How to cover malign influence campaigns: a view from the Philippines

By Regine Cabato

Contents

Introduction

I was decapitated at Oxford University. At least, that's what an anonymous sender told my mother in a text message.

"She was seen inside her room naked and her neck was cut off!" they wrote, claiming to be from the Manila Embassy.

I blinked at a photo of the message forwarded by my sister. I was wrapped in four layers, but that January morning suddenly felt chillier.
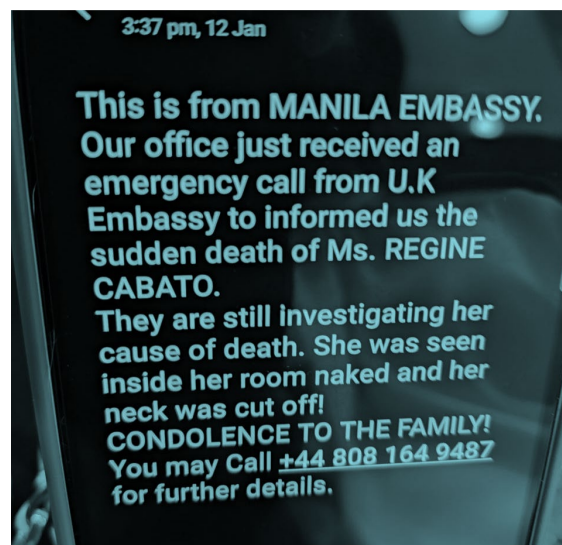
A typo-riddled text message my mother received in January 2023, informing her of my alleged sudden and violent death in the United Kingdom.

On instinct, I looked around the uncrowded bus cruising on the way to London. After determining there was no threat from the elderly couple near the door, I let myself feel the shock first… and then the laughter. 'A *head* is cut off,' I thought, 'not a *neck*.'

When I began my fellowship at the Reuters Institute for the Study of Journalism last autumn – exhausted by marathon coverage of a disinformation-riddled election in my home country – I kept a low profile in the hope harassment wouldn't follow me.

In the Philippines, a colourful archipelago of some 100 million people just west of the Pacific, a young and digitally savvy population has helped power a growing digital economy that includes call centres, platform gig work, and an underground industry of disinformation and influence operations.

The country is one of the world's most chronically online nations, and has earned the moniker of "patient zero" of the post-truth era.[1]

In 2016, it elected Rodrigo Duterte: a gun-toting, tough-talking populist whom the Oxford Internet Institute discovered spent a whopping PhP 10 million – or $200,000 – on a keyboard army to prop up his image.[2] His rise to power thrust local disinformation operations into the public consciousness. When U.S. president Donald Trump was elected that same year, the Philippines was widely seen as a precedent and branded a testing ground – a "petri dish" – for disinformation tactics that would be exported around the world.[3]

What Duterte began in 2016 was perfected a presidency later. Ferdinand Marcos Jr., son and namesake of a dictator who plundered up to $10 billion from the country, emerged the victor of the 2022 presidential election.[4] The win came on the back of an extensive online and offline campaign to whitewash his family history.[5] His running mate, vice president Sara Duterte-Carpio, is Duterte's daughter.

It took a decades-long effort to return the Marcos family to power following the dictator's ousting in the 1986 revolution. After living in exile in Hawaii, the family returned to the Philippines – and subsequently, their political careers – in the 90s. Their historical revisionism project moved online in the early 2000s, and social media algorithms that favour sensationalism gave them unprecedented reach over a decade later. Each member of the Marcos political dynasty has cultivated their own online personality and support base. Their network of official pages collaborates with each other and with hyper-partisan channels, anonymous personalities, and mega- and micro-influencers across Facebook, YouTube, and TikTok.[6]

Marcos has denied using troll farms, but data shows evidence of coordinated, manipulated networked propaganda surrounding much of his online support.[7] He was the biggest beneficiary of disinformation, while opposition candidate Leni

[1] *Facebook's Katie Harbath on Protecting Election Integrity.* Rappler, 24 June 2018, www.youtube.com/watch?v=dJ1wcpsOtS4&t=76s.

[2] Bradshaw, Samantha, & Howard, Philip. *Troops, Trolls, and Troublemakers: A Global Inventory of Organized Social Media Manipulation.* Oxford Internet Institute, 2017.

[3] Occeñola, Paige. "PH Was Cambridge Analytica's "Petri Dish" – Whistle-Blower Christopher Wylie." *Rappler*, 10 Sept. 2019, www.rappler.com/technology/social-media/239606-cambridge-analytica-philippines-online-propaganda-christopher-wylie/

[4] Davies, Nick. *The $10 Billion Question: What Happened to the Marcos Millions?* The Guardian, 7 May 2016, www.theguardian.com/world/2016/may/07/10bn-dollar-question-marcos-millions-nick-davies

[5] Cabato, Regine, & Mahtani, Shibani. *How the Philippines' Brutal History Is Being Whitewashed for Voters.* The Washington Post, 12 Apr. 2022, www.washingtonpost.com/world/2022/04/12/philippines-marcos-memory-election/.

[6] Soriano, C.R.R., & Gaw, F. (2022). Platforms, alternative influence, and networked political brokerage on YouTube. Convergence, 28(3), 781-803. https://doi-org.ezproxy-prd.bodleian.ox.ac.uk/10.1177/13548565211029769

[7] Tuquero, Loreben. *Tracking the Marcos Disinformation and Propaganda Machinery.* Rappler, 14 Apr. 2022, www.rappler.com/newsbreak/iq/stories-tracking-marcos-disinformation-propaganda-machinery/.

Robredo was its biggest target.[8] In January 2022, Twitter suspended hundreds of pro-Marcos accounts for violating rules on spam and manipulation.[9] And his support base dominates Facebook and YouTube, with eight out of ten videos on the latter platform seeking to rewrite the Marcos family's history. [10, 11]

What began as a fairly unsophisticated copy-paste trolling scheme using bots has morphed to incorporate real influencers – from celebrities with lifestyle vlogs to self-anointed political commentators to TikTok creators with mass appeal.

Inadequate campaign finance laws mean there's no clear money trail to track either political endorsements-for-hire or creators who cash in through platforms' built-in monetizing incentives. The propaganda network has become so encompassing that it's almost impossible to tell the inauthentic from the organic. "Perception is real," Imelda Marcos herself said in Lauren Greenfield's 2019 documentary *The Kingmaker*, "and the truth is not".

I have been reporting on the Marcos family since 2016, and on the local disinformation landscape, too: sitting in a troll farm ahead of the 2019 midterms, and monitoring fake fan cams revering Marcos' son as a heartthrob. From this front-row seat to the Philippine's changing psyche both on- and offline, it's been clear that journalists are no longer the gatekeepers. There is a fracture in Filipinos' shared reality, creating what academics have called "a political multiverse of madness" and "parallel public spheres". In one sphere, documentation from criminal courts shows the source of Marcos wealth to be embezzlement; in another, court records are rejected in favour of believing that the Marcos's wealth was handed down by a non-existent king.[12]

What can journalists reporting from the Global South do when covering such fractured realities, particularly when they lack the data analysis and OSINT teams that large Western outlets can afford to counter disinformation? I wanted to create a realistic guide for colleagues. To do so, I have archived journalists' experiences of the 2022 campaign, compiled lessons, and compared notes with foreign reportage on the same themes. I spoke to 23 journalists who covered the Philippine election and disinformation ecosystem, both independently and across ten news organisations in the Philippines and the foreign press. I also interviewed a political strategist and drew from the recommendations of various academics. Finally, I spent time with the disinformation team at BBC to see what might be transposable.

[8] Tsek.ph. *Disinformation Attempts to Rehabilitate the Marcoses, Disparage EDSA*. Tsek PH, 25 Feb. 2022, www.tsek.ph/disinformation-attempts-to-rehabilitate-the-marcoses-disparage-edsa/.

[9] Reuters. Twitter suspends hundreds of accounts promoting Philippines' Marcos. 21 Jan. 2022. https://www.reuters.com/world/asia-pacific/twitter-suspends-hundreds-accounts-promoting-philippines-marcos-2022-01-21/

[10] https://www.facebook.com/watch/live/?ref=watch_permalink&v=1725969461069725

[11] Soriano, C. R. R., & Gaw, F. (2022). See footnote 6.

[12] Teehankee, J.C., 2023. Beyond nostalgia: the Marcos political comeback in the Philippines.

The pro-Marcos online network may have started off as obviously inauthentic, but it is now a well-funded, well-oiled, self-sustaining machine. Like a hydra, if one head falls, more grow in its place. When I began my fellowship project, I wanted to find out: how does one kill a hydra?

Defining & measuring influence operations

In the Philippines, influence operations that prop up politicians on social media are largely the result of work by the advertising and public relations industry. The phenomenon and its role in the 2016 elections was first extensively documented in the Philippines in [Architects of Networked Disinformation](#), by Jonathan Ong and Jason Cabanes.[13] Since then, these operations have only expanded and grown more sophisticated, adapting to evade social media company crackdowns and present a more convincing, authentic-passing front to audiences on social media.

In the mainstream media, the bulk of news coverage on the topic consisted of frequent fact-checks to debunk disinformation, and the occasional analytical or investigative report examining the phenomenon.

Following 2022's presidential elections, Filipino academics have called for the reframing of disinformation through the lens of influence operations. Researchers argue that a broader framework would look at actions that were "neither illegal nor deceitful but […] exploitative" of grey areas in laws, platform policies, and journalistic rules (Parallel Public Spheres, 2022).[14] Similarly, a 2023 [report supported by Internews](#) proposed a new metric to assess inauthentic behaviour and, by this measure, found election-related digital spending of PhP 1 billion ($17 million) in paid political endorsements from influencers, with one content creator being paid PhP 3 million ($52,000) for a single video.[15]

Plenty of problematic information nowadays does not fall strictly under the definition of disinformation, which presupposes malicious intent. This intent can be harder to establish in an information landscape where political discourse is more scattered across a cottage industry of micro- and nano-influencers on more varied platforms than it was six years ago, when political discourse was more concentrated around fewer but prominent serial disinformers.

On the 2022 campaign trail, for example, some smaller content creators were genuine fans as opposed to political hacks directly on a politician's payroll.[16] That

[13] Ong, Jonathan Corpus, and Jason Vincent Cabañes. "Architects of Networked Disinformation: Behind the Scenes of Troll Accounts and Fake News Production in the Philippines." *Architects of Networked Disinformation: Behind the Scenes of Troll Accounts and Fake News Production in the Philippines*, 2018 https://doi.org/10.7275/2cq4-5396

[14] Ong, J.C., Fallorina, R., Lanuza, J.M.H., Sanchez, F., and Curato, N. (2022). "Parallel Public Spheres: Influence Operations in the 2022 Philippine Elections". The Media Manipulation Case Book. Internews and Harvard Kennedy School Shorenstein Center. https://mediamanipulation.org/research/unmasking-influence-operations-in-the-philippines.

[15] Gaw, F., Bunquin, J.B., Cabbuag, S., Lanuza, J.M., Sapalo, N., & Yusoph, A. (2023). Political Economy of Covert Influence Operations in the 2022 Philippine Elections. Internews. https://internews.org/resource/political-economy-of-covert-influence-operationsin-the-2022-philippine-elections

[16] Talabong, Rambo. "[The New Media of the New Society](#)." 2022.

said, many of those fans were also taking advantage of YouTube's Partnership Program to reap a profit.[17]

One definition of influence operations, offered by Thomas, Thompson, Wanless, describes them as "organised attempts to achieve a specific effect among a target audience."[18] While malign influence operations dominate the news, the study points out that positive influence operations like voter registration campaigns also exist.

Other political influence operations in the Philippines also occupy a grey area. These include white-hat operations – in my previous reporting, one operator self-identified as managing "white trolls"– who also participate in coordinated inauthentic behaviour, but who self-impose their own ethos like avoiding cursing and black propaganda, and referencing news sites to debunk their opponents.

During the 2022 election campaign, I spoke to a self-organised, pro-opposition volunteer "truth army" who said they used their real accounts to counter malign operations by flocking to posts that were swarmed with hate.

Considering influence as opposed to disinformation also allows us to look at the broader structures that enable and shape problematic information, which narratives are at play, and its movement both online and offline.

In the Philippines, many factors influence elections: name recall, money, command votes from political figures in power, and so on. Alan German, president of Agents International, a public relations and political communications company in Manila describes himself as a "persuasion specialist". He said the electoral market could be captured in two parts: retail votes won by advertising and public relations strategies; and negotiated votes, won by political alliances, endorsements, and ground operations. He described his job as overseeing a cohesive campaign both in and out of social media, coordinating with creatives, influencers, logistics, and ground operators like neighbourhood officials who mobilise voters: "I am the conductor... I'm the guy that says 'green light', 'red light'."

"Influence operations are neither inherently good nor bad," Thomas, Thompson, and Wanless wrote, "and it is up to societies themselves to decide what conduct and responses are and are not acceptable."

Introducing the 'lie machine'

Some campaigns, though, are designed to be misleading, dehumanising, and even vicious. One way to frame malign influence operations is through looking at them as "lie machines", described by Oxford Internet Institute professor Philip Howard as

---

[17] Elemia, Camille. "In the Philippines, a Flourishing Ecosystem for Political Lies." *The New York Times*, 6 May 2022, www.nytimes.com/2022/05/06/business/philippines-election-disinformation.html.

[18] Thomas, Thompson, Wanless. "The Challenges of Countering Influence Operations." 2020. https://carnegieendowment.org/2020/06/10/challenges-of-countering-influence-operations-pub-82031

"social and technical mechanisms for putting an untrue claim into service of political ideology".[19]

Howard divides the lie machine into three parts: the producer, who is typically a political personality, party, or entity crafting or commissioning the lie; the distributor, which is the platform and social media algorithm amplifying the lie; and the marketer, typically a firm or strategist "who [profits] by selling the lie to you."
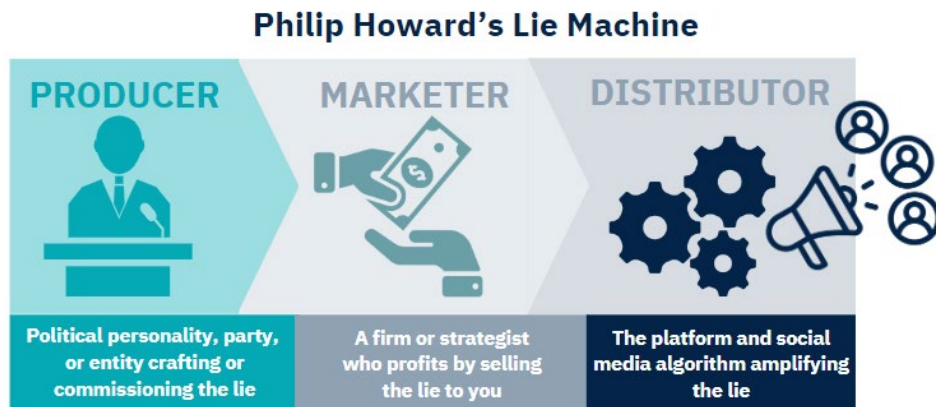
Philip Howard describes how producers, marketers, and distributors work together to disseminate lies in a digital age.

"The motivations for building lie machines are likely to remain constant: they are built either to make money or to promote ideology, and they are most powerful when both motivations are in play," Howard wrote. "Tracing this detailed process exposes the global political economy of big lies but also reveals what makes some democracies more resistant and resilient than others to the phenomenon."

That sounds like a job for good journalism. But those who undertake this work are often targeted by malicious influence campaigns themselves – across all parts of the production, distribution, and marketing cycle. There may be direct attacks by politicians, influencers and trolls repeating or fashioning new attacks, and social media algorithms amplifying the vitriol.

How much influence do influence operations really have?

Some figures, most famously Nobel Peace Prize winner and Rappler CEO Maria Ressa, have sounded the alarm on big tech's role in the deterioration of democracies. Ressa has argued that social media algorithms, and tech companies' inaction, has directly influenced elections and resulted in real-world violence.[20]

Others are hesitant to draw the same conclusion. In "Parallel Public Spheres," Filipino academics argued that social media "are powerful drivers of conversation

---

[19] Howard, Philip N. *Lie Machines: How to Save Democracy from Troll Armies, Deceitful Robots, Junk News Operations, and Political Operatives*. Yale University Press, 2020.
[20] Ressa, M. (2023) How to stand up to a dictator. S.l.: W H Allen.

but have no totalizing or 'brainwashing' effects."[21] They urged the local press to turn toward not just "platform takedowns or bot-busting but holding chief disinformation architects accountable and exposing the global and local political economic arrangements that allow them to thrive."

I asked campaign strategist German to what extent he thought online campaigns actually affected votes, given disinformation alone seems to have low conversion rates. He said, "the bandwagon effect [...] is real". In other words: the illusion of mass support (citing surveys, for example) can be a tool if the claim being supported is credible, authentic, and the perceived support is persistent.[22]

Bombardment, or *kuyog* ('swarm' in Filipino') can also affect a campaign, said German. The massive trolling of someone who posts in favour of a political candidate, for example, can intimidate individuals from speaking up, effectively "killing promotion".

In business, influencers tapped to promote brands and products are measured against a conversion rate, or the likelihood of audiences converted to buyers from engaging with content. While it is tricky to conclusively measure the same type of conversion in politics, it may be a factor: "Since particular emotional appeals are more effective at different times during the campaign, directing influencers to heighten particular sentiments at the right time with the right messaging can make voter conversion more likely," Filipino academics wrote in a recent study on influence operations in 2022.[23]

"People's beliefs are a product of accumulated information," which could include information normalized by propaganda and disinformation seeded and repeated through machinery over the years, one of the study's authors, Fatima Gaw, said in an interview.

Disinformation watchers may continue to wring hands over causality, but Howard points out in *Lie Machines* that the gradual and incremental change in communities, as well as the deterioration of public trust and institutions, is enough: [24]

> *"Average effects across an entire country are not what modern lie machines are designed to create. It is network-specific effects, sited in particular electoral districts, among subpopulations, that are sought. The ideal outcome for political combatants is not a massive swing in the popular vote but small changes in sensitive*

[21] Ong, J.C., Fallorina, R., Lanuza, J.M.H., Sanchez, F., and Curato, N. (2022). Parallel Public Spheres.
[22] Kate Starbird, Ahmer Arif, and Tom Wilson. 2019. Disinformation as Collaborative Work: Surfacing the Participatory Nature of Strategic Information Operations. Proc. ACM Hum.-Comput. Interact. 3, CSCW, Article 127 (November 2019), 26 pages. https://doi.org/10.1145/3359229
[23] Gaw, F., Bunquin, J.B., Cabbuag, S., Lanuza, J.M., Sapalo, N., & Yusoph, A. (2023). Political Economy of Covert Influence Operations in the 2022 Philippine Elections.
[24] Howard, P.N. *Lie Machines.* p. 110.

> *neighborhoods, diminished trust in democracy and elections,*
> *voter suppression, and long-term political polarization."*

As these operations – and the economy and technology surrounding it – become more complex, so too must the reporting that investigates the phenomenon.

Journalists can no longer afford to play whack-a-troll. A responsible and well-rounded response to influence campaigns will require civil society, socio-political organisations, and the media to target different parts of the lie machine in complementary and, ideally, coordinated ways.

In newsrooms, this might look like reportage that alternates between, for example:

- data or evidence connecting political actors to malign influence campaigns, or demonstrating how these campaigns are mounted (such as Tsek PH's analysis of fact-checks pointing to Marcos as a primary beneficiary, or an ABS-CBN investigation into how Facebook pages are bought and repurposed for elections);[25, 26]

- exposés targeting platform design and algorithms (such as Rappler's data analysis that found YouTube amplified a pro-Marcos influencer after he murdered someone in broad daylight);[27]

- and accountability journalism zooming in on prominent disinformation figures (such as Rappler's research revealing a prominent pro-Marcos YouTuber was spreading covid conspiracy theories, resulting in him taking down a significant number of his own videos).[28]

The media must carefully consider whose opinions, narratives, and interests they are giving mileage to, and why.

---

[25] Soriano, C. R. R., & Gaw, F. (2022). See footnote 6
[26] https://news.abs-cbn.com/business/11/08/21/fb-pages-stans-and-likes-for-sale-ahead-of-halalan-2022
[27] https://www.rappler.com/newsbreak/investigative/sympathizers-defend-ateneo-shooting-gunman-online/
[28] https://www.rappler.com/newsbreak/in-depth/youtube-vlogger-sangkay-janjan-takes-down-videos-following-rappler-investigation

Anatomy of a 'lie machine'

Influence campaigns abound in the Philippines for various political, commercial, promotional, and even geo-political purposes. Military and police networks have been involved in red-tagging activists and other civilians (see page 21), and a Chinese network was taken down by Facebook for posting content supportive of China's occupation of islands in South China Sea.[29]

But the most consequential influence machinery in the country is that of the Marcos family, rivalled only (and sometimes complemented) by the Dutertes. This chapter will summarise what we know so far about this machine, and the tactics it has deployed in the past.[30]

Seeding influence

How did the Marcos family whitewash a two-decade regime fraught with corruption and human rights abuse? Unlike the Dutertes – whose campaign strategist, Nic Gabunada, was name-dropped by Facebook when they took down his spammy network in 2019 – the marketers behind the Marcos machinery have largely succeeded in evading investigative exposés.[31] This may be a benefit of deploying a decentralised network of strategists, influencers, and myths.

The mythology surrounding the Marcoses did not start online, but is a continuation of propaganda the dynasty built around itself during the period of dictatorship. This included denial of corruption, defence of Marcos-era policies, and distorted statistics claiming the economy prospered under their rule.[32]

The Marcoses were among the first Filipino political figures to venture into the digital age, according to UP TWSC researchers Miguel Reyes and Joel Ariate, with a presence on websites like Flickr.[33] Some of the earliest conspiracy videos, posted over a decade ago, posited that the Marcos family did not steal wealth from government coffers, but instead received it from a non-existent ancient royal – a conspiracy theory that would come to be known as "Tallano gold."

With the invention of social media networks, what were once fringe conspiracies were put on blast. Disparate talking points defending the Marcos family became an elaborate web of lies inextricable from the other, amplified by algorithms and poor content moderation.

---

[29] Nimmo, B, Shawn Eib, C., Ronzaud, L. Operation Naval Gazing. Graphika, 2020. https://graphika.com/reports/operation-naval-gazing

[30] Soriano, C. R. R., & Gaw, F. (2022). See footnote 6

[31] https://www.nytimes.com/2020/09/29/business/rodrigo-duterte-facebook-philippines.html

[32] https://verafiles.org/articles/marcos-lies-still

[33] Interview with Reyes and Ariate in 2022, see also https://diktadura.upd.edu.ph/2023/05/09/buy-marcos-lies-book/

The mainstream press and public started paying more attention to historical distortion online in 2016, when Marcos Jr. sought the vice presidency and his father, the late dictator, was given a stealthy and controversial burial in the Cemetery of Heroes.[34] Although Marcos Jr. narrowly lost the vice presidency that year, the online campaign helped create an impression that most Filipinos did not mind, or were even in favour of giving the dictator state honours.

By the time Marcos announced his bid for the presidency in 2021, his campaign had a legacy of offline operations and over a decade of online propaganda under its belt. Then he scored a political alliance with presidential daughter and Davao mayor Sara Duterte Carpio. Experts have partially credited their election win to this alliance, which allowed political machineries of both dynasties from the northern and southern regions of the country to cooperate online and offline.[35]

In 2022, fact-checking collective Tsek PH found that Marcos was the top beneficiary of disinformation over election season, despite having registered zero spending on Facebook advertisements.[36,37]

Over the past six years, the Marcos family has grown its following across all major social media platforms. Here is an overview of the family's presence on the top platforms where disinformation spread.

Tsek PH found that up to 70% of false and misleading claims it corrected originated on Facebook.

In the year leading up to the election, the Digital Public Pulse Project (DPP) of the University of the Philippines-Diliman monitored online activity in relation to the elections; it found that Facebook communities were partisan.[38]

Although Marcos Jr. was among the politicians with the highest engagement on the platform, DPP said alone he was initially "a minority in cluster size" compared to president Rodrigo Duterte, who had first campaigned on the platform in 2016. However, Marcos got a boost by proxy by teaming up with the outgoing president's daughter Sara Duterte-Carpio, who was supported by the existing pro-Duterte network.

---

[34] https://www.rappler.com/newsbreak/iq/151667-timeline-ferdinand-marcos-burial-controversy/
[35] https://www.theguardian.com/world/2022/may/26/philippines-election-result-is-a-win-for-dynasty-politics
[36] Tsek PH (see footnote 9).
[37] Salazar, Cherry. *Robredo Leads, Marcos Snubs Advertising on Facebook.* Philippine Center for Investigative Journalism, 13 Jan. 2022, https://www.pcij.org/article/7746/robredo-leads-marcos-snubs-advertising-on-facebook.
[38] Bunquin, J. B. & Gaw, F. (2022). Digital Public Pulse: Mapping the digital landscape of Philippine politics. Philippine Communication Society Review, pp181-194.

DPP also found the prevalence of partisan Facebook groups using content produced by the politicians they supported. Meanwhile, non-partisan accounts and groups – like hobby, interest, and trade groups – also slid in partisan messaging in more subtle ways, an indication they were possibly run by "cloaked actors". The groups "may be strategically put up to pool together diverse users that serve as captive audiences once political content is shared in these spaces," DPP said.[39]

Some of these pages or groups might even change their names, sometimes expressing support for a politician outright, and inheriting all the likes and follows initially harvested for a non-partisan interest. One group for history enthusiasts with over half a million members was turned over to pro-Marcos administrators, who began sharing false claims and bullying members who spoke up against it.[40]

YouTube

Google's video-sharing platform was Marcos's online bailiwick, according to DPP.[41] His campaign "cemented its ownership of the election network" through three key communities: mainstream political actors, including his own channel and media channels affiliated with religious groups; entertainment channels and personalities; and smaller hyper-partisan influencers amplifying his campaign.

The president's own official channel boasts 2.73 million followers. During the campaign, he adapted family-friendly, feel-good content that borrowed from Buzzfeed-like games: a lie detector challenge with his family, quick questions with his sons, and reaction videos to messages from supporters. In December 2022, Marcos posted that he had received a Gold Play Button from YouTube, a tiered creator award depending on a number of subscribers – an indicator the tech giant was aware of Marcos's following and even considered him a partner in profit.[42]

Marcos and his family members were also guests on the channels of mega-influencers, including sympathetic celebrities and hyper-partisan vloggers notorious for spreading disinformation.[43] Meanwhile, micro-influencers also cashed in on the popularity, using pro-Marcos content to rack up views and

[39] Bunquin, J. B. & Gaw, F. (2022). Digital Public Pulse: Mapping the digital landscape of Philippine politics. Philippine Communication Society Review, p52.

[40] Baizas, Gaby. *Facebook Group Sheds Philippine History Brand, Now Spreads Marcos Propaganda*. Rappler, 20 Jan. 2022, www.rappler.com/nation/elections/facebook-group-sheds-philippine-history-brand-spreads-marcos-family-propaganda/.

[41] Digital Public Pulse: 2022 Philippine General Elections. Philippine Media Monitoring Laboratory, College of Mass Communication, University of the Philippines. pp, 118-119, p319.

[42] https://www.youtube.com/watch?app=desktop&v=pgWqRvO3TwA

[43] https://www.youtube.com/watch?v=1EwMAiqLUhM

subscribers.[44] Some turned a profit by monetizing the YouTube Partnership Program.[45]

TikTok

The Chinese short video platform had both "good and bad" information, according to a 2022 study supported by Internews.[46] But only 38% of respondents said they saw content from reliable media personalities and organisations, as the app was still primarily used for leisure and entertainment. In the same study, the most widely reported piece of disinformation encountered on the platform – encountered by a third of respondents – was a myth from the Marcos lie machine: that the Philippines was the "richest country in Asia" during the Marcos regime.[47] Around 7 to 13% of misleading claims fact-checked by Tsek PH originated from TikTok.[48]

Micro-influencers from various walks of life endorsed Marcos, shared vlogs from his rallies, and engaged in social media challenges that supported him. Over half of the videos coded in the Internews study were "politically charged". But, as with YouTube, provocative content did not come from Marcos himself. His early TikToks did not mention politics or his father, the late dictator, save for passing mentions, said Jacque Manabat, then a reporter for broadcaster ABS-CBN who followed Marcos' campaign and analysed his TikTok account for her master's thesis. Instead, Marcos focused on lifestyle content. "He behaved like […] a celebrity influencer, not a politician," said Manabat. "If your target includes people born long after martial law, you have a clean slate of followers."

When I monitored pro-Marcos content on the platform ahead of the election, there was also an abundance of anonymous accounts dedicated to posting fan edits. Some used archival photos to wax nostalgic and glorify various members of the Marcos dynasty. Others spliced content from Marcos' official YouTube channel and distributed shorter clips on TikTok. Others fawned over Marcos Jr.'s sons, one of whom was running for Congress. Incendiary content was typically left to hyper-partisan influencers.

Common elements and tactics in messaging

Across all platforms, there were some common trends seen in pro-Marcos messaging:

[44] Soriano, Cheryll Ruth R, and Fatima Gaw. "Platforms, Alternative Influence, and Networked Political Brokerage on YouTube." *Convergence: The International Journal of Research into New Media Technologies*, vol. 28, no. 3, 28 July 2021, pp. 781–803, https://doi.org/10.1177/13548565211029769.
[45] Elemia, Camille. (See footnote 18.)
[46] https://internews.org/wp-content/uploads/2021/12/Internews_Understudied-Digital-Platforms-PH_December_2021.pdf
[47] Lanuza, J.M.H., Fallorina, R. Cabbuag, S., *Understudied Digital Platforms in the Philippines*. Internews, 2021. https://internews.org/wp-content/uploads/2021/12/Internews_Understudied-Digital-Platforms-PH_December_2021.pdf
[48] https://www.tsek.ph/firehose-of-disinformation-floods-run-up-to-election/

- **Historical distortion** On top of lies shared in the form of trivia, archival photos and footage were spun and edited into revisionist fancams that spread across TikTok.

- **Mythologising** Modern Marcos lies are often steeped in folklore and even mythology, a practice that dates back to state-commissioned cultural work and propaganda under the dictatorship. The most prominent example is the Tallano gold conspiracy.

- **Toxic positivity** Content on the official channels of Marcos shows family-friendly, feel-good content, glossing over the various controversies the family faces – from a record of corruption to tax evasion. Phrases like "respect my opinion" were weaponized to put an end to discussion, bringing conversations to a deadlock. Sociologist Nicole Curato wrote Marcos' rallying cry for unity "reduces critical arguments to the perpetuation of bad vibes, and ignores unease because cynicism cripples hope." [49, 50]

- **Harassment and black operations** While Marcos has preached a message of unity, his flock of hyper-partisan influencers, bots and trolls, and even organic followers took on the role of attack dog: verbally abusing and harassing his critics, opposition figures and supporters, and members of the press.

One need not look too far beyond the family tree for proof of a direct smear campaign. The president's sister, senator Imee Marcos, acted in a series of incendiary videos directed by propagandist Darryl Yap and depicting a veiled caricature of Robredo as an obnoxious and stupid puppet. Marcos told broadcaster ANC that the series was "my way of answering issues." The senator's participation suggests the series may have been commissioned.[51, 52]

On the last leg of the campaign, a fake link implicating one of Robredo's daughters in a non-existent sex scandal emerged. The source of this link was not determinable, but pro-Marcos accounts played up the lie and echoed an idea that the Robredo camp manufactured it themselves.[53]

---

[49] Curato, Nicole. The Philippines: Erasing History through Good Vibes and Toxic Positivity https://th.boell.org/en/2022/03/28/philippines-good-vibes-toxic-positivity

[50] https://opinion.inquirer.net/150205/respect-my-opinion

[51] https://www.rappler.com/nation/filipino-workers-online-share-stories-sweat-labor-after-marcos-insult/

[52] https://news.abs-cbn.com/spotlight/06/30/22/the-destabilizing-charisma-of-imee-marcos

[53] See footnote 51

How Filipino journalism is responding

Filipinos are generally aware of the threat of disinformation: nine out of 10 say "fake news" is a problem, according to a [2022 Pulse Asia poll](#).[54] The bad news is that the same survey found almost half think journalists are to blame.

Lowered trust in the press increased during the pandemic and settled at 38% this year, according to the [RISJ Digital News Report](#).[55] Relatively lower trust in brands targeted by coordinated harassment – such as ABS-CBN and Rappler – suggest the possible impact of malign influence campaigns in public opinion.

Apart from disinformation, influence campaigning, and falling trust, journalists must also contend with the effects of a business model under pressure, increased news avoidance, and labour and welfare issues.

In the [*Oxygen of Amplification*](#), Data & Society argues that internal newsroom issues such as churnalism, overwork, and quotas for deliverables have a direct link on the susceptibility and vulnerability of media institutions to disinformation amplification and the spread of ideas that otherwise would have stayed in the dark.[56] In a newsroom where journalists must put out several stories a day, who is left to check the facts?

Conflict of interest and corruption within the media industry are additional factors that fracture a credible and united press. "Envelope journalism" is a prevalent practice in a culture of corruption – and with low pay and precarity in the industry, some would view integrity as a luxury.

Not even six months after the election of Marcos Jr, [a handful of reporters](#) from leading newsrooms – assigned to cover him on the campaign trail – quit to join Marcos's communications team or other appointed government positions.[57] While a career shift from journalism to political communications is not uncommon, it is unclear whether newsrooms closely examine, sanction, or even actively condemn potential conflicts of interest that could call into question the impartiality of reporting made over the campaign.

Despite the stress and the growing fractures, many local journalists have maintained a defensive stance. Here are some actions they have taken over the past two years.

---

[54] https://pulseasia.ph/updates/september-2022-nationwide-survey-on-fake-news/
[55] https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2023/philippines
[56] Phillips, Whitney. *The Oxygen of Amplification*. Data & Society, 2018. https://datasociety.net/library/oxygen-of-amplification/
[57] https://news.abs-cbn.com/news/06/29/22/andanar-tells-successor-to-protect-the-press-promote-foi

Fact-checking initiatives

The Philippines is home to several fact-checking initiatives, including four that are accredited by the International Fact Checking Network: Vera Files, Rappler, MindaNews, and PressOne. Tsek PH is a collaborative, cross-sectoral fact-checking network that coalesces around the period of elections.

One of the ongoing debates in both the local and global media industry is the effectiveness of fact-checking operations. Some critics argue that these checks are reactive and don't reach the audiences most affected by fake news. However, fact-checks also serve important functions: providing the literature and references needed for quick debunking, and daily monitoring and data gathering that eventually helps point toward larger trends.

Dedicated disinformation and influence reporting

When I began research for this project, Rappler was the only Filipino outlet I spoke to that had a dedicated team working to produce regular longform stories on disinformation. In most newsrooms, disinformation issues are still usually assigned to general affairs, or to reporters who cover other related beats.

Rappler's data team is composed of a four-person unit, coordinated by reporter Rambo Talabong until he left Rappler in March 2023. Their stories include an expose of a top pro-Marcos YouTuber that resulted in the takedown of over 150 videos, and analysis of online sympathy for a micro-influencer who assassinated a mayor in a graduation ceremony.

Pumapodcast's popular production, *Catch Me If You Can*, hosted by Kat Ventura and Jonathan Ong, explains disinformation and influence operation machinery, and features guests who are self-confessed influence operators. It is among the company's top-performing podcasts, clearly meeting an appetite and curiosity about the shadow industry from listeners both in the Philippines and abroad.

The podcast uses English—with some casual Filipino—as a primary language, and Pumapodcast's following primarily caters to the ABC social class demographic. Given this reach, there is still room for similar content and programmes that could appeal or target other audiences.

In post-election efforts, GMA News' project InoculatED focuses on pre-empting disinformation, or prebunking, with a concentration on health, climate, and even entertainment.[58] Because the network did not have a dedicated team looking at data and disinformation prior to this, the project's managers sought funding from Google News Initiative for its first round of operations and hopes it will be successful enough for newsroom management to sustain support for later.

---

[58] https://www.gmanetwork.com/corporate/articles/2022-12-12/764/gma-networks-digital-video-lab-wins-google-news-initiative-challenge-for-inoculated/

International nonprofit Internews has since also launched a collaborative rapid response network to support reporting on foreign influence operations in the Philippines. Among the newsrooms that have joined this effort include online news sites Mindanews, PressOne.PH, and PhilStar.com, and production company Probe Productions Inc. The network held a training session in December, and shares and syndicates each other's content. It has so far produced reports on pro-Duterte accounts' push for the former president to run for senator, Mindanews' expose of attempts to reproduce its site and content, and a video and text explainer of influence operations.

Statements of support

In the past, statements of support from newsrooms and support networks have helped in providing solidarity for the target of malicious campaigns, while also signalling institutional and organisational intolerance for the harassment of journalists. The National Union of Journalists of the Philippines routinely releases such statements in support of its members when faced with harassment, red-tagging, and other threats.

After the verbal assault of reporter Raissa Robles by senatorial candidate Larry Gadon, the Foreign Correspondents Association of the Philippines issued a statement calling for Gadon to be disciplined for his actions.[59] In less than a month, the Supreme Court issued a preventive suspension, saying "the call of the public did not fall on deaf ears."[60] In June 2023, it unanimously disbarred Gadon.[61]

Expansion to new platforms

Some journalists have struck out independently, and sometimes have found more freedom on new platforms than in old newsrooms. Broadcast journalist Christian Esguerra, formerly of ABS-CBN, launched his YouTube show Facts First, which has since grown a following of over 190,000 subscribers.

Talabong, previously of Rappler, has also launched his own YouTube channel. He and other reporters such as Jacque Manabat cultivate active presences on TikTok.

---

[59] Foreign Correspondents Association of the Philippines, 2021. https://twitter.com/FOCAP2020/status/1471722250676355073
[60] https://sc.judiciary.gov.ph/sc-suspends-gadon-and-orders-him-to-show-cause-why-he-should-not-be-disbarred-over-viral-video-against-journalist/
[61] https://sc.judiciary.gov.ph/court-unanimously-disbars-atty-lorenzo-larry-gadon-for-misogynistic-sexist-abusive-and-repeated-intemperate-language/

Alternative mechanisms for monitoring harassment

In the absence of human resource or security initiatives and digital scans, newsrooms in the Philippines are not likely to keep an active record of the influence operation-related violence they face. Most Filipino journalists, with good reason, simply ignore the vitriol they receive or are tagged in.

However, documentation of this phenomenon is important, especially when assessing security threats for future coverage. It is also important in establishing patterns that might be helpful for a legal case in the future. Some workplaces have developed alternative mechanisms for monitoring this.

At Rappler, the task of monitoring online attacks against reporters is assigned to the data forensics team. This is because "it's also a story that can be told... there's really a larger pattern," said Don Kevin Hapal, who heads Rappler's data cluster. The team noticed that sometimes, harassment eased after they published reports calling out and situating these online attacks in coordinated campaigns. This record-keeping is especially useful for possible legal action.

At the *Philippine Daily Inquirer*, the employees' union conducts this documentation, then decides whether to release a statement or what other actions to push for. Reporter Krixia Subingsubing noted that the presence of this alternative should not be an excuse for management to forgo hiring an in-house security specialist.

The impact on journalists

It is no secret that those who reported critically about Duterte and Marcos were magnets for often-coordinated malicious influence operations. The most famous case is that of Rappler CEO, Maria Ressa. Analysis by the [International Center for Journalists](#) showed that 60% of the attacks against her were designed to undermine her credibility, and the other 40% were personal and designed to – in her words – "[tear down my spirit](#)".[62,63]

Similar attacks are echoed in experiences of other journalists. This chapter documents lesser-known instances of journalists targeted by harassment campaigns, and other intimidation tactics in relation to the 2022 election.

Online abuse and violence

The most common tools of influence operations against journalists is [astroturfing](#), [brigading](#), trolling, and online attacks.[64,65] The barrage of hate typically comes after an influencer puts the journalist on blast, often providing talking points that are echoed by smaller accounts through sending harassing personal messages and flooding comment sections. Some journalists have decreased their public presence on social media as a result of these attacks.

One tactic of malign influence operations is to target journalists in real time through brigading comment sections of livestreams, setting an impression of public opinion while an event is still ongoing. "Sometimes the trolls would be looking for me even before I asked any question, to pre-emptively belittle me even before I spoke, setting the narrative for other viewers so to speak," said Pia Ranada, a Rappler reporter previously assigned to cover then-president Duterte.

Other times, the attacks came retroactively, such as when hyper-partisan accounts splice reports or questions posed by journalists out of context. The content is then edited with captions and text publicly shaming the journalists or framing them as being biased, even when the questions they ask are fair, standard, and even benign.

Among the biggest targets of these attacks during the campaign trail was Rappler reporter Lian Buan, who was assigned to shadow Marcos on the campaign trail.

Another high-profile incident was the verbal assault of South China Morning Post reporter and *Marcos Martial Law: Never Again* author Raissa Robles by senatorial candidate Larry Gadon, who was running under Marcos' slate. The video [went viral](#)

---

[62] https://www.icfj.org/our-work/maria-ressa-big-data-analysis

[63] Maria Ressa, "We're All Being Manipulated the Same Way," *The Atlantic*, April 2022. https://www.theatlantic.com/ideas/archive/2022/04/maria-ressa-disinformation-manipulation/629483/

[64] https://www.theguardian.com/commentisfree/2012/feb/08/what-is-astroturfing

[65] https://newsliteracy.psu.edu/glossary/brigading

on social media.[66] Despite his eventual disbarment due to conduct, Gadon was appointed a presidential adviser for poverty alleviation.[67]

In June 2022, after the election, senator Imee Marcos teased ABS-CBN news anchor Karen Davila over her brother's presidential win during the latter's morning show. "It's nice to see you here in the Philippines. I thought you would migrate if a Marcos won," she said in Filipino, although it is unclear that Davila made any such remark. Davila replied light-heartedly: "Always hoping for the best for the country." Although Davila later posted they were "all good" and Marcos apologised via a text message, the online machinery was quick to harass her, splicing and spinning the video and framing her as humiliated by the senator. False news about the anchor being asked to leave the country circulated online. Meanwhile, Marcos doubled down on opposition supporters by posting a sarcastic video about where Filipinos could migrate to.

Personal harassment, assault, and intimidation

In some instances, the effect of influence operations crossed the line from online to offline, lapsing into personal harassment and physical intimidation.

On the campaign trail, reporters covering Marcos were pressured into producing good press and isolated if they did otherwise.[68] The tension bubbled over when a handful of women reporters, including Rappler's Lian Buan and TV5's Marianne Enriquez, were shoved and roughhoused by Marcos's security.[69] Later, after the election, Buan was deliberately snubbed by Marcos's spokesman after posing a question at a formal press conference.

Jacque Manabat and her team from ABS-CBN were surrounded by angry supporters while in two different vehicles during a pro-Marcos rally, she told me in an interview. Some waved cameras in front of the team, trying to provoke a reaction. When one of the staff lost his cool, video of the incident went viral, prompting hyper-partisan media outlet SMNI to target ABS-CBN in its report. "Journalists should not be experiencing this," said Manabat. "We're there to cover, we're there to observe."

BBC correspondent Howard Johnson was also widely harassed after doorstepping Marcos Jr. on the campaign trail. One pro-Marcos British blogger, who previously co-organised and livestreamed a 2017 "protest" against the BBC's coverage of the Philippines at the network's headquarters in London, started a Change.org petition

---

[66] https://www.philstar.com/headlines/2022/02/08/2159368/suspended-lawyer-gadon-faces-criminal-raps-over-vulgar-video-vs-journalist

[67] https://pco.gov.ph/news_releases/pbbm-appoints-atty-larry-gadon-as-presidential-adviser-for-poverty-alleviation/

[68] https://www.rappler.com/newsbreak/in-depth/marcos-jr-campaign-where-journalists-blocked-boxed-out/

[69] https://www.rappler.com/nation/elections/rappler-to-marcos-camp-stop-harassing-journalists/

calling for the network to apologize for Johnson's interview. The petition drew almost 7,000 signatures.

Red-tagging

Red-tagging, the practice of labelling individuals or groups as communists or terrorists, has often led to both online and offline harassment in the Philippines. In recent years, red-tagging has become more dangerous as it often precedes arrests and extrajudicial killings of activists, journalists, and other members of civil society.

Although the practice has been in place for years – used often by law enforcement and the military against activists – the violence and fatalities associated with it significantly ramped up under former president Rodrigo Duterte's term. It was often used as a tactic to discredit journalists before and after the 2022 election.[70]

Attention hacking and hijacking safe spaces

Filipino podcast start-up Pumapodcast held Twitter Spaces to discuss the elections and other pressing political issues with their followers. Hyper-partisan users, typically supporting Duterte and Marcos, began joining these sessions.

At first, organisers sought to hear out dissenting views. While some interactions were decent, they eventually had to revoke speaking access from personalities who came in to curse others and pick a fight. "Our listeners come to us because they trust us to create safe spaces," said CEO Carljoe Javier. "If people are coming in with the wrong mindset, you're going to make the place unsafe."

Eventually, they took feedback through comments to better moderate the discussion. Pumapodcast also began to keep a record of repeat hecklers, to blacklist them in future online events. After stricter moderation rules were observed, those actors would hold a competing Twitter Space, either simultaneously or directly after Pumapodcast's, to discredit and malign the earlier conversation.

Threats to digital security

In March 2022, freelance journalist Anthony Esguerra's Facebook account was hacked, posting pro-Marcos propaganda and child pornography. Given the nature of the content posted, Esguerra suspects the hacker may have been trying to get his account, which was verified with a blue check, suspended. This attack could also be viewed as another method aimed at targeting a journalist's credibility.

"The first thing I experienced was paranoia [...] to the point I couldn't sleep," he said, adding he was alerted to a strange log-in from Vietnam.

At the time, Esguerra was running an online production called Factual PH. To date, he has been unable to boost journalistic content there, likely due to a record of

---

[70] https://cpj.org/2022/05/red-tagging-of-journalists-looms-over-philippine-elections/

violating Facebook community guidelines. He has also noticed strange accounts have access to his ad management features. Esguerra has been unable to reach the tech giant directly.

He said the event had dented his morale, as he was blamed for his own hacking online. He eventually took down a post chronicling the hackers' posts on his account, as the screenshots were taken out of context and weaponized against him.

Vulnerability of freelancers

Freelancers like Esguerra are especially vulnerable to online attacks by influence campaigns because there is often no supporting body to offer resources and support.

In another case, photojournalist Martin San Diego took a photo of then-presidential candidate Marcos Jr. and film director Paul Soriano during the campaign, while on assignment for Rappler. Soriano lifted the photo, cropping the watermark, and shared it on his Instagram account with an endorsement of Marcos. When San Diego raised concern regarding his work's copyright and credit – first through a comment on Soriano's post and then in his own social media profiles – he was met with a wave of harassment. He estimates around a hundred comments poured in on his personal Instagram, a hundred more on his Facebook account, and even more on a news article sharing Soriano's post. At the height of the harassment, he received "five to 10" comments or notifications an hour.

San Diego regularly contributes to both foreign and local press. As he was on assignment for a local outlet at the time, it was not among the cases mentioned when the Foreign Correspondents Association of the Philippines released a statement of concern regarding harassment toward journalists.

In the end, Soriano deleted his post without an apology. He has since been appointed Marcos's "adviser on creative communications".

Recommendations

This section compiles expert advice, best practices from various newsrooms, and resources to create a comprehensive guide for covering disinformation and influence operations in the Philippine context.

My recommendations cover both editorial or content and organisational interventions.

- First, I offer tips for crafting a newsroom-wide disinformation and influence reporting strategy.

- Second, I suggest editorial approaches that can strike a balance between informing the public and not amplifying manipulative actors.

- Next, I compiled some experiences and best practices about how to sensitively approach and report about hyper-partisan characters.

- Lastly, I argue for the instatement of a standard protocol should journalists come on the receiving end of malign influence campaigns.

1. Define your malign influence reporting strategy

"We've all got the same aunt who sends us the thing […] [saying], 'Is it right?'" said Rebecca Skippage, who co-pioneered and leads BBC's disinformation unit. "We all have the same universal experience of disinformation and being inadvertently pulled into it."

As disinformation and influence permeates all aspects of life – politics, finance, health — Filipino newsrooms could benefit from pivoting away from covering the issue as breaking news, and toward more analytical and investigative pieces, explaining larger trends to their audiences, and targeting malign networks as opposed to playing whack-a-mole with individual issues.

"If you're a responsible news organisation, why are you not investing in this?" Skippage added.

Critical analysis should be incorporated into overall news production, even in light features such as entertainment items. While not every piece of journalism has to be investigative in depth, every piece should have thorough fact-checking, be aware of and avoid playing into narratives possibly planted by malign influence operations, and adhere to the principle of "doing no harm".

On the next page, I outline some tips for how to get started.

Newsrooms need a resident expert – preferably more than one

Most Filipino newsrooms use a beat system, and disinformation-related news is still largely delegated to general affairs or the beat most closely related to the topic. Proposed fake news laws are typically covered by Congress or Senate reporters, health-related misinformation goes to health reporters, and so on. While fact-checking should be a muscle for reporters covering all beats, experts still advise having a dedicated reporter or point person – or even a whole team – who specialise on this topic in-depth.

This expert should ideally be able to conduct basic digital forensic scans, navigate the information landscape on various platforms, and identify bigger and evolving disinformation and influence patterns and trends. This almost certainly means taking at least one person out of the daily grind to concentrate on deep reporting.

Some global newsrooms have dedicated entire teams to disinformation and influence reporting. Peer support is helpful, as the workload and vulnerability that comes with uncovering sensitive and volatile topics are shared. A team format also enriches reporting with more rigorous cross-checking and information-sharing, especially when team members have different skills or specialties that cut across country or area expertise (covering different languages, sociocultural and political contexts), and methods (such as fact-checking, open-source intelligence, and mastery of tools).

Researcher Fatima Gaw told me there are still plenty of "blind spots" on the subject of disinformation, and newsrooms have to invest in investigative journalism and analysis. "It's a matter of resources and infrastructure. An individual journalist can't do that without a data team," said Gaw. "The effort to report disinformation should be institutionalised and not on the journalist [alone]." As stated on page 19: Rappler is the only Filipino outlet with a dedicated team as described above, while GMA News' InoculatED and the Internews collaborative initiative show early promise.

Put good information back on the same platform as bad information

When countering or exposing disinformation and influence operations, always return to the same platforms where they circulate, and in the same language, Skippage advised.

Citing Rappler data, Hapal said that – in the Philippine context – fact-checks rarely penetrate the echo-chambers where they were initially shared, and often don't reach the same audience.

Invest in training and upskilling journalists

Covering disinformation and influence requires a combination of hard and soft skills, on top of basic reporting. A reporter in this field should have a grasp of how to scrape information and use analytical tools on various

platforms. "You have to be a data analyst, a social media monitor," said Skippage. "You have to be a really good storyteller… be able to find case studies, go across multiple platforms."

The journalist must also be capable of analysing online trends and putting them in context, whether in reference to influence operation tactics in the past or elsewhere in the world. They should also be able to navigate reporting in a nuanced and sensitive way, without playing into narratives that might benefit malign influence operators or oversimplify complicated issues.

Outsource and partner with academics, data experts, and other disciplines across civil society

In the event of a lack of in-house data experts, media outlets can work together with various academics and experts who may be able to help in sourcing and making sense of both quantitative and qualitative information. The Internews network on foreign influence operations has partnered with DoubleThink Lab, a digital intelligence organization based in Taiwan.

Journalists also want more solidarity and even cross-newsroom collaboration across the industry. This would allow reporters to work with others whose skill and resources complement their own, as well as expand reach across various audiences. For example, ABS-CBN's Manabat worked with Manila Bulletin tech editor Art Samaniego to produce an expose on Facebook pages and groups being sold.

BBC's disinformation unit has also collaborated with Bellingcat. Disinformation is a "team game," said Skippage. "People want to help each other out because it's really rare you see the whole picture. You need to work together to crack the nut."

Collaborations need not only be editorial in nature, but could also be strategic – timing publication and sharing resources for maximum reach, for example.

Non-journalistic anti-influence efforts to collaborate with

Here are a few ongoing media-adjacent efforts and initiatives that are open for collaboration and coordination with the press.

- Filipino behavioural science start-up, AHA, is looking into what actions can compel people to unfollow misinforming actors or leave closed groups with hyper-partisan content. An initial study found that participants who received educational tools had the potential to block more fake news sites.

- The Movement Against Disinformation is a civil sector initiative that has been assisting in calling for more actions from tech platforms, as well as lending journalists legal support in cases that could test the justice system.

- Out of the Box is a media literacy initiative and nonprofit that has reached hundreds of schools and thousands of students and teachers.

- [Martial Law Museum](#) is an online learning resource providing teaching modules for basic education about the Marcos dictatorship.
- [Bantayog ng mga Bayani](#), a memorial that contains an archive of martial law victims and survivors whose stories have been sidelined by pro-Marcos propaganda.
- [Wiki Society of the Philippines](#), a volunteer group of Wikipedia editors who have been on the frontlines of correcting disinformation.

2. Take a structural editorial approach

Influence operations and disinformation must be examined as a structural issue, and may require a reorientation toward some established newsroom cultures and habits. Not every story is worth breaking, and some stories require deeper dives.

In a 2023 [report](#), Filipino academics called on journalists to shed light on "the insidiousness of covert influence operations" and "support researchers in surfacing and materializing their political economic workings through investigative journalism." They urged the press "to follow and sustain leads on influence operations beyond elections as part of their political beats, focusing not only on national-level propaganda campaigns but also their counterparts in local politics, especially on anti-media influence campaigns."[71]

Here are some ways to begin:

Weigh newsworthiness

Every step of the editorial process entails decision-making and risk assessment. When covering a lie or influence campaign, journalists and their editors must weigh out whether the stories are worth covering at all.

> *The Oxygen of Amplification* **offers three criteria for establishing newsworthiness:** [72]
>
> - **Tipping point** Has the story extended beyond the interests of the community being discussed?
> - **Social benefit** Will the story have a positive social benefit, open up a new conversation, or add weight … to an existing conversation?
> - **Potential harms** Will the story produce harm (embarrassment, re-traumatization, professional damage), or could an audience use the story to cause harm (attacking sources, imitating crimes)?

Among many factors, reporting about harmful information might increase the likelihood of harassment, and make bad actors seem "more visible, more influential" than they really are, researcher [Whitney Philips](#) wrote.

---

[71] "Political Economy of Covert Influence Operations in the 2022 Philippine Elections," 2023.
[72] Phillips, Whitney. *The Oxygen of Amplification*. Data & Society, 2018.
https://datasociety.net/library/oxygen-of-amplification/

Conversely, not reporting about it could "allow poisonous ideology to flourish, and [cede] cultural territory to […] manipulators." [73]

Another method of measuring the impact of an influence campaign is using the Brooking's "breakout scale," a comparative model that evaluates how much of a threat is present.[74] The scale has six categories and considers whether an influence operation has moved from a small reach on a small platform, to outside initial target communities across various platforms, with a possible incitement for violence. Considering these various degrees of risk could help in editorial decision-making as to whether or not an issue should be covered.

Tell stories about the human cost

One of the things that elevates disinformation journalism, said Jack Goodman, a senior journalist at the BBC disinformation unit, is to showcase human interest – the harm caused to people affected directly. This could include fractures in polarised communities and families, and the offline impact of online abuse.

In the Philippines, where a fifth of the population lives below the poverty line, connection of disinformation to "gut issues" like transport and the prices of goods are important. Migrant workers are particularly vulnerable to influence operations in part due to targeted marketing and organising. Reportage on the intersection of influence and relatable socio-economic issues, like finance scams, might also be valued among Filipino audiences. In 2017 for example, thousands gathered at a university south of Manila to claim rumoured dole outs from the Marcos gold.[75, 76]

Although the Marcoses have distanced themselves from these operations, Duterte himself has lent credence to the gold conspiracy theory, and related scams have proliferated.[77]

Tell stories about platform accountability

Another angle that uplifts disinformation journalism, Goodman said, is platform accountability. How do algorithms and big tech amplify malign influence, or overlook it?

---

[73] Phillips, Whitney. *The Oxygen of Amplification*.
[74] https://www.brookings.edu/articles/the-breakout-scale-measuring-the-impact-of-influence-operations/
[75] https://newsinfo.inquirer.net/932919/marcos-ill-gotten-wealth-ferdinand-marcos-jr-bongbong-marcos-uplb-gathering-one-social-family-credit-cooperative-marcos-gold-bullion-buyer-ltd-ferdinand-marcos-manuel-cereno-jr-uplb-studen
[76] Ariate, Joel F, et al. Marcos Lies. University of the Philippines Third World Studies Center, 2022, pp. 235–239.
[77] https://apnews.com/general-news-0f756f80439c4f8ea65b8b9311be0d25

Reports, for example, about now-president Marcos receiving a gold star from YouTube under its partnership program must also contain the context that the platform has allowed for disinformation to thrive.

Tell stories about institutional accountability

Sociologist Nicole Curato notes that the appeal of political myths like those surrounding the Marcos family are related to a disillusionment with failing institutions and systems.

> *"A society with credible institutions would allow ordinary citizens to interrogate unexplained wealth. Seeing corrupt public officials placed behind bars provides closure to the story of accountability. But this is not the Philippine society. We turn to myths, urban legends, and rumours to make sense of suspicious situations because we cannot count on our institutions to make sense of this for us and give closure to the story."*[78]

Reportage on influence operations should be complemented by ongoing efforts to seek the accountability of institutions – both from the various branches of government and businesses.

Filipino academics, including Curato, have called for more media coverage that focuses on accountability and unmasking disinformation architects.[79] This is easier said than done, especially in the Philippines – but Filipino journalists can look to collaborative examples abroad for models, such as "The Story Killers" project, which gathered skills from 30 news organisations around the world, including *The Guardian*'s undercover investigation into an Israeli organisation that claimed to have manipulated more than 30 presidential elections around the world.[80]

Cover proposed solutions critically

Reportage should include critical analysis of possible policy interventions and solutions to the infodemic.

Oxford professor Howard, in Lie Machines, says such influence operations "have flourished on social media such that we are past the point of industry self-regulation. Ultimately the solution must involve public policy oversight and some market-shaping initiatives."[81]

---

[78] Nicole, Curato. "The Golden Age of Disinformation." *LSE Southeast Asia Blog*, London School of Economics, 18 May 2023, blogs.lse.ac.uk/seac/2023/05/18/the-golden-age-of-disinformation/
[79] Parallel Public Spheres, page 20.
[80] https://www.theguardian.com/world/2023/feb/15/revealed-disinformation-team-jorge-claim-meddling-elections-tal-hanan
[81] Howard, Lie Machines.

Filipino academics take a more cautious approach by urging bodies like the national advertising council to self-regulate. In a region where fake news laws can be weaponized against freedom of speech, local academics have also warned that proposed public policies or legislation can be lopsided and abused, especially as lawmakers most benefit from the disinformation and influence ecosystem.[82]

Academics also cautioned journalists against pinning blame singularly on platforms and portraying the public as "brainwashed" victims with no agency.[83] While it is also important to keep tabs on international developments, policies, and interventions on tech platforms, local experts advise focusing on bridging gaps and depolarising communities.[84]

Here, it is useful for journalists to remember that their role is not in the creation of policy, but in providing the information and public accountability that will:

- make the case for policy intervention clear,

- provide information that informs policymakers,

- critically assess who the policy has been drafted by and seek expert comment on potential unintended consequences of the policy.

It also helps to be forward-looking, so coverage of civil society efforts – like the impact of legal cases filed by the Movement Against Disinformation in support of journalists and the jurisprudence it hopes to create, and the work done by content creators who are allies of media literacy and press freedom – receives recognition.

Be more selective about who you choose to amplify (and how)

In a world where clicks drive commercial success, it's understandable that many news outlets choose to amplify whatever is the most fun or infuriating. But academics at Harvard Kennedy School and Microsoft Research argue that the press should move toward "strategic amplification," recognising that "amplifying information is never neutral and those who amplify information must recognize the costs and consequences of publication."[85]

Philippine news outlets should also wean itself from overreliance on "quote cards," a popular visual tool where a quote of the day is slapped against a photo of the prominent figure who said it. The graphic format is a magnet for

---

[82] Gaw, F., Bunquin, J.B., Cabbuag, S., Lanuza, J.M., Sapalo, N., & Yusoph, A. (2023). Political Economy of Covert Influence Operations in the 2022 Philippine Elections, page 89.
[83] Parallel Public Spheres, page 12.
[84] https://www.gmanetwork.com/news/opinion/content/861875/collaborators-invited-for-community-engagement-fund-vs-disinformation/story/
[85] Donovan, J., & boyd, danah. (2021). Stop the Presses? Moving From Strategic Silence to Strategic Amplification in a Networked Media Ecosystem. American Behavioral Scientist, 65(2), 333-350. https://doi-org.ezproxy-prd.bodleian.ox.ac.uk/10.1177/0002764219878229

social media engagement, as Facebook compels users to react to what is said; but it is also often abused as quotes can be spun out of context or sensationalised. Furthermore, the format has also been hijacked by disinformers to attribute false quotes to their political targets.

Filipino outlets should also carefully weigh inviting serial disinformers and bad faith actors to programmes where they can forward their agenda unchecked, especially in live interviews and livestreams, formats that do not undergo rigorous post-production fact-checking.

Case study: A narrative made in Malacañang

Movie director Daryll Yap runs an incendiary channel, previously slapped with a community guideline violation notification by Facebook, that posted a black propaganda video series against opposition leader Leni Robredo, including having a puppet with her likeness tied to a bed and exorcised.

His content regularly contained indicators of covert political campaigning – from the casting of the president's sister, senator Imee Marcos, in the videos against Robredo, to political advertisements for senatorial candidate Jinggoy Estrada being stitched into supposedly non-partisan content.

Yet, when his propaganda film *Maid in Malacañang*, a fictionalised account of the dictatorship, was released it garnered a lot of press in the Philippines.

Local news company News5 reposted screenshots of the video from the senator's page, racking up 25,000 engagements and 15,000 shares. Other outlets published quote cards depicting a war of words between an actress who likened history to gossip (a quote from an interview on Imee Marcos' vlog) and outraged historians responding to her.

All of this buzz served as unwitting promotion for the propaganda movie. It also made it easy for pro-Marcos voices to cast the actress as a victim of mob mentality and cancel culture.

In hindsight, more responsible coverage of this event would have investigated the role of Senator Imee Marcos in its production, speculation that moviegoers were given free tickets to drive traffic to the box office, and human-interest stories spotlighting survivors of martial law with direct counter-narratives to the lies planted in the film.

Do not talk down to audiences and sources

Experts warn against framing voters in reports as people without agency, or pandering to a stereotype of a "*bobotante*" (a portmanteau that uses the Filipino words for "dumb voter").[86] At BBC, Skippage said "empathy" in reporting their stories is key: "We are really clear that we're not saying to

---

[86] Parallel Public Spheres, page 12.

people, you are dumb for believing this. If people are pulled down a rabbit hole […] plenty of people go down that route."

Bernice Sibucao, a senior manager and deputy head of the Digital Video Lab at GMA News, overseeing the production of the upcoming series InoculatED, pushed for a "more conversational type of journalism", and emphasised the importance of talking to (not down to) audiences. "At the end of the day, nobody likes to admit they were fooled," she said.

InoculatED aims to focus on less partisan topics – like health misinformation – to win a following and build trust among general audiences ahead of the next election, when polarisation kicks in.

When interviewing potentially misinformed characters, avoid both sides-ism, false equivalences, and stereotypes. Disinformers tend to bank on caricatures and bloat them, possibly to attack the opposing ideology or further discredit the press. The industry might benefit from showing more examples of regular people across various sectors who are affected by influence operations.

> **Some underreported angles and stories to explore in the Philippines**
> - Post-election coverage for single-issue policies
> - Explanatory pieces on how data collection, computational propaganda, and algorithms work
> - Accountability journalism involving disinformation architects and exploring self-regulation mechanisms in the advertising and public relations industry
> - Influence operations in politics at local government levels
> - Underreported trends on TikTok and YouTube

3. Mark out boundaries for dealing with influence actors

In the Philippines, a myriad of business models for influence operations abounds: from in-house operations to agency-run troll farms to independent influencers-for-hire. Different actors in the hierarchy of influence operations should be handled differently in reportage.

Approach critically and with caution

Before reaching out to manipulative actors such as online abusers and hyper-partisan personalities, consider first whether the story absolutely requires their voice.[87] What added value would this conversation give?

Second, consider the position of the source in the potential "lie machine" and the hierarchy of influence and information operations. Some nuances to consider: are they a state actor, or are they working privately? If they are

---

[87] Phillips, Whitney. *The Oxygen of Amplification*.

working privately, are they likely to be receiving incentives – financial or otherwise – from political clients, or are they genuine supporters who may be profiting via other means, such as social media business models? Do they have a history of misleading their viewers, and attacking the press?

Third, conduct thorough research. If they are a political influencer, who are the figures they have worked with before? How wide is their reach? What's their brand, their style?

These considerations can help assess the journalist's security risk, the possible motivations of the source, and the newsworthiness of the content. Personalities with a wide reach and a record of harassment or serial disinformation should also be treated differently, from say, organic and working-class micro-influencers.

For sources much further up the hierarchy of disinformation architecture – head honchos and political strategists directly coordinating with political clients – consider offering anonymity and a chance to speak off the record, advised strategist Alan German.

German advises journalists to "talk shop" and familiarise themselves with industry terminology: influencers are known as key online or opinion leaders, or KOLs. A line troll network, or LTN, is a team operation typically led by a moderator and, like a call centre, has a number of 'seats' for the heads involved in the operation; their online narrative is shaped by talking points plucked from a 'message tree'. Knowing these details can help in establishing rapport, as well as provide operational details.

However, academics have also urged more reportage that might actually hold strategists and advertising executives accountable. Approaches could vary depending on the story angle and developments surrounding it. If the piece is an accountability expose, architects and influencers may have to be named, possibly through a paper or online trail, whistleblowers, or other investigative techniques. If the planned report is more explanatory in nature – say, about new tactics in influence – leeway in anonymity is allowable.

Conversely, if sources with possible commercial interests want to be explicitly named or drop many hints about their services, consider their motivations: their media presence might be a dog whistle that they are open for business.

In my [own reporting](#) following the 2019 Philippine midterms, we were unable to name the operator and political client of the troll farm where we embedded due to promised confidentiality. However, through online scraping, we found less coordinated Twitter networks supporting three other senatorial candidates using similar, though slightly less sophisticated techniques. A reverse image search revealed that these troll accounts used photos of real people, including that of a Filipino victim of a killing in the United States.

The photo was recycled across fake accounts supporting each of the three candidates, a possible indicator that the politicians' networks were run by the same operator. Since finding this operator was near impossible based solely on the online trail, we reached out to the politicians' camps – all of whom denied they availed such services. As the report was explanatory in nature, unpacking how availing such services became common across the political aisle, we tried to balance both confidentiality and accountability by presenting different case studies.

Case study: Meeting Marcos micro-vloggers on the campaign trail

Filipino journalists Camille Elemia and Rambo Talabong both followed pro-Marcos video content creators for the *New York Times* and Rappler respectively. At the time, the rise of organic, mostly working-class vloggers cashing in on YouTube's Partnership Program and making a following by riding on the platform's dominantly pro-Marcos algorithm was new.

"I would just calmly respond to their rants about 'biased' media, about Rappler. One time I even allowed a vlogger to air our Q&A live," said Elemia, who previously headed Rappler's disinformation cluster before becoming an independent journalist. "The comments of the viewers were expectedly harsh. I said yes to that so they can see I am a 'legit' journalist and [I don't want to lose out]."

Rambo Talabong encountered similar attempts to put him on the spot. He said he maintained a friendly demeanour with vloggers, waving at their cameras, and not taking their approach or questions personally. Being male helped, he said.

"To them, I am content. Even demonising me was content. They think in the currency of content, as long as I can be of use to them [...] When the cameras were off for both sides, we understood each other. If the camera is on, it's like a tradeoff," said Talabong. "I think in their minds, it's like a collab [...] I did not take [how they portrayed me] against them. They were taking advantage of what would get more views."

"Others who featured me negatively, they stayed as sources. They tell me it's just... [their] public persona," said Talabong, who was later able to break stories about in-fighting in pro-Marcos groups as these content creators tried to organise for accreditation to cover the presidential palace.[88, 89]

This example demonstrates how engaging with hyper-partisan content creators also requires soft skills in social interaction, including compartmentalising and defusing conversations that could potentially escalate to hostility. Identity privileges like gender and social class are also a factor. Talabong's previous background as a labour reporter helped in his understanding of these sources, while Elemia watched their content and used this as a conversation opener.

---

[88] https://www.rappler.com/newsbreak/in-depth/reasons-why-marcos-jr-vloggers-fighting/
[89] https://www.rappler.com/newsbreak/in-depth/marcos-vloggers-group-disbands-months-into-ferdinand-marcos-jr-presidency/

Avoid outsourcing news from questionable actors

When politicians snub the traditional press in favour of hyper-partisan vloggers, legacy newsrooms should resist the temptation to parrot comments made on those channels to a broader audience – especially if these are designed to stoke an emotional public reaction. By publishing quote cards or leading news stories sourcing questionable influencers' content, newsrooms are legitimising, amplifying, and possibly driving traffic to their platforms. This not only accelerates the spread of the information, but also plays into the hands of the politician's interests as well as lines the pockets of the influencer with profit from increased views and engagement.

If content from a politically compromised influencer must be cited, it is best to apply the newsworthiness test. The issue must also be weighed on the scales of accountability as opposed to amplification.

Brace for a hostile reception

It's no secret that reporters focusing on disinformation are among the most vulnerable targets for harassment. Newsrooms should strategize and discuss safety measures that can minimise risks for those involved, such as rotating assignments and considering advantages and disadvantages based on factors like gender identity, ethnic or regional identity, language, and so on.

Expect to negotiate a tradeoff in these interactions. Chad de Guzman, a Filipino journalist at TIME, reached out to around a hundred micro- and nano-influencers and supporters for his piece about pro-Marcos propaganda on TikTok. Most refused to be interviewed.

Journalists should also anticipate a chance that their coverage will be spun to a creator's advantage.

Case study: Pumapodcast controversy

In May 2023, "Catch Me If You Can" hosted hyper-partisan social media personality Sass Sasot — a pro-Duterte and pro-Marcos influencer with over 880,000 followers — in two back-to-back episodes set in a dominantly interview format. The releases drew backlash, including from journalists and Pumapodcast followers, for platforming a serial disinformer.[90] This prompted the team to release a statement, and one of its hosts went on interviews with at least two major broadcasters and another online show to explain the decision to interview Sasot.[91]

Critics of the decision, including Rappler editors Inday Varona and Gemma Mendoza, were dissatisfied with how the interview was framed on the podcast, and believed Sasot should not have been afforded the space considering an unapologetic history of harassing the press.[92]

[90] https://www.rappler.com/nation/backlash-against-pumapodcast-hosting-disinformation-peddler/
[91] https://twitter.com/PumaPodcastPH/status/1657296163664703488
[92] John Nery, In the Public Square. https://youtu.be/iu17QzsNX-8

As the controversy gained traction online, Sasot alluded to the podcast on Facebook. Without naming the production, Sasot bragged about not even sharing it and proceeded to baselessly accuse Pumapodcast of being funded by the U.S. government. Accusations of foreign funding have also been levelled by Duterte and his supporters against other embattled local media companies, Rappler and Vera Files, and it has preceded coordinated harassment in the past.

Some hyper-partisan influencers have pushed for their interviews with journalists covering them for their own interests, to provide content and drive engagement to their channels. In these cases, journalists need to be prepared to be put on the spot or misrepresented. However, this is not to say that journalists must engage; journalists should not be obligated to respond if they feel unsafe.

Frame hyper-partisan actors in context

Hyper-partisan sources are allowed a right of reply, but it's also the journalist's job to question them and subsequently frame any lies and personal interests. Some relevant tips when dealing with manipulative sources from *Data & Society* include:

- Put their statements in historical and ideological context. Minimise the inclusion of "euphemistic dog whistles";
- Weave the performative nature of their actions into the story;
- Avoid deferring to their chosen language, explanations, or justifications.[93]

Case study: Marianna Spring

In BBC podcasts *Disaster Trolls* and *Marianna in Conspiracyland*, sources who are anti-vaxxers and conspiracy theorists are called such, even when they deny the label. Spring's reporting also distinguishes between influencers who are openly hostile, reformed conspiracy theorists, and followers who may be just going along for the ride and may be willing to change their minds.

Although sharing a human side to these characters helps in understanding the phenomenon, journalism must also not deodorise the actions of malign actors. Do not allow manipulative personalities to hijack or seize the narrative.

4. Pursue a standardised organisational protocol

Journalists in traditional Filipino newsrooms said they would benefit from a clearer protocol when they become targets of malign influence campaigns. Most journalists reported that while they received moral and verbal support from colleagues or management, instances of online abuse were treated on a case-to-case basis.

Often, journalists reported directly to their superiors and there was no clear designated security or human resources officer who specialised in the issue.

---

[93] https://datasociety.net/library/oxygen-of-amplification/

Philippine Daily Inquirer reporter Krixia Subingsubing said that while a clear protocol was in place in event of disasters like armed conflict and climate crisis coverage, that infrastructure is not yet present with disinformation and online violence.

Local newsrooms "still don't consider harassment a sort of disaster," said Subingsubing. She added this could be in part because of a perception that online violence might not reach the level of physical harm, whereas covering hazards and disasters often entails preparations for possible threats to physical safety, hospitalizations, and even legal concerns.

There is more awareness now that online violence should not be treated as less 'real,' and that it often precedes physical threats.

"It would be better to have a systematic response protocol if employees are harassed or threatened. It's not enough to have an impulsive, case-to-case basis response," said Subingsubing. "It would be better to have on paper a written system that a reporter should go through – these are your guarantees, your possible benefits: you can go on leave, you can avail of a counsellor."

Here are some best practices and hard rules that newsrooms can apply:

Track attacks against the company and employees
Appoint a person – possibly from human resources or security – to document and track the attacks against the company and its employees. This officer must be able to run risk assessments, or partner with an expert who can, to determine what safety precautions the targeted journalist must take (i.e. if their address is doxxed, they should move temporarily to a safehouse).

BBC Monitoring found that 80% of hate it received was targeted toward only its disinformation reporter Marianna Spring.[94] Similarly tracked data can also later prove useful in establishing patterns and finding common offenders. These can be used for accountability mechanisms, from legal evidence in building a case, or demanding takedowns of inauthentic behaviour and malicious pages from tech companies, or even reporting a future story.

Provide regular security assessments, tools, and training
On top of regular refreshers, pre-emptive measures to protect journalists' online accounts should also be taken before a story that might trigger attacks breaks. This could include changing passwords and installing two-factor authentication, closing open inboxes, removing public mentions of family members and personal contacts who might be harassed, and even temporary account deactivation.

---

[94] https://www.thetimes.co.uk/article/the-bbcs-marianna-spring-its-really-normal-to-really-hate-me-g8nkbmfvl

Newsrooms and organisations should provide access to security tools such as password managers, VPNs, and privacy services like DeleteMe, which combs the dark web for leaked data and removes it.

There must also be a ready protocol for when security threats bleed from online to offline — such as the release of a statement condemning the accusation, the provision of a safehouse, and other support systems.

Access to welfare options

Newsrooms should institute a welfare protocol, and let journalists know the types of support they offer including leave, psychological first aid or debriefings, mental health counselling, legal counselling, and support, and so on. If an organisation cannot provide one or more of these services, it should partner with or direct the journalist to resources that can.

They must also rethink broad and restrictive rules that might re-traumatize victims of malign influence campaigns. Some local newsrooms, at least on paper, encourage (if not impose) total non-engagement with online attacks, but may not have the internal mechanisms in place to adequately protect journalists or hold online abusers accountable. While journalists should be discouraged from engaging directly with malicious posts, they should not be punished for calling out these attacks, putting forward their own counter-narrative, or clearing their name.

Mob censorship and silence is one of the intended effects of lie machines. Completely ignoring these attacks, and leaving malicious content to dominate comment sections or search results, also allow for an impression of the diminished credibility of journalists.

Consider releasing a statement of support

Newsrooms must defend their employees and condemn disinformation campaigns against them, especially when attacks follow dog-whistles by prominent government figures. However, the consent and security of the journalist involved is key, as such statements also risk publicising the case even further.

Shawn Crispin, senior Southeast Asia representative of the Committee to Protect Journalists, said that the statements should do no harm. When evaluating the harm publicity can cause, consider: "Is it a street thug, or a senator who is blowing hot air? Who's dispensing these threats and how credible is it that they would act on [it]?" In some cases, when state leaders and officials at the national level pick on individual journalists, a statement can push for accountability, signal support for the journalist, and exert external pressure. This may not always translate at local and regional levels, where journalists may be more vulnerable to actual threats on their life.

In the case of hyper-partisan influencers, "you can do them most harm by taking away their platforms," Crispin added. If the target of online abuse

consents, one can "name and shame," as well as call on platforms to sanction the influencer for abusing community guidelines or terms of service.

The best statements are strongly worded and strategically timed. "You don't want to make a statement that doesn't go anywhere," he said. "Media pickup is key. You get the algorithm working in your favour."

Clear calls for action – such as the takedown of a post or account, call for investigation – can strengthen a statement. Background context and data gathered from documenting these attacks, or noting patterns of abuse, can also lend credence to the case.

Cultivate a supportive culture

Targets of online violence should not have to comb through the vitriol they receive themselves. When abuse gets overwhelming, a targeted journalist should be able to turn over their account to a friend or colleague who could wade through the notifications, report, block, and hide malicious accounts and comments on their behalf. Offering solidarity to colleagues online and offline (reverse swarming) is cited by the ICFJ among the most effective way to counter online harassment.[95]

---

[95] https://www.icfj.org/our-work/chilling-global-study-online-violence-against-women-journalists

Conclusion

The Philippines has been branded as a site of innovation for disinformation and influence techniques. Filipinos are at risk of exposure to or being targeted by malign influence campaigns – not just from politicians but state actors such as police, the military, and foreign operatives like China.

Although such campaigns have been normalised and are practised across different sides of the political aisle, the most impactful influence operations are those of President Ferdinand Marcos Jr. and his family, who once ruled the Philippines in a bloody and corrupt two-decade dictatorship.

It is a testament to the power of marketing and public relations that within the span of less than four decades, a dictator's family has gone from being seen as political pariahs to underdogs making a royal comeback. The cumulative mythos the Marcos family relied on – from a conspiracy theory involving gold to their return to power being prophesied by Nostradamus – has captured the Filipino public imagination far more than old, boring facts.

To shed light on this warped mythology, journalists should work to expose – and protect themselves from – the different cogs of what Oxford Internet Institute professor Philip Howard called the "lie machine": the politicians behind the lies, the platform and algorithm distributing it, and the commercial agents profiting from it.

In this paper, I've used case studies, interviews, and research to highlight how journalists in the Philippines are responding to, and how they are target of, malign influence campaigns. I also propose some ways to move forward, practices and systems we can adapt, and editorial angles we can explore.

Journalism must graduate from reactive interventions and simplistic disinformation coverage, and move toward making the press more resilient to influence operations both within the organisational structure and editorial production.

I proposed four major recommendations:

1. Institute an organisational disinformation and influence reporting strategy
- News organisations need a resident expert, if not a whole team, with specialised skills who can lead the charge.
- Echoing Rebecca Skippage, good information needs to go back on the same platform as bad information, and in the same language.
- News organisations must invest in training journalists with specialised skills, such as data analysis and knowledge of tools.
- Partner with the academe and civil society — both for news content, such as in examining data for investigations, but also for other initiatives to assure good information has the highest possible reach.

2. Cover disinformation as a structural issue

- Weigh newsworthiness. This can be done by examining criteria like a tipping point, social benefits, and potential harms (Data and Society). Journalists can also look at the breakout scale (Nimmo) to assess the damage caused by influence operations.
- Tell stories about human interest, platform accountability, and institutional accountability. Each of these angles are almost like the antithesis to different parts of the lie machine: the human harm that is caused by malicious political ideologies and intentions, platforms and their algorithms, and businesses and institutions profiting from or complicit with these influence operations.
- Exercise strategic amplification. Consider if you might be needlessly amplifying bad faith narratives and voices.
- Do not talk down to sources and audiences. Treat audiences and sources with empathy and equality.

3. Combining previous literature and journalists' experience, I compiled tips for how to sensitively approach and report about hyper-partisan characters

- Approach critically and with caution. Consider whether the hyper-partisan voice is really needed in your report. Examine their place in the disinformation or influence hierarchy and other nuances of their character, and run a thorough background check.
- Avoid outsourcing news from manipulative actors. Putting hyper-partisan and questionable personalities on blast only serves to give them attention.
- Brace for hostile behaviour. Prepare to face volatility on the field and online, and take necessary precautions.
- Put hyper-partisan and manipulative actors in context. Explain their motivations, interests, and ideology. Avoid pandering to their language and justifications.

4. Lastly, I argued for the institutionalisation of a standard protocol in preparation for and response to malign influence operations targeting journalists

- Track attacks against the company and employees. Gathering data points might be helpful for possible evidence later – whether for journalistic investigations, legal cases, or platform takedowns.
- Provide regular security assessments, tools, and training. In the Philippines, news organisations might have protocols in place for armed conflict, natural disasters, and climate hazards – but it's time to update that protocol to include malign influence operations.
- Provide and connect journalists to welfare options such as vacation leave, psychological debriefings and counselling, legal support, and so on.
- Consider releasing a statement of support. This signals an institutional stand, which expresses solidarity with the targeted journalist. It could also help put

enough pressure on institutions to hold the attacker accountable, or in the case of politicians, even issue an apology.

- Cultivate a supportive culture. Journalists facing online abuse need the support of their colleagues; lend a helping hand.

Despite dwindling trust in the news, the past few years are not without wins. People still look for the clarity of journalists in times of conflict, health and climate crises, and tragedy.

How does one dismantle a lie machine? A truth machine, Howard proposes. Hopefully the above action points can contribute to the construction of such a machine, at least for those in the Philippines.

There's only so much that journalists can do: our part is to tell true stories, deeply and sensitively. And our work ends where the work of others – academics and educators, big tech, civil society, and the public – begins.

But in the same way that the lie machine is a many-headed thing, our truth machine should – through innovation and collaboration – be similarly dexterous.

Resources

Training resources

- [Verification Handbook 3: For Disinformation And Media Manipulation](#) by EJC

- [Digital Forensic Research Lab training](#) by Atlantic Council

- [A Handbook for Journalism Education and Training](#) by UNESCO Journalism, 'Fake News' and Disinformation

- [FIMI: Towards a European Redefinition of Foreign Interference](#) by EU Disinfo Lab

Handbooks and guides

- [CPJ Digital Safety Guide](#)

- Protecting Newsrooms and Journalists Against Online Violence by IWMF

- [The Philippine Journalists' Safety Guide](#) (2020), has a chapter dedicated to digital and information security, with a few tips for online harassment

- UNESCO's [The Chilling: what more can news organisations do to combat gendered online violence?](#)

- [The Media Manipulation Casebook](#)

- [10 Tips for Reporting on Disinformation](#) from Data & Society

- CPJ Digital Safety [Working from Home](#)

Video resources

- Craig Silverman on [Exposing the people and funding behind websites and digital ads](#)

- Jane Lytvynenko and Craig Silverman on [Investigating media manipulation campaigns](#)

Online tools and support organisations

- [The Breakout Scale: Measuring the Impact of Influence Operations](#) by Brookings

- The DISARM (Disinformation Analysis and Risk Management) framework

- Reporters Without Borders' [forensic services for digital attacks](#) on journalists

- [The Online Response Hub](#) by the Coalition Against Online Violence

Research and books

- [Disinformation 2.0: Trends for 2021 and beyond](#) by the European Centre of Excellence for Countering Hybrid Threats

- Harvard Kennedy School's [Misinformation Review](#)

- [How to Be a Woman Online: Surviving Abuse and Harassment, and How to Fight Back](#), by disinformation expert Nina Jankowicz, published in 2022, contains practical tips for digital security and fostering safe spaces.

Acknowledgements