



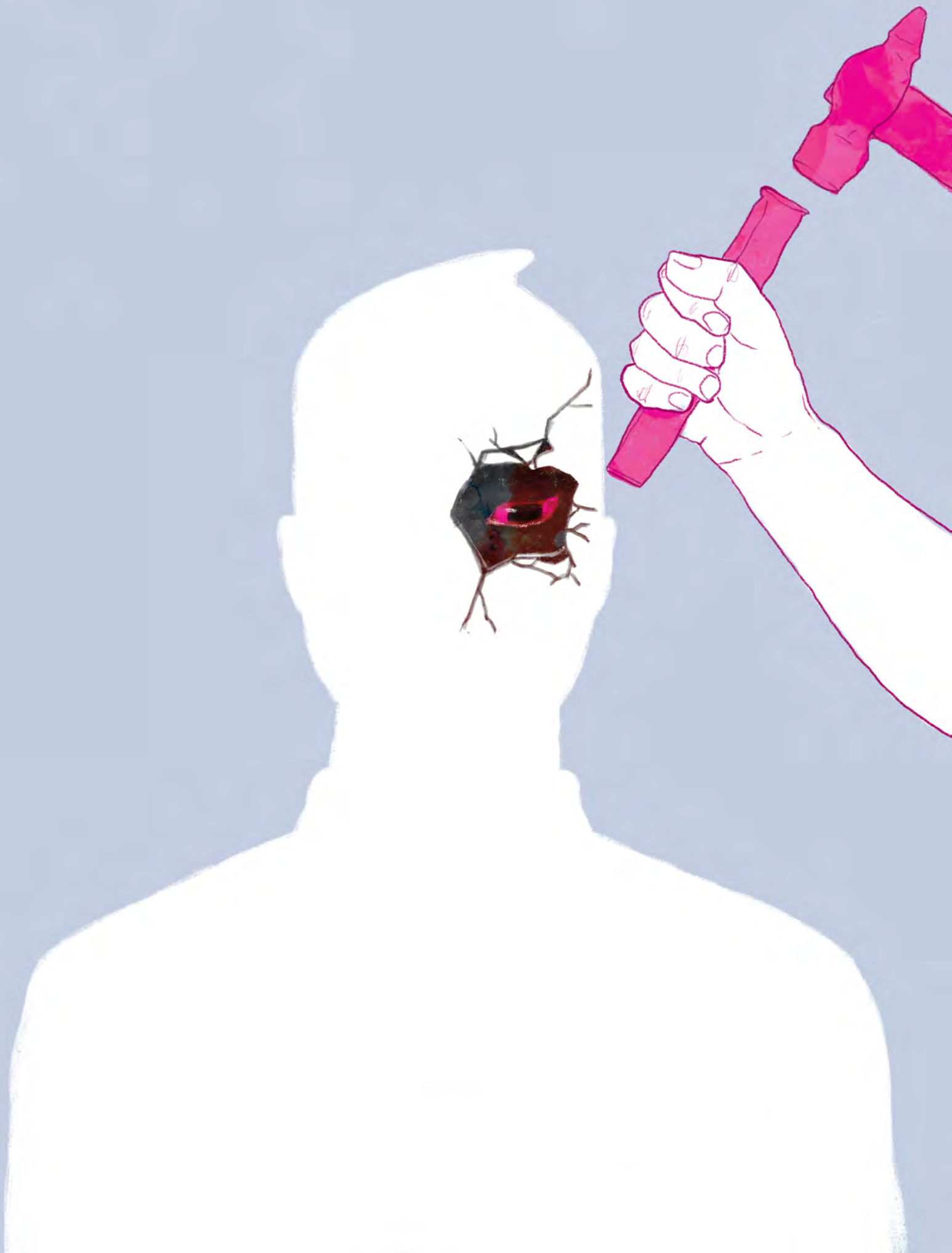
Architects of Networked Disinformation

Behind the Scenes of Troll Accounts and
Fake News Production in the Philippines

By Jonathan Corpus Ong & Jason Vincent A. Cabañes



UNIVERSITY OF LEEDS



Architects of Networked Disinformation

Behind the Scenes of Troll Accounts and
Fake News Production in the Philippines

Jonathan Corpus Ong
Jason Vincent A. Cabañes

An Executive Summary

About the Authors

Dr Jonathan Corpus Ong is Associate Professor in Global Digital Media at the University of Massachusetts. He is the Convenor of the British Council-funded Newton Tech4Dev Network (www.newtontechfordev.com), a consortium of global researchers investigating emerging technologies in low- to middle-income countries. He is the author of the book *The Poverty of Television: The Mediation of Suffering in Class-Divided Philippines* (Anthem Press 2015) and co-editor of the volume *Taking the Square: Mediated Dissent and Occupations of Public Space* (Rowman and Littlefield 2016; with Maria Rovisco). He is the Co-Editor-in-Chief of the journal *Television & New Media*.

Dr Jason Vincent A. Cabañes is Lecturer in International Communication at the School of Media and Communication, University of Leeds. In the Newton Tech4Dev Network, he co-leads (with Cheryll Soriano) the research stream on digital labor in the global South. He has previously co-authored (with Jayeel Cornelio) a chapter on the rise of trolling in the Philippines in the book *A Duterte Reader*. He has also published in top-tier journals such as *New Media and Society*, *Media, Culture, and Society*, and *South East Asia Research*.

Acknowledgments

This report is produced by the Newton Tech4Dev Network—a global consortium of scholars that aims to advance research on digital technologies in low- and middle- income countries. Funded by a British Council Newton Fund grant, the network supports international collaborations around research areas of digital labor in the global South, disaster management and digital media, and crisis cultures. The Newton Tech4Dev Network is co-hosted by the University of Leicester in the UK and De La Salle University in the Philippines. Visit our website www.newtontechfordev.com for more details about the network members and their various projects.

This report includes contributions from our project research assistant Pamela Combinido based in De La Salle University Philippines. It also features additional reporting from Deepa Paul-Plazo.

The report layout is by Kay Aranzanso, and report illustrations are produced by Tokwa Peñaflorida.

We would like to thank the generous and insightful group of academics, journalists, tech experts, lawyers and policymakers who served as advisers to this project and gave extensive feedback on an early draft of this public report.

Overview

No technology has been weaponized at such an unprecedented global scale as social media.

Diverse research approaches now attempt to decipher how laptop screens and smartphones around the world are used to manipulate public debate, hijack mainstream media agenda, and influence political outcomes.

In the Philippines, the names on the lips of political pundits and the mainstream press are those of the social media influencers who command large “troll armies” credited with sweeping Rodrigo Duterte into unforeseen election victory in 2016. Under Duterte’s presidency, “trolls”, or “Dutertards” as his fanatic followers have been dubbed, are seen to have debased political discourse and silenced dissidents in their vociferous sharing of fake news and

amplification of hate speech. But who exactly is responsible for leading disinformation campaigns? What kind of people sign up to become “trolls”, and what industries and social groups do they come from? What technical skill sets do they have, and how do their labor conditions shape the campaigns they create? Most crucially, how do they justify the work that they do and live with the stigma of being labeled a “troll”?

Through interviews and participant observation of people we call **architects of networked disinformation**, our study uncovers the professionalized and hierarchized group of political operators who design disinformation campaigns, mobilize click armies, and execute innovative “digital black ops” and “signal scrambling” techniques for any interested political client. This report reveals that disinformation architects are a crucial and common part of Filipino political campaigns at both national and local levels—employed by many politicians regardless of party and ideology.

Stemming from the Philippines’ image-based political system, the chief architects of networked disinformation come from the advertising and PR industry, whose mastery in corporate marketing hyperextends to an unregulated and highly profitable industry of digital political campaigning.

Conveniently hidden behind the smoke-screen of prominent social media influencers who fan the flames of political divisiveness and stoke the public’s moral panics about trolling,

ad and PR executives assemble their own teams of anonymous digital influencers and community-level fake account operators.

Tasked to translate core campaign messages, these anonymous influencers and fake account operators weaponize their fluency with the popular vernacular in covert digital operations designed to mobilize populist public sentiment. While the basic blueprint of political disinformation campaigns may appear similar to corporate brand strategy, networked disinformation involves many moral compromises: from seeding revisionist history narratives to silencing political opponents to hijacking news media attention through artificially trending hashtags.

Drawing from the tradition of production studies in media and communications research, our study contributes an ethnographically informed understanding of how the personal motivations, social backgrounds, and everyday routines of digital workers impact on their production of networked disinformation and fake news. Our study aims to move beyond easy dichotomies of heroes and villains and does not seek to name and shame individual paid trolls.

The problem of disinformation production goes deeper than any one caricatured hero or celebrity villain; it is systemic, deeply rooted, and entwined in the cultural fabric of Philippine society. Behind the madness is an invisible machine: industrial in its scope and organization, strategic in its outlook and expertise, and exploitative in its morality and ethics. At the helm of the machine, the chief architects of disinformation hide in plain sight, wearing respectable faces, sidestepping accountability while the public's moral panics about trolling are directed elsewhere.

This report thus aims to develop a critique of ecological vulnerabilities in the creative industries in the Philippines that enable politicians to recruit highly skilled, if corruptible, disinformation architects to collude with them without industry self-regulatory mechanisms and sanctions in place. This report also identifies large gaps in Philippine campaign finance legislation and digital platform regulation, and proposes preliminary recommendations to address these issues. Finally, the study aims to invite ethical reflection about the process in which ordinary people become complicit in deception work as they aspire for financial gain or seek political and symbolic power. Through the set of preliminary recommendations we present, we open the conversation as to how we can reinvigorate professional ethics, uphold worker justice, and create cross-sectoral advisory groups with lawyers, academics, platform designers, and creative professionals to address our individual, social and cultural complicity in networked disinformation.

In this report, we:

- 1** Narrate “deep stories” of individual workers positioned at different levels of the hierarchy, not to vilify them but to understand their motivations and social backgrounds
- 2** Discuss the labor arrangements that underpin networked disinformation in order to reveal the vulnerabilities of professional industries and institutions to political deception work
- 3** Discuss the persuasive techniques that architects of networked disinformation deploy in mobilizing populist sentiment to further clients’ elite agendas and for their own economic and political gain
- 4** List preliminary recommendations aimed at every level of fake news production’s hierarchical structure

Seven Key Findings

1

The use of fake accounts and paid influencers on Facebook and Twitter for political operations is widespread. Multiple political parties at both national and local levels make use of “click armies”. Previous reports¹ have spotlighted only Duterte’s Partido Demokratiko Pilipino Lakas ng Bayan (PDP-Laban) as hiring fake account operators, thus overlooking the systematic manipulation of political discussions on Facebook and Twitter by various players across the political spectrum.

2

Politicians often employ campaign strategists from local “boutique” advertising and PR agencies as chief architects of networked disinformation campaigns. These experts use tried-and-tested corporate branding techniques such as “core campaign messaging” and “brand bibles”, while exploiting ecological vulnerabilities in the unregulated industry of political marketing. They also deploy sketchy new digital strategies such as what we label “signal scrambling” and “digital black ops” to distort trending rankings and hack the attention of mainstream media.

¹ Bradshaw, Samantha and Philip Howard. 2017. Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation. <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/07/Troops-Trolls-and-Troublemakers.pdf>; Jenkins, Ciaran. 2017. “Fake News: Trolling in the Philippines.” Channel 4, 10 February. <https://www.channel4.com/news/fake-news-trolling-in-the-philippines>; Ressa, Maria. 2016. “Propaganda War: Weaponizing the Internet.” Rappler, 3 October. <https://www.rappler.com/nation/148007-propaganda-war-weaponizing-internet>

3

Ad and PR strategists delegate political marketing responsibility. They rely heavily on the promotional labor of digital influencers (who have between 50,000 to 2,000,000 followers on Facebook and Twitter) and community-level fake account operators (who manually operate fake profiles to infiltrate community groups and news pages)–and very minimally on automated bots. While the ad and PR strategists are usually paid a lump sum by their political clients on a per-project basis, they subcontract work to influencers and pay them following PR industry standard matrices of reach and engagement. Community-level fake account operators are paid a fixed daily rate based on a set quota of online posts or comments. The incentive scheme that the strategists set for the influencers and account operators maximizes potentials for the signal boosting of communication messages that strategically use popular vernaculars resonant with populist public sentiments.

4

Disinformation workers are financially, politically, socially, and psychologically driven in different ways. The infrastructure of networked disinformation is built on a relationship of competitive collegialities. The people we interviewed are primarily driven by financial motivations, but most of them are actually politically aligned with their client. A few justified their engagement in the work of fake news production with the psychosocial fulfillment of doing edgy, experimental work. Many of our respondents entered digital underground work after being disillusioned in the creative industries from witnessing first-hand systematic corruption and/or experiencing exploitative work arrangements.

5

Operating fake accounts for politicians involves similar modes of always-on, flexible, and (self-)exploitative arrangements as other online freelance work. However it is accompanied by the stressful emotional labor of justifying this work both to others and themselves. Workers develop implicit rules to establish a moral order and distance themselves from the stigma around trolling. The workers we met claimed to avoid using sexist language in their fake accounts and attributed the use of offensive language to “real people” or “real supporters” of their political client. Many workers view their fake account operations as one sideline among many, and expressed a wish to be judged based on their primary jobs and/or future aspirations.

6

Networked disinformation campaigns operate with two opposing dynamics in play. On the one hand, **controlled interactivity aims for collective participation and cooperation among disinformation workers who are informed by a common script; on the other, volatile virality relies on these workers' individual insight and creativity in translating a script into social media posts that achieve maximum, if uncontrolled, spreadability** across decentralized networks of communicative exchange. Emotionally charged campaigns tapping into populist sentiments of anger and resentment may thus achieve their strategic goals, but inadvertently unleash uncivil expressions of misogyny, anti-intellectualism, and other forms of offensive speech into the public discourse.

7

While nobody really admits to being a troll, everyone in the disinformation hierarchy seems to be engaged in various degrees of trolling. Many workers justify their use of fake accounts on Facebook and Twitter as a mere extension of the longstanding ad and PR practice of using spin. Ordinary citizens have long tolerated deception work from “legitimate sources”—corporate brands, celebrities, journalists and oligarchic media. This long-standing acceptance paved the way for political disinformation to thrive unregulated in a digital underground that only very subtly hides from plain sight.

Our full report argues that there is no one-size-fits-all solution to the complex problem of networked disinformation. While efforts to blacklist fake news websites, expose fake accounts, or vilify divisive digital influencers may sometimes be well-meaning, these solutions do not address the institutions and systems that professionalize and incentivize disinformation production. The project reflects on the vulnerabilities in the Philippine political culture and creative industries that entice digital workers to become architects of disinformation. It also reflects on the potentially far-reaching consequences of the “stockpile of digital weapons” in the Philippines, with its highly organized online freelance labor force, to democratic countries in the West, and vice versa. Policy recommendations in the report address:

- Advertising and PR industry self-regulation
- Political campaign finance regulation
- News media coverage and investigative reportage of fake news and trolling
- Platform intermediary regulation responding to particular concerns of fragile democracies in the global South

We also advance new concepts to deepen our understanding of the operations of networked disinformation:

1 Disinformation interface: the porous boundaries between so-called “paid troll” work executed by influencers and fake account operators and the enthusiastic zeal of political fans and grassroots intermediaries

2 Moral justifications: the denial strategies and mental acrobatics that disinformation architects perform to dismiss their own personal responsibility to democratic processes and political exchange. Moral justifications include discourses of normalization (“I do the same thing for corporate brands”), fictionalization (“This is straight out of Game of Thrones”), and splitting (“This is just a sideline and not my real job”).

3 Competitive collegialities: the working relationships that disinformation producers have with each other. This is about how they are compelled to work together even if their professional aspirations are ultimately at odds with each other. While chief disinformation architects want to maintain the work hierarchy they established, lower-level political operators aim to overcome the power asymmetries.

4 Volatile virality: a key principle of promotional communication that runs alongside what Jennifer Stromer-Galley has called “controlled interactivity”². While controlled interactivity is concerned about message discipline throughout a campaign, we identify that it is challenged by an opposing principle of volatile virality, which is concerned with message amplification through the use of populist rhetorical styles resonant with political fans and grassroots supporters.

For a free download of the full report, please visit: www.newtontechfordev.com

² Stromer-Galley, Jennifer. 2014. *Presidential Campaigning in the Internet Age*. Oxford: Oxford University Press.

THE ARCHITECTURE OF NETWORKED DISINFORMATION

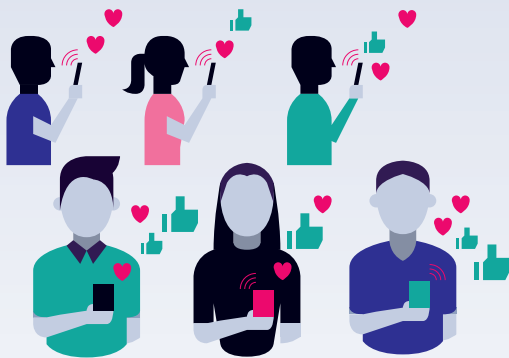
REFERENCE: Ong, J.C. & Cabanes, J.V. (2018). Architects of Networked Disinformation

Political Clients



Chief Architects of Networked Disinformation

Elite advertising and PR strategists: they liaise with political clients and set campaign objectives



Digital Influencers

Anonymous Influencers:

Anonymous operators of social media pages with humorous/ inspirational/ pop culture content: they translate campaign messages into viral posts

Key Opinion Leaders:

Celebrities and pundits with highly engaged fans and followers on social media: they carry core campaign messages



Community-Level Fake Account Operators

Precarious middle-class workers subcontracted by ad and PR strategists or hired by politicians' chief-of-staff: they amplify reach and create "illusions of engagement"



Grassroots Intermediaries

Politician's fan page moderators, unpaid opinion leaders, volunteer political organizers

PUBLIC

PAID

VOLUNTARY



