorità e presentate nella sua interfaccia online per facilitare e ottimizzare l'accesso alle stesse da parte dei destinatari del servizio. Ciò avviene, ad esempio, suggerendo, classificando e mettendo in ordine di priorità le informazioni in base ad algoritmi, distinguendole attraverso testo o altre rappresentazioni visive oppure selezionando in altro modo le informazioni fornite dai destinatari. Tali sistemi di raccomandazione possono avere un impatto significativo sulla capacità dei destinatari di recuperare e interagire con le informazioni online. Come viene riconosciuto dal medesimo Considerando, essi svolgono inoltre un ruolo importante nell'amplificazione di determinati messaggi, nella diffusione virale delle informazioni e nella sollecitazione del comportamento *online*. Di conseguenza, le piattaforme dovrebbero provvedere in modo coerente affinché i destinatari del loro servizio siano adeguatamente informati in ordine a tutti questi aspetti. Esse dovrebbero indicare chiaramente i parametri di tali sistemi di raccomandazione in modo che i destinatari del servizio comprendano la modalità con cui le informazioni loro presentate vengono messe in ordine di priorità.

Inoltre, dal punto di vista procedurale e delle regole sulla trasparenza, la legge sui servizi digitali stabilisce che le piattaforme online devono indicare nei loro termini e condizioni come utilizzeranno i processi decisionali algoritmici e come questi processi debbano rispettare i diritti fondamentali previsti dalla Carta dei diritti fondamentali dell'UE⁸⁰, inoltre viene stabilito che i processi devono essere equi, non arbitrari e trasparenti (considerando 38b)⁸¹.

In termini generali, il regolamento introduce una serie di obblighi per le piattaforme, cui corrispondono altrettanti diritti in capo agli utenti: la piattaforma deve fornire all'utente una dichiarazione chiara e specifica delle ragioni per cui applica la restrizione (art. 15); deve essere garantita un'efficace procedura interna di reclamo (art. 17) che va spiegata nei termini e condizioni (art. 12); occorre disciplinare il meccanismo per la risoluzione extragiudiziale delle controversie (art. 18) e, infine, che sia specificato il diritto a un ricorso giudiziario (art. 15, par. 2, lett. f)). Queste garanzie

⁷⁹ DSA, considerando 52.

⁸⁰ DSA, considerando 38.

⁸¹ DSA, considerando 38/b.

procedurali si applicano indipendentemente dal fatto che il contenuto fosse o meno illegale o che la restrizione fosse imposta in base ai termini e alle condizioni della piattaforma stessa.

Il DSA delinea un quadro normativo applicabile a tutte le forme di moderazione dei contenuti denominato meccanismo di segnalazione e azione (*notice-and-action framework*), di particolare interesse per il tema che qui ci occupa⁸². Tale meccanismo si applica a tutte le azioni di moderazione dei contenuti, in un'accezione ampia dell'espressione. I primi interventi normativi sul tema si erano concentrati quasi esclusivamente sulla rimozione dei contenuti o sulla sospensione degli *account*, al punto che in dottrina si è parlato di «*binary leave up/take down paradigm*»⁸³. Il DSA si pone in modo innovativo da questo punto di vista, in quanto accoglie una nozione molto più ampia di moderazione dei contenuti, tale da ricomprendere – come rilevato nel considerando 55 – anche lo *shadowban*. Anche se l'accezione nella quale sembra essere stata accolta rimanda al senso originario (e più ristretto) di sospensione degli *account* non notificata, ciò non modifica il quadro giuridico predisposto dal regolamento, in quanto nel testo dell'articolo tale espressione non ricorre più esplicitamente⁸⁴.

Il meccanismo *notice-and-action* viene disciplinato dall'art. 16 del DSA, che stabilisce due principi fondamentali: in primo luogo, le regole relative alla moderazione dei contenuti riguardano «le politiche, le procedure, le misure e gli strumenti utilizzati ai fini della moderazione dei contenuti, compresi il processo decisionale algoritmico e la verifica umana, nonché le regole procedurali del loro sistema interno di gestione dei reclami» e devono essere pubblicate dalle piattaforme alla voce “Termini e condizioni”, redatte «in un linguaggio chiaro, semplice, comprensibile, facilmente fruibile e privo di ambiguità».

Tale regola riflette nel DSA il principio di legalità, prevedibilità e trasparenza nell'imposizione delle sanzioni. L'art. 14 del DSA non introduce principi fino a quel momento ignoti alla governance delle piattaforme, ma serve piuttosto a chiarire e a rafforzare obblighi già sussistenti in capo alle medesime. La gran parte di esse infatti pubblicava le regole sulla moderazione dei contenuti: pur divenendo via via più dettagliate nel corso del tempo, sono state comunque soggette a critiche proprio per la non piena trasparenza e per l'insufficiente grado di dettaglio, e l'art. 14 potrebbe costituire un incentivo a disciplinare in modo più puntuale le regole sulla moderazione dei contenuti, soprattutto con riferimento a quei rimedi – come lo *shadowban* – che hanno spesso il loro presupposto in contenuti definiti *borderline* o genericamente inappropriati⁸⁵.

L'art. 17 del DSA richiede che ogni attività di moderazione dei contenuti sia accompagnata da una motivazione chiara e specifica per le restrizioni imposte a causa dell'asserita illegalità dei contenuti o della incompatibilità

⁸² P. Leerseen, *An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation*, in *Computer Law & Security Review*, (48)2023, 5.

⁸³ E. Goldman, *Content moderation remedies*, in *Michigan Technol Law Rev*, (28)1, 2021.

⁸⁴ P. Leerseen, *An end to shadow banning?*, cit., 2.

⁸⁵ Ivi, 6; A. Heldt, *Borderline speech: Caught in a free speech limbo?*, cit.; M. Horten, *Algorithms patrolling content: where's the harm?*, cit.

con le condizioni generali imposte dalla piattaforma. La motivazione deve includere, in particolare, le seguenti informazioni: il tipo di misura assunta; la violazione contrattuale o legale che il contenuto ha determinato; se la decisione si basa su un procedimento decisionale automatizzato; se la decisione sia stata assunta a seguito di segnalazione proveniente da un soggetto terzo; le possibilità di replica dell'utente. La disposizione in esame assolve quindi due compiti distinti, individuabili nella notificazione della decisione e nella sua motivazione. Dal momento che lo *shadowban* si caratterizza proprio per la totale mancanza di una notificazione all'utente che ne viene colpito, è possibile dedurre un divieto di utilizzo di tale pratica a norma del DSA. L'applicazione di tale misura non dovrebbe essere imposta in modo arbitrario, gli utenti devono essere informati, le motivazioni devono essere fornite, e andrebbe garantita l'intera gamma delle opzioni di ricorso, compresa quella giudiziaria⁸⁶.

Il punto vero della questione è che ad oggi non è ancora chiaro in cosa consista il risultato della *demotion* di un elemento all'interno di una piattaforma *online*. Numerosi commentatori avevano già rilevato il problema di definire in via di fatto in cosa consista tale pratica e, oltre a ciò, come provarlo dal punto di vista fattuale⁸⁷. Tali violazioni sarebbero difficili da inquadrare, anche perché non è possibile definire il livello di visibilità cui qualcuno ha diritto⁸⁸.

La difficoltà di rendere trasparenti le politiche sulla moderazione dei contenuti nel caso della *demotion* sta proprio nella mancanza di un parametro in grado di indicare il livello medio di comparazione tale da provare che quel contenuto è stato degradato nella classifica. Una ricostruzione del concetto di *demotion* inquadrabile anche dal punto di vista giuridico dovrebbe partire dall'idea che il processo di classificazione della piattaforma non consiste nell'applicazione di un singolo algoritmo monolitico, ma si basa sull'applicazione di una serie di algoritmi che lavorano in sequenza su più fasi e con funzioni diverse. Nell'ambito di tali meccanismi, è possibile distinguere alcuni sistemi diretti ad attribuire maggiore visibilità ai contenuti, nella prospettiva del massimo coinvolgimento degli utenti, e altri che impongono a posteriori *malus* o *bonus* su questi punteggi basati su ulteriori processi di moderazione, rilevando i fenomeni di *clickbait* o di *bate speech*. In altre parole, alcuni algoritmi attribuiscono punteggi che aumentano la visibilità, mentre altri la riducono. Il primo meccanismo punta tutto sulla massimizzazione dell'*engagement*, mentre il secondo interviene prevalentemente per assicurare il rispetto dei termini e delle condizioni della piattaforma, costituendo in questo secondo caso un'azione di moderazione vera e propria: questa interpretazione inquadra il problema dello *shadowban* prendendo l'ottimizzazione dell'*engagement*

⁸⁶ P. Leerseen, *An end to shadow banning?*, cit., 7.

⁸⁷ R. Griffin, *Rethinking rights in social media governance: human rights, ideology and inequality*, in *SSRN Draft Paper*; T. Gillespie, *Do not recommend? Reduction as a form of content moderation*, *Social Media + Society*, (8)3, 2022; D. Keller, *Amplification and its discontents: why regulating the reach of online content is hard*, in *J Free Speech Law*, (227)1, 2021, 68 ss.; L. Thorburne-J. Stray-P. Bengagina, *What will "amplification" mean in court?*, in *Tech Policy Press*, 2022; P. Leerseen, *An end to shadow banning?*, cit., 8.

⁸⁸ R. Griffin, *Rethinking rights in social media governance: human rights, ideology and inequality*, cit.

come trattamento di base, rispetto al quale si possono poi individuare le riduzioni⁸⁹.

Un primo passo, già compiuto con il DSA, è stata l'imposizione di una regola generale sulla trasparenza nei sistemi di raccomandazione dei contenuti, in forza della quale le piattaforme devono rendere noti «i principali parametri utilizzati nei loro sistemi di raccomandazione, nonché qualunque opzione a disposizione dei destinatari del servizio che consente loro di modificare o influenzare tali parametri principali». Inoltre, questa disposizione prevede che le piattaforme indichino «le ragioni per l'importanza relativa di tali parametri»⁹⁰: il punto meritevole di attenzione è che tali regole non si riferiscono alla moderazione dei contenuti – di cui la retrocessione costituisce un particolare esempio – ma alla *curation of contents* in un senso più ampio, che non si limita esclusivamente alla loro *demotion*⁹¹.

Regole di codifica generali come queste, tuttavia, incontrano ancora importanti limitazioni. Anzitutto, non sembrano sufficienti a gettare una luce significativa sul funzionamento concreto di tali algoritmi, prevalentemente a causa della complessità che li caratterizza. Inoltre, l'art. 27 non richiede nemmeno spiegazioni approfondite, essendo richiesta una mera descrizione dei parametri principali degli algoritmi. Nel peggiore dei casi, queste descrizioni potrebbero essere a tal punto generiche da non offrire alcun chiarimento effettivo. Tuttavia, ipotizzando un'implementazione più solida, tali regole potrebbero comunque funzionare come un incentivo alla trasparenza nella moderazione dei contenuti individuali: qualora un utente sperimentasse un improvviso calo di visibilità senza aver ricevuto alcuna notifica di azioni di moderazione individuali ai sensi dell'art. 17 delle DSA, ciò potrebbe indurlo a verificare la presenza di aggiornamenti generali delle politiche di curatela ai sensi dell'art. 27.

7. Quale destino per lo *shadowban*? Verso un (auspicabile) superamento, per salvaguardare il dibattito politico e la partecipazione democratica

In questa sede si è cercato di analizzare in una prospettiva costituzionalistica la recentissima imposizione di uno *shadowban* sui contenuti di natura politica applicata dal Gruppo Meta.

All'esito della ricerca, è possibile formulare alcune – sia pure parziali – considerazioni conclusive. Come si è avuto modo di analizzare, lo *shadowban* è una tecnica di moderazione dei contenuti ampiamente utilizzata da tutte le piattaforme *online*, che consiste nella limitazione della visibilità dei contenuti degli utenti senza che questi ne siano messi a conoscenza.

Sono noti i benefici derivanti da tale pratica, come la de-prioritizzazione delle *fake news*, dei contenuti di *hate speech* o addirittura in contrasto con termini e condizioni predisposti dalle piattaforme, fino allo stemperamento

⁸⁹ P. Leerseen, *An end to shadow banning?*, cit., 8 ss.

⁹⁰ DSA, art. 27.

⁹¹ P. Leerseen, *An end to shadow banning?*, cit., 12.

della polarizzazione tra gli utenti.

Tuttavia, si è rilevato come tale pratica presenti numerosi profili problematici, soprattutto quando abbia ad oggetto contenuti di carattere politico, peraltro da individuare secondo una definizione vaga e imprecisa come quella fornita dal Gruppo Meta, che lascia un margine di discrezionalità enorme in capo alle piattaforme su cosa includere (o escludere) nella restrizione. Ad oggi non risulta dimostrato che lo *shadowban* abbia avvantaggiato alcune ideologie politiche o singoli candidati, anche se l'indagine aperta dalla Commissione europea nei confronti del Gruppo Meta sulla questione non sembra ancora essersi conclusa, lasciando insoluti i dubbi circa gli effetti che tale pratica sia in grado di produrre sulle campagne elettorali, come evidenziato dalle interrogazioni parlamentari presentate da alcuni eurodeputati.

Fermo restando quanto appena precisato, è però altrettanto dimostrato che lo *shadowban* è capace di influenzare il processo di formazione delle opinioni attraverso i *social network*, sia nel senso di sviare che di indirizzare l'attenzione degli utenti in relazione a determinati argomenti. Occorre poi sottolineare che la circostanza che tale tecnica di moderazione dipenda dalla frequenza con cui vengono postati i contenuti nonché, in alcuni casi, dalla loro diffusività, si pone in contrasto con la funzione tipica della propaganda politica tramite i *social network*, che vede nelle due caratteristiche appena elencate la propria ragion d'essere primaria.

Vi sono poi effetti ancor più insidiosi che lo *shadowban* può determinare sulla partecipazione politica dei cittadini. Anzitutto, il fatto che le regole da cui dipende la sua applicazione siano per l'appunto oscure genera un senso di frustrazione negli utenti che si vedono imposte tali limitazioni, che può condurre a sfiducia nelle piattaforme, radicamento di teorie del complotto, isolamento dei singoli e maggiore contrapposizione tra utenti, ottenendo effetti opposti rispetto alla più volte evidenziata funzione di stemperamento del dibattito.

Proiettando tali effetti su una dimensione collettiva, lo *shadowban* sui contenuti politici impoverisce il panorama informativo nell'ambito dei *social media* a favore di contenuti a carattere esclusivamente commerciale, impedisce ai cittadini la formazione di un'opinione attraverso il confronto tra notizie e contenuti di carattere politico, alimentando anche la creazione delle *echo chambers*, contribuendo, in ultima analisi, al disinteresse nei confronti del dibattito pubblico: tutto ciò si traduce in una disaffezione dei cittadini verso la politica e, in particolare, verso il momento elettorale, contribuendo all'aumento dell'astensionismo già in sé preoccupante per la legittimazione delle assemblee rappresentative e per la tenuta di un sistema democratico.

Come si è avuto modo di dimostrare, lo *shadowban* non è soltanto un fenomeno rilevante dal punto di vista del funzionamento delle piattaforme di *social media* e delle conseguenze che è in grado di produrre sul piano sociologico e politologico, ma è in grado di tradursi anche in specifiche violazioni di diritti fondamentali. Dal punto di vista procedurale, la circostanza che si tratti di limitazioni imposte in via unilaterale, arbitraria, senza preavviso e in assenza di una violazione esplicita dei termini e delle condizioni compromette il diritto dell'utente di impugnarle dinanzi alle

medesime piattaforme. La più evidente e immediata tra le lesioni dei diritti fondamentali è la libertà di manifestazione del pensiero, compreso nella misura in cui l'applicazione di uno *shadowban* equivale nei fatti ad una censura vera e propria. A essere compromessa è poi la libertà d'informazione, da intendersi sotto il versante del diritto ad essere informati. Inoltre, il fatto che ne risultino colpiti con più facilità profili recenti, non verificati e non appartenenti a figure politiche pone anche un problema dal punto di vista della tutela del principio di eguaglianza. La lesione di rilevanti diritti fondamentali a causa dell'applicazione dello *shadowban* ai danni di un politico belga è stata riconosciuta dal giudice nazionale, in quanto tale misura di moderazione dei contenuti era il risultato di un procedimento automatizzato e previa profilazione sulla base delle preferenze politiche di un utente è stato ritenuto in contrasto con il GDPR.

Ad oggi, rimangono insoluti numerosi interrogativi sulla portata e sugli effetti che tale tecnica di moderazione dei contenuti è in grado di produrre sul dibattito pubblico e politico. Se da un lato non è provato che possa influenzare le preferenze degli elettori a favore o contro determinati candidati o ideologie politiche, dall'altro non possono essere taciute le preoccupazioni – peraltro efficacemente espresse nelle interrogazioni presentate al Parlamento europeo da alcuni eurodeputati – sui suoi effetti nelle campagne elettorali, sia a causa della vaghezza della definizione di contenuto politico adottata dal Gruppo Meta, sia in ragione della opacità dei meccanismi di funzionamento sottesi alla sua applicazione. È fin troppo facile immaginare come attraverso l'oscuramento di determinate notizie o dei profili di alcuni candidati o partiti politici sarebbe possibile orientare il consenso degli elettori in altre direzioni, facendo leva sulla vaghezza della definizione di contenuto politico e, soprattutto, sull'inesistente trasparenza dei meccanismi applicati dalle piattaforme.

In questa sede non è possibile indicare quali potrebbero essere le soluzioni sul futuro dello *shadowban*, ma è possibile tracciare alcune coordinate dentro le quali il fenomeno potrebbe (e, forse, dovrebbe) essere indirizzato. Muovendo dalla disciplina introdotta dal DSA, gli obblighi di trasparenza imposti alle piattaforme sulla moderazione dei contenuti devono includere le informazioni relative al procedimento decisionale algoritmico e sulla verifica umana; nonché, qualora venga applicata una restrizione ad un contenuto – tra cui rientra anche la retrocessione dello stesso – deve essere fornita all'utente una motivazione chiara e specifica: tale tecnica di moderazione dei contenuti dovrebbe essere destinata a cadere sotto la scure del DSA, pur non conoscendo l'esito dell'istruttoria avviata dalla Commissione sulle attività del Gruppo Meta anche con riferimento allo *shadowban* dei contenuti politici.

Rimane sullo sfondo il grande interrogativo su che cosa ci si aspetta che i *social network* diventino. È chiaro che considerandoli mere imprese commerciali operanti sul mercato non si può pretendere che esse si facciano carico degli effetti di quanto avviene al loro interno sulla collettività, dal punto di vista del dibattito pubblico e politico. Tenendo conto però del ruolo che essi hanno assunto nell'informazione (anche politica), degli effetti che sono in grado di produrre sul dibattito pubblico e dell'influenza che esercitano sulle campagne elettorali non è più immaginabile ignorarne

il ruolo, continuando a ridurli a meri operatori privati. Questa sembra essere la direzione intrapresa dall'ordinamento europeo, che attraverso il regolamento (UE) 2024/900 ha introdotto disposizioni armonizzate sulla trasparenza, sugli obblighi e doveri di diligenza per la prestazione di servizi di pubblicità politica, sull'uso di tecniche di *targeting* e di consegna dei messaggi pubblicitari che comportano il trattamento dei dati personali nell'ambito della pubblicità politica online⁹².

Come si è cercato di dimostrare in queste pagine, lo *shadowban* dei contenuti politici costituisce uno snodo rilevante nello studio delle campagne elettorali *online* e dei loro effetti sui diritti di partecipazione politica: se la direzione verso cui l'ordinamento europeo sembra muovere è quella di rendere i *social network* ambienti nei quali trovare un'informazione politica di qualità, garantendo il pluralismo e contribuendo allo sviluppo di un dibattito pubblico sano, il primo passo da compiere è quello di preservare l'esistenza di un ecosistema informativo nel quale sia assicurata l'esistenza dei contenuti di carattere politico.

⁹² C. Lobascio, *Micro-targeting politico e protezione dei dati personali. Il Regolamento (UE) 2024/900*, in *Riv. it. inf. dir.*, 1, 2025, 205 ss.

Abstract

Il saggio si pone l'obiettivo di analizzare l'impatto che lo *shadowban* dei contenuti politici applicato dal Gruppo Meta è in grado di produrre sulle campagne elettorali *online* e, in termini più generali, sulla partecipazione democratica. La ricerca si sofferma anzitutto sulla definizione di tale tecnica di moderazione dei contenuti, analizzando la sua evoluzione nel tempo. Il lavoro si concentra poi sulle conseguenze dello *shadowban*, esaminando dapprima gli effetti positivi che vengono addotti dalle piattaforme a giustificazione della sua adozione, e soffermandosi poi sugli aspetti negativi che determina sul dibattito pubblico. Il saggio esamina la lesione dei diritti fondamentali che derivano dallo *shadowban*, prendendo in considerazione la recente sentenza di un giudice belga che ha condannato Meta al risarcimento del danno a favore di un politico i cui contenuti erano stati oggetto di retrocessione illegittima. Infine, la ricerca si concentra sull'introduzione del divieto di *shadowban* nel Digital Services Act, formulando alcune considerazioni conclusive sulle possibili evoluzioni di tale tecnica di moderazione dei contenuti.

The essay aims to analyse the impact that Meta Group's shadowbanning of political content can have on online election campaigns and, more generally, on democratic participation. The research focuses first on defining this content moderation technique, analysing its evolution over time. The work then focuses on the consequences of shadowbanning, first examining the positive effects that platforms cite as justification for its adoption, and then focusing on the negative aspects it can have on public debate. The work examines the violation of fundamental rights resulting from shadowbanning, taking into account the recent ruling by a Belgian judge who ordered Meta to pay damages to a politician whose content had been unlawfully downgraded. Finally, the research focuses on the introduction of the shadowbanning ban in the Digital Services Act, offering some concluding remarks on the possible developments of this content moderation technique.

Keywords

shadowban – campagne elettorali *online* – moderazione dei contenuti – Gruppo Meta – Digital Services Act