Digital Public Pulse: Mapping the digital landscape of Philippine politics

Jon Benedik A. Bunquin
Fatima Gaw
University of the Philippines Diliman

Throughout history, politics in the Philippines has always been about establishing political bulwarks, forging elite alliances, and mobilizing grassroots communities *on the ground.* It was only in 2016 when then-presidential candidate Rodrigo Duterte organized his campaign around social media and rallied individuals who were unseen, unheard, and marginalized in mainstream discourse to propel his presidential bid to popularity (Tapsell, 2021; Sinpeng et al., 2020). Emergent digital publics such as those who supported Duterte, as well as those who were critical of his administration's policies, have become a force in contemporary Philippines politics by setting agenda, expressing dissent, and calling for political action (Auethavornpipat & Tanyag, 2021; Bunquin & Gaw, 2021; Aranda, 2021). Their capacity to exercise political agency materialize in Twitter trending topics, Facebook groups, YouTube vlogs, and other digital spaces without the gatekeeping of news media, nor the mediation of experts, professionals, and other institutional actors.

It was inevitable that the same enabling digital media environment was used to manipulate publics by groups with vested political interests. Disinformation online has been the hallmark of the Duterte regime (Lees, 2018), underpinned by political manipulation operatives from the PR and marketing industry (Ong & Tapsell, 2022; Ong & Cabañes, 2018) and institutionalized into government operations (Davidson & Fonbuena, 2020; Bradshaw & Howard, 2017). The co-option of the digital public sphere by ill-intent entities has become endemic to Philippine politics in later elections, benefitting local and national politicians alike (Ong et al., 2019). This has also paved the way for digital influencers as new political intermediaries to advance political agenda in parallel and more often than not, in conflict with mainstream media, scholarly institutions,

and historical accounts (Soriano & Gaw, 2022a, 2022b; Ong & Cabbuag, 2022). Together, they weave complex digital networks of actors that purport to represent publics in mainstream political discourse but are in fact, working to undermine their political agency and manufacture political outcomes.

Designing a big data election research

These realities serve as the *raison d'etre* of our research project *Digital Public Pulse*. Digital Public Pulse (DPP) is a pioneering, interdisciplinary media monitoring collaboration that inquires about the intersection of digital publics, social networks, and media convergence. In its pilot year, DPP investigated the digital landscape of the 2022 Philippine General Elections across the major social media platforms Facebook, Twitter, and YouTube. It mapped the network actors engaged in election-related conversation, examined salient topics, agenda, and issues they discuss, and analyzed how publics, platforms and discourse characterize contemporary political engagement on digital media. Along the way, the study also took a deep dive into the indicators of networked political manipulation that emerged in our preliminary findings, which were publicly disseminated through the Facts First PH initiative.

The research was the first of its kind in many ways, engaging in interdisciplinary perspectives from communication and media studies, political science, and data science to perform large scale multi-method analysis using network analysis, textual analysis, and other computational methods. It oscillated between different modalities (e.g., audio-visual, textual, interactive), scope (e.g., actors, communities, networks), and interests (e.g., political influence, messaging frames, disinformation), all of which constitute the vastness and heterogeneity of mediated political engagement online. The defining character of the research was its cross-platform analysis that drew critical insights both from each social media platform having its distinct architectures that shape political experiences, and across platforms and the ways they configure digital political cultures in the Philippines. Digital Public Pulse also covered twelve months of data to have a longitudinal view of the developments in election discourses, from the beginning of election year until the election day itself.

The design of the research is intricate and deliberate. As a media monitoring project, it documented the digital public sphere encompassing multiple phenomena, platforms, and concepts. As much as it aimed to

capture these events and interactions in their entirety, the nature of digital data as partial, semi-public, and contingent on access provided by the platforms restricted us from full view of political activities and engagement online. As such, we developed a methodology that pushed the boundaries of data sourcing, collecting, and analysis and employed triangulation by using multiple methods across spaces and across time to ensure the rigor in our research. We describe our approach as *partial* but *comparatively extensive and exhaustive* in a way that it encompassed both major and minor forces that shape political relations online, including those with intent to distort, manipulate, and ultimately misrepresent publics through covert influence operations. Such endeavors to innovate research methods did not come easy and we discuss our challenges later on in this article.

The research also instigated rethinking and reframing of how we understand political processes as mediated by digital media technologies. Existing election research in the form of pre-electoral surveys and more recently, sentiment analysis often construct the voting public based on normative categories of socio-economic class, region, gender and other immutable attributes. These may inform political behavior but are ultimately limited in giving insight into how Filipinos' political dispositions have been formed in the first place, as they are entrenched in social networks, participate in interactive discourses, and use digital technologies in their everyday lives. Digital Public Pulse regarded the members of the public in this intersection as people who exercise their political agency beyond the dictates of their social positions and through the new tools for expression afforded to them by contemporary media environments. In doing so, we decentered our interests from the traditional political interlocutors who are assumed to strongly shape public opinion, to those new interlocutors who partake in the political discourses with the knowledge and skills to navigate the vernaculars of the platforms and engage in the digital cultures that animate mediated political processes online.

Digital Public Pulse translated its research findings into actionable recommendations to relevant stakeholders in pursuit of policy directions and program strategies that uphold democratic values and protect electoral processes from undue political influence. Despite the undeniable impact of social media during the last two national elections, interventions and policies remain to be siloed in specific agencies and sectors and often, they engage in more normative practices that fail to

account for the distinct ways mediated political processes operate. Our research project presented evidence that insists that stakeholders need to break off from business-as-usual and develop more nuanced skills and new capacities to address issues in these uncharted, technically complex digital spaces. Stakeholders such as state agencies, the media, civil society, academics and the public are all enjoined to reflect and criticize their policies and practices and their impact (or the lack thereof) on the election process, and to reform and evolve their perspectives and strategies in response to a digitally-mediated political environment.

Challenges of doing big data research to study politics in the Philippines

Our use of digital trace data enabled us to map the Philippine election landscape by disentangling the complex networks formed through online discourse and reveal the unique contours and textures of electoral politics as they transpire in social media platforms. We investigated traces of human behavior that are abundant in digital spaces, left each time users interact in and with platforms, and engage with digital objects and actors. However, embarking on such a research journey requires a consideration of the researchers' skills and resources, an understanding of the data and its nature, a consciousness of the contexts where the phenomenon is taking place, and a cognizance of users' rights to privacy and audience characteristics in the analysis and reporting of findings. We outline each of these elements here and the challenges we encountered, citing specific instances during our study as examples and learnings that can be derived from them.

1. Harnessing skills and resources appropriate for mapping digital spaces. The endeavor to map the uncharted digital terrain of electoral politics in the Philippines requires an understanding of cultures and ways of communicating in online spaces, as well as tools to grapple unique elements that make up these environments. Our experience in using various open-source and affordable tools to collect data from YouTube and Twitter data, having done previous work on digital media and social network analysis (e.g., Bunquin & Gaw, 2021; Soriano & Gaw, 2022b) provided us with an initial set of tools to collect and examine digital data. We also initiated mini-workshops and training sessions to equip other members with a basic understanding of how these tools work, so we can better understand what questions can be asked, what types of data can be collected, and what analytical approaches can be performed.

But the tools were limited in scope, and while we possessed domain expertise in the area of digital communication and media research, as well as extensive experience doing quantitative and qualitative research, we were entering the territory of big data, which required competencies beyond traditional statistical approaches and qualitative coding. Technical expertise was also warranted. Thus, we sought collaborations with various institutions who shared the same vision of upholding and strengthening democratic processes and institutions. We reached out to the University of the Philippines' Data Science for Society program, which connected us to data science experts in the disciplines of mathematics and political science to serve as collaborators and technical consultants in the project. We also collaborated with AI4GOV, a start-up company that utilizes artificial intelligence in building "solutions for public service and citizen engagement" (AI4GOV, n.d.), as they provided their expertise in developing machine learning models to classify textual data. Moreover, we were able to collaborate with our students through the laboratory courses of the Department of Communication Research, as they actively took part in the research during its data collection and analysis phases. The project also served as a venue for experiential learning and expanding their research competencies. We also reached out to Rappler, who provided us access to Crowdtangle, a social media listening tool developed by Meta, Facebook's holding company. Rappler also provided our research project a venue to disseminate its findings, so we can reach a broader audience. From these experiences, we found that interdisciplinary collaborations and academe-industry partnerships were pivotal for the success of the big data project, as they foster creative and effective ways of using digital data to answer relevant and important questions in the field (Cesare et al., 2018).

    2. Understanding digital trace data and its limitations. The large, always-on, and accessible features of digital data (Salganik, 2018) provide opportunities for researchers to study phenomena as they happen real-time. Digital public pulse collected data on a rolling basis, allowing us to capture behaviors at the point of action and observe the crests and troughs of various behaviors throughout the election year. However, because they manifest in algorithmically confounded spaces (Salganik, 2018), with users engaging in performative acts before an imagined audience (Marwick & boyd, 2011), and computational propaganda shaping digital behaviors and manipulating discourses (Bolsover & Howard, 2019), we had to be cautious of analyzing and reporting these behaviors as purely

authentic and ensure that influences of algorithms, performativity, and manipulation were considered in the design of the research, development of procedures, coding of data, and making inferences. To illustrate, in our Facebook data, we found pages that are seemingly irrelevant in election conversations, as they appear to be posting content about romance or music based on their page names and descriptions. But our knowledge of pseudonymous actors (Ong and Cabbuag, 2022), obfuscated propaganda in digital media (Farkas and Neumayer, 2018), and the affordances of the platform prompted us to further investigate, and we found these pages to be posting the same content about the elections, all at the same exact time in what appears to be a coordinated effort to push for specific messages about a candidate. We also found many others that were doing the same in our dataset.

These insights are not automatically observable in the data, and they can get lost in the millions of rows of the dataset. But because of our understanding of platform affordances and cultures, we were able to surmise various reasons why these behaviors exist and further investigate. Big data does not automatically equate to good research. It is also necessary to state that digital trace data, no matter how large it may be, is not generalizable without the application of complex estimation processes, as data generated by platforms are typically skewed towards a certain population (Salganik, 2018), and not every member of the population has access to the internet. Researchers must be conceptually clear on why they are engaging in big data research, and posses a comprehensive understanding of how such data is generated and structured, and what underlying assumptions and insights can be unearthed from them. In the field of data science, this is referred to as *data understanding* (Wirth & Hipp, 2000), and it operates alongside *business understanding*, or the knowledge of one's objectives and requirements in the data project, so that data can be harnessed properly to achieve goals.

    3.   Contextualizing findings within the dynamics of social and mediated realities. Because digital trace data is derived from agents and their online behaviors made accessible through platforms, they are constantly changing. Salganik (2018) calls this *drift*, or the changing patterns in the data that may be caused by a change in the population of users from whom data is collected or *population drift*, behaviors of people whose data is being collected or *behavioral drift*, and systems that make data collection possible or *system drift*. In studying the elections, we had to contend with the various drifts that posed several challenges in collecting and

interpreting data. Fluctuations in the volume and types of data collected during the election cycle were observed, peaking during specific points of the election season, such as the filing of candidacy, airing of election debates, and start of the election season. In analyzing the data, we had to be mindful of these events – understanding how digital election texts were operating within these contexts, and how patterns of linkages among online users can be connected to events happening outside of digital platforms. Our analyses were performed with these contexts in mind, and segmenting our data into periods based on key events allowed us to formulate context-based insights.

But aside from changes with what is happening 'out there,' system drifts also posed challenges in our research. During our data collection period, platforms were also actively engaged in taking down manipulative actors within them (e.g., Reuters, 2022; Gregorio, 2022), and making changes within their systems themselves (e.g., The YouTube Team, 2021). When our findings showed lower proportions of suspicious actors during the second period of our data collection, we had to be conscious of the take-downs implemented by platforms prior to our data analysis, and include the context of the finding in our reports. By contextualizing data within the changes taking place inside and outside of platforms, researchers analyzing digital trace data exercise rigor and can report findings in a more nuanced and accurate manner. Moreover, by being agile in our research, we minimized the effects of drifts due to changes in the system and ensured that we can provide explanations for sudden shifts in our findings.

4. Confronting the consequences of doing and reporting big data research in contentious contexts. Aside from the technical and operational aspects of doing big data research, we also had to contend with ethical dilemmas emerging as we conducted our study. Salganik (2018) likens the role of digital researchers to the person in the watchtower of a panoptic prison (Foucault, 1995, as cited in Salganik, 2018), with the ability to observe behaviors of multiple people at once and create a database of individuals, thinking of various research studies that can be produced. Such databases can be detrimental, especially if they are used for surveillance purposes with the end goal of manipulating user behavior. To avoid unanticipated secondary use (Salganik, 2018), we had to ensure that data was stored in a place that can only be accessed by selected members of the research project. We valued user privacy by ensuring that no private user information will be shared in research reports, even if such a user turned

out to be non-human or a bot. This was especially true for platforms like Twitter, which had documented evidence of computational propaganda (Bolsover & Howard, 2019), and whose privacy settings enabled users to make their posts private or public anytime they want. We also only collected publicly available information, and made sure to report findings at the aggregate-level. Posts among users considered as public figures, such as politicians and news media organizations, were the only ones reported in public venues. We also constantly exercised reflexivity by asking why we were collecting specific types of information, and how such information can serve public interest and our larger goal of providing insights for various stakeholders in upholding democratic values. We regularly convened with all members of the laboratory, ensuring that multiple perspectives were taken into account before arriving at decisions related to ethics and user protection.

In reporting our findings, we were cognizant of the technicalities involved in big data research. Hence, we focused on communicating insights that can be derived from the numerical information in our public presentations, and directed recommendations towards specific stakeholder groups who can create interventions and act on the results of our study. We engaged representatives from the government, media, civil society, the academe, and the public in weekly research fora as part of our dissemination efforts. Our partnership with Rappler also allowed us to extend the reach of our research beyond online webinars, and our findings were eventually picked up by several media organizations, and prompted several appearances in the television programs, where we were able to talk about our research before audiences outside of online media.

But going public with our study invited attacks from hyper-partisan actors and rabid supporters, who disagreed with or diminished the value of the results of our research. These attacks ranged from being the subject of ridicule in comment sections and getting hateful direct messages in personal social media platforms, to being actual topics of YouTube videos with the aim of discrediting our work. To protect ourselves, we had to ensure that our social media platforms were locked and private. One of us also had to deactivate their channels to avoid being tagged and targeted. But beyond these self-protection mechanisms at the individual level, we believe that higher educational institutions must establish clearer protocols to protect its faculty members and researchers from acts that ultimately damage the academic institution and what it stands

for. We also believe in creating more institutional efforts to actively push back against malicious actors who aim to discredit academia.

Digital trace data carry uncertainties that may be difficult to demystify for researchers unfamiliar with the digital terrain, and much like the blackbox that characterizes algorithms in digital platforms, researchers are never 100% certain of the various forks in the paths (Gelman & Loken, 2014) that they will have to confront as they go through a series of decision-points in navigating this territory. However, these challenges provide researchers an "opportunity to expand researchers' investigative potential" (Cesare et al., 2018) and we believe that this can be done by harnessing novel ways of thinking and doing that leverage on interdisciplinary tools and expertise, engaging in collaborations between academics and the industry based on a mutual interest of social welfare, having a clear understanding of data and research contexts, and establishing ethical protocols that protect participants and researchers from the unanticipated harm that may emerge during the process.

Conclusion

We moved mountains to do this research because we know it is urgent, important, and critical in building knowledge and instigating action in the next years and elections to come. After establishing a working framework on performing social media monitoring research for tentpole democratic events such as the elections, we call on universities, researchers, and students to build their capacities and conduct similar kinds of research to institute a community of practice among scholars. Apart from the novel and critical analysis of mediated political processes advanced by the research, Digital Public Pulse serves as a comparative scholarly work and an evidence-rich artifact that can be the basis and/ or point of reference for future election research. In the same way we owe much from the works of pioneering scholars in Philippine digital cultures, we also think that researchers who build on our work are paving the way for new trajectories in scholarship in political communication, digital research, and in the communication field in the Philippines, more broadly.

Digital media is always changing and we are subject to the conditions and limitations of access and capacity so we need to work together to develop theory-first approaches and creative methodologies to navigate these contingencies. This responsibility should not lie with researchers alone, as many of the challenges in doing this kind of research, as we

have documented here, are because of technical limitations, resource constraints, and political realities of a hyper-partisan political culture. Communication and media departments should provide their researchers with up-to-date method training and supply them with the necessary equipment and tools to analyze big data. Universities should already pre-empt the need to fund cutting-edge election research years before the election year to ensure the continuity and quality of such initiative. Lastly, researchers are already taking the risk to perform research where they might be the subject of partisan attacks and politicking, and institutions must have a clear support system to protect the researchers and stand with them when they put themselves out in the open.

—

References

AI4GOV (n.d.) About us. http://ai4gov.net

Aranda, D. (2021). "Die-hard supporters": Overseas Filipino Workers' online grassroots campaign for Duterte in the 2016 Philippines elections. *Cornell International Affairs Review*, *14*(2), 93-127. https://doi.org/10.37513/ciar.v14i2.618

Auethavornpipat, R., & Tanyag, M. (2021). Protests and pandemics: Civil society mobilisation in Thailand and the Philippines during COVID-19 (Policy brief). *The New Mandala.* https://www.newmandala.org/protests-and-pandemics-civil-society-mobilisation-in-thailand-and-the-philippine/

Bolsover, G., & Howard, P. (2019). Chinese computational propaganda: Automation, algorithms and the manipulation of information about Chinese politics on Twitter and Weibo. *Information, communication & society*, *22*(14), 2063-2080. https://doi.org/10.1080/1369118X.2018.1476576

Bradshaw, S., & Howard, P. (2017). Troops, trolls and troublemakers: A global inventory of organized social media manipulation. In *Computational Propaganda Research Project* (pp. 1–37). Oxford Internet Institute.

Bunquin, J. B., & Gaw, F. (2021). Digital witnessing in a pandemic: An assemblage analysis of the Twitter hashtag network #MassTestingPH. *Southeast Asian Media Studies Journal*, *3*(1), 23–52. https://www.ejournals.ph/article.php?id=16530

Cesare, N., Lee, H., McCormick, T., Spiro, E., & Zagheni, E. (2018). Promises and pitfalls of using digital traces for demographic research. *Demography*, *55*(5), 1979-1999. https://doi.org/10.1007%2Fs13524-018-0715-2

Davidson, H. & Fonbuena, C. (2020, September 23). Facebook removes fake accounts with links to China and Philippines. *The Guardian. https://www.theguardian.com/technology/2020/sep/23/facebook-removes-fake-accounts-with-links-to-china-and-philippines*

Farkas, J., & Neumayer, C. (2018). Disguised propaganda from digital to social Media. In *Second International Handbook of Internet Research* (pp. 1–17). Springer Netherlands. https://doi.org/10.1007/978-94-024-1202-4_33-1

Gelman, A. & Loken, E. (2014). The statistical crisis in science. *American Scientist*, *102*(6). https://doi.org/10.1511/2014.111.460

Gregorio, X. (2022, April 7). Hundreds of 'malicious' Facebook accounts in Philippines taken down. *The Philippine Star*. https://www.philstar.com/headlines/2022/04/07/2172950/hundreds-malicious-facebook-accounts-philippines-taken-down

Lees, C. (2018). Fake news: the global silencer: The term has become a useful weapon in the dictator's toolkit against the media. Just look at the Philippines. *Index on Censorship*, *47*(1), 88-91. https://doi.org/10.1177/030642201876957

Marwick, A. E., & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New media & society*, *13*(1), 114-133. https://doi.org/10.1177/1461444810365313

Ong, J.C. & Cabañes, J.V.A. (2018). Architects of networked disinformation: Behind the scenes of troll accounts and fake news production in the Philippines. Report, Newton Tech4Dev Network. https://doi.org/10.7275/2cq4-5396

Ong, J.C. & Cabbuag, S. (2022). Pseudonymous influencers and horny 'alts' in the Philippines: Media manipulation beyond 'fake news'. *The Journal of Socio-Technical Critique*. 132. https://scholarworks.umass.edu/communication_faculty_pubs/132

Ong, J. C., & Tapsell, R. (2022). Demystifying disinformation shadow economies: fake news work models in Indonesia and the Philippines. *Asian Journal of Communication*, *32*(3), 251-267. https://doi.org/10.1080/01292986.2021.1971270

Ong, J.C., Tapsell R., & Curato, N. (2019). Tracking digital disinformation in the 2019 Philippine midterm election (Report). *The New Mandala*. https://www.newmandala.org/wp-content/uploads/2019/08/Digital-Disinformation-2019-Midterms.pdf

Reuters. (2022, January 21). Twitter suspends hundreds of accounts promoting Philippines' Marcos. *Reuters*. https://www.reuters.com/world/asia-pacific/twitter-suspends-hundreds-accounts-promoting-philippines-marcos-2022-01-21/

Salganik, M. J. (2018). Bit by bit. *Social research in the digital age*.

Sinpeng, A., Gueorguiev, D., & Arugay, A. A. (2020). Strong fans, weak campaigns: social media and Duterte in the 2016 Philippine election. *Journal of East Asian Studies*, *20*(3), 353-374. https://doi.org/10.1017/jea.2020.11

Soriano, C. R. R., & Gaw, F. (2022a). Platforms, alternative influence, and networked political brokerage on YouTube. *Convergence*, *28*(3), 781–803. https://doi.org/10.1177/13548565211029769

Soriano, C. R. R., & Gaw, M. F. (2022b). Broadcasting anti-media populism in the Philippines: YouTube influencers, networked political brokerage, and implications for governance. *Policy & Internet*, *14*, 508– 524. https://doi.org/10.1002/poi3.322

Tapsell, R. (2021). Social media and elections in Southeast Asia: The emergence of subversive, underground campaigning. *Asian Studies Review*, *45*(1), 117-134. https://doi.org/10.1080/10357823.2020.1841093

The YouTube Team (2021, November 10).An update to dislikes on YouTube. YouTube Official Blog. https://blog.youtube/news-and-events/update-to-youtube/

Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining* (Vol. 1, pp. 29-39). http://cs.unibo.it/~danilo.montesi/CBD/Beatriz/10.1.1.198.5133.pdf

About the Authors

JON BENEDIK BUNQUIN is an assistant professor at the University of the Philippines, Department of Communication Research. He is currently taking his PhD in Communication and Media Studies at the University of Oregon as a Fulbright scholar. His research focus lies at the intersection of science communication and political communication in new media environments.

FATIMA GAW is a PhD student in the Media, Technology and Society program at Northwestern University, Illinois, and a former Assistant Professor at the University of Philippines, Diliman. Her research centers on the mediation of platforms, algorithms and digital technologies in cultural production, politics, and public discourse in the Global South.