Tweets That Chill: Analyzing Online Violence Against Women in Politics

NDI
NATIONAL DEMOCRATIC INSTITUTE

© National Democratic Institute (NDI)
Website: www.ndi.org
Design and Printing: Dana Sulit, 2019

Copyright © National Democratic Institute for International Affairs (NDI) 2019. All rights reserved. Portions of this work may be reproduced and/or translated for non-commercial purposes with the prior written permission of NDI provided NDI is acknowledged as the source of the material and is sent copies of any translation. Send publication permission requests to legal@ndi.org.

Tweets That Chill: Analyzing Online Violence Against Women In Politics

Report of case study research in **Indonesia**, **Colombia**, and **Kenya**

Compiled and analyzed by: Kirsten Zeiter, Sandra Pepera, and Molly Middlehurst (NDI)

Technical lead: Dr. Derek Ruths (Charitable Analytics International)

ACKNOWLEDGMENTS

This report represents the commitment and collaboration of many people dedicated to promoting gender equality, women's political participation, and protecting online democratic spaces. The underlying project was developed by the National Democratic Institute (NDI) and implemented with the support of Dr. Derek Ruths, Chief Architect of Charitable Analytics International. Overall leadership and technical inputs were provided by NDI's Director for Gender, Women and Democracy (GWD), Sandra Pepera, and Kirsten Zeiter (GWD). Important contributions to the report writing and production were received from two other GWD team members, Molly Middlehurst and Tamar Eisen. Recognition is also due to colleagues on NDI teams in Washington DC, and in-country, who supported the case studies in Indonesia, Colombia, and Kenya, including Amanda Domingues.

This report has benefited from a committed production team, including our graphic designer, Dana Sulit.

Finally, the Institute gratefully acknowledges the support of the National Endowment for Democracy (NED) for providing the funding for this program.

ABOUT CHARITABLE ANALYTICS INTERNATIONAL

Charitable Analytics International (CAI) helps social good initiatives use data and analytics to achieve greater reach and greater impact. Since its launch in 2016, CAI has pursued this aim both through projects and training. Projects are concrete collaborations with an NGO or other social good organization in which CAI brings innovative data technology to help the partner realize greater impact. Through workshops, CAI shares best practices and shows how data technology can and should be a key tool in the mission to make the world a better place.

FOREWORD

The rise of the digital 'town square' is an undeniable force affecting democracies around the world. Citizens and governments alike are grappling with the impact of these technologies and platforms on political organizing and activism, which has the power to both embolden and silence voices in the digital realm. The long-standing debate about how to promote unity, connection, and conversations between and across diverse groups of citizens, and provide access to decision-making processes for many who have been traditionally excluded, is now being taken online.

The National Democratic Institute (NDI) has long served as a leader in supporting women around the world to overcome barriers to their political participation. Far too often, violence against women in politics (VAW-P), in all its manifestations, creates a 'chilling effect' that drives politically-active women offline and in some cases out of the political realm entirely.

In response to this growing challenge, NDI undertook this unique study to explore the effect of gender-based violence as expressed toward college-age, politically-active women on Twitter in three countries – Indonesia, Colombia, and Kenya. This first-of-its-kind study should serve as a tribute to women activists everywhere who are struggling to express their political views online.

We hope this study will set the groundwork for continued research on the impact of online violence toward politically-active women, and inform broader efforts to achieve more open and inclusive political environments around the world that promote and protect democratic values.

Derek Mitchell
President
National Democratic Institute

May 2019

TABLE OF CONTENTS

Executive Summary	4
Introduction	12
Methodology	13
Surveys	13
Workshops	13
Twitter Data Analysis	15
Key Findings	16
Prevalence of Online VAW-P	16
Type of Online Violence	18
Impact of Online Violence	19
Lessons Learned	22
Conclusion	24
Appendix	26
Endnotes	27

EXECUTIVE SUMMARY

English

As political discourse increasingly shifts online, a free, open, and inclusive internet where all citizens can engage in dialogue is critical to modern democracy. Online spaces can be particularly important in helping women and other marginalized groups to overcome barriers to becoming politically-active, thereby broadening the inclusive and participatory nature of democracy. However, the digital world has also become a forum for disinformation, hate speech, abuse, and harassment targeting politically-active women. This online violence against women in politics (VAW-P) creates a hostile environment which causes women to self-censor, be silent, or otherwise withdraw from online political discourse. It is one of the most prevalent forms of violence hindering women's active political engagement, but remains largely under-documented and under-researched.

When attacks against politically-active women are channeled online, the expansive reach of social media platforms magnifies the effects of psychological abuse by making those effects seem anonymous, borderless, and sustained, undermining women's sense of personal security in ways not experienced by men. Many of the state and non-state actors who perpetrate this online VAW-P are mobilizing across transnational networks. The misuse—by states, organizations, and individuals—of the very freedoms that the information space is supposed to enable, has become one of the greatest threats to its integrity. Not all forms of online VAW-P are crimes, but all impact on the human, civic, and political rights of women and girls.[1] In response to this new and growing challenge to the integrity of democracy, NDI undertook a program to understand and raise awareness about the nature and impact of online violence on women's political engagement and ambition. NDI intends that the findings from this program will inform changes to social media platform policies and national and global legal and political frameworks.

In Indonesia, Colombia, and Kenya, NDI worked with a number of women in politics, and civic technology and women's rights organizations to develop a way to examine the country- and context-specific challenges facing women as they engaged in online political discourse. The outcomes of these three country case studies enabled NDI to develop a robust typology of online VAW-P, identifying its dominant forms as **insults and hate speech, embarrassment and reputational risk, physical threats,** and **sexualized distortion**. Working with its partners in-country, NDI developed lexicons of both gender-based harassing language and the political language of the moment, in order to examine the online violence experienced by politically-active women. These lexicons, each developed in local languages (Bahasa for Indonesia, Colombian Spanish, and a mix of Swahili and English in Kenya), were then used to conduct data scraping[2] of a sample group of Twitter accounts within the target population of college-aged women and men who took part in the research. Twitter was selected as the social media platform for quantitative data analysis because the majority of Twitter interactions are public, thereby enabling NDI to analyze large data sets retrospectively. The timeframe for the Twitter scraping was set within a six-month window of a significant political event - an election, referendum, political scandal or crisis - in each country. This quantitative Twitter analysis was complemented by qualitative analysis of the workshop discussions and responses from surveys administered to the same populations.

NDI uncovered key findings from each country, and taken together, these results confirmed that online VAW-P had a "chilling effect" on politically-active women's social media engagement. Responses to online violence included pausing, decreasing, or completely stopping their Twitter activity. This finding, along with other results of the program, emphasizes the need for further studies to help understand and mitigate the impact of online violence on women's political participation, and provide the evidence for advocacy efforts. Future research, advocacy, and policy development must ensure that women are able to meaningfully participate in two key spaces - the political realm and the online world - and the areas where they intersect. Women's political engagement and the integrity of the information space must be protected so that inclusive, resilient, and sustainable democratic governance can be achieved.

EXECUTIVE SUMMARY

Bahasa

Ketika wacana politik semakin bergeser ke arah *online*, internet yang bebas, terbuka, dan inklusif yang memungkinkan keterlibatan semua warga negara dalam dialog menjadi hal yang sangat penting bagi demokrasi modern. Ruang *online* dapat menjadi sangat penting dalam membantu perempuan dan kelompok terpinggirkan lainnya untuk mengatasi hambatan guna menjadi aktif secara politik, dengan demikian menjadikan demokrasi semakin inklusif dan partisipatif. Namun, dunia digital juga telah menjadi wadah bagi disinformasi, ujaran kebencian, intimidasi, dan pelecehan yang menyasar para perempuan yang aktif secara politik. Kekerasan terhadap perempuan dalam politik *(violence against women in politics* atau VAW-P), yang dilakukan secara *online*, menciptakan lingkungan yang tidak bersahabat yang menyebabkan perempuan membatasi diri sendiri, diam, atau menarik diri dari wacana politik secara *online*. Ini adalah salah satu bentuk kekerasan yang paling jelas terlihat yang menghalangi keterlibatan perempuan secara aktif dalam politik. Sayangnya, dokumentasi dan penelitian dalam hal itu masih sangat kurang.

Ketika serangan terhadap perempuan yang aktif secara politik disalurkan secara *online*, jangkauan luas dari platform media sosial memperbesar efek dari pelecehan psikologis dengan membuat efek tersebut tampak anonim, tanpa batas, dan berkelanjutan, merusak rasa aman pribadi perempuan dengan cara-cara yang biasanya tidak dialami oleh laki-laki. Banyak dari aktor negara *(state actors)* and non-negara *(non-state actors)* yang melakukan VAW-P secara *online* bergerak dalam jaringan transnasional. Penyalahgunaan yang dilakukan oleh negara, organisasi, dan individu terhadap kebebasan yang semestinya dimungkinkan dengan adanya ruang informasi telah menjadi salah satu ancaman terbesar terhadap integritasnya. Tidak semua bentuk VAW-P *online* adalah kejahatan, tetapi semua berdampak terhadap hak asasi manusia, sipil, dan politik semua perempuan, dewasa maupun gadis remaja.[1] Menanggapi tantangan yang baru dan semakin berkembang terhadap integritas demokrasi ini, NDI menjalankan program untuk memahami dan meningkatkan kesadaran tentang sifat dan dampak kekerasan *online* terhadap keterlibatan dan aspirasi perempuan dalam politik. NDI menyatakan bahwa temuan dari program ini akan menginformasikan perubahan pada kebijakan platform media sosial serta kerangka hukum dan politik secara nasional maupun global.

Di Kolombia, Indonesia, dan Kenya, NDI bekerja sama dengan sejumlah perempuan di bidang politik, serta organisasi-organisasi hak perempuan dan teknologi sipil *(civil technology)* untuk mengembangkan suatu cara guna meneliti tantangan spesifik konteks dan negara yang dihadapi perempuan ketika mereka berusaha untuk terlibat dalam wacana politik secara *online*. Hasil dari studi kasus di ketiga negara tersebut memungkinkan NDI untuk mengembangkan tipologi yang baik terkait VAW-P *online*, mengidentifikasi bentuk-bentuk utamanya sebagai **penghinaan dan ujaran kebencian, risiko reputasi dan rasa-malu, ancaman fisik, dan distorsi seksual**. Bekerja dengan para mitranya di dalam negeri, NDI mengembangkan kamus bahasa pelecehan yang berbasis gender dan bahasa politik saat itu, guna menyelidiki kekerasan *online* yang dialami oleh para wanita yang aktif secara politik. Kamus ini, masing-masing dikembangkan dalam bahasa lokal (bahasa Indonesia, bahasa Spanyol Kolombia, serta campuran bahasa Swahili dan Inggris di Kenya), kemudian digunakan untuk melakukan ekstraksi data *(data scraping)*[2] sekumpulan sampel akun Twitter dari perempuan dan laki-laki usia kuliah yang berpartisipasi dalam penelitian. Twitter dipilih sebagai platform media sosial untuk analisis data kuantitatif karena sebagian besar interaksi Twitter bersifat publik, sehingga memungkinkan NDI

untuk menganalisis sekumpulan besar data secara retrospektif. Kerangka waktu untuk ekstraksi data dari Twitter ditetapkan dalam masa enam bulan dari peristiwa politik yang signifikan – pemilihan umum, referendum, krisis atau skandal politik – pada masing-masing negara. Analisis Twitter yang dilakukan secara kuantitatif ini dilengkapi dengan analisis kualitatif dari diskusi dalam lokakarya dan tanggapan dari survei yang dilakukan pada populasi yang sama.

NDI mengungkap temuan utama di tiap negara dan, bila digabungkan, hasilnya mengonfirmasikan bahwa VAW-P memiliki "efek mengerikan"[3] pada keterlibatan perempuan yang aktif secara politik dalam media sosial. Reaksi terhadap kekerasan *online* umpamanya mengurangi, menghentikan sejenak, atau benar-benar menghentikan aktivitas mereka di Twitter. Temuan ini, bersama dengan hasil lainnya dari program tersebut, menekankan perlunya studi lebih lanjut untuk membantu memahami dan mengurangi dampak dari intimidasi secara *online* terhadap partisipasi politik perempuan, dan memberikan bukti untuk upaya advokasi. Pengembangan kebijakan, advokasi, dan penelitian di masa mendatang harus memastikan bahwa perempuan dapat berpartisipasi secara bermakna dalam dua ruang utama - ranah politik dan dunia *online* - serta bidang-bidang yang bersinggungan dengan keduanya. Keterlibatan politik perempuan dan integritas ruang informasi harus dilindungi sehingga tata kelola demokratis yang inklusif, tangguh, dan berkelanjutan dapat tercapai.

EXECUTIVE SUMMARY

Spanish

A medida que el discurso político cambia cada vez más en línea, una internet libre, abierta e inclusiva donde todos los ciudadanos puedan participar en el diálogo es fundamental para la democracia moderna. Los espacios en línea pueden ser particularmente importantes para ayudar a las mujeres y otros grupos marginados a superar las barreras para volverse políticamente activos, ampliando así la naturaleza inclusiva y participativa de la democracia. Sin embargo, el mundo digital también se ha convertido en un foro para la desinformación, el discurso de odio, el abuso y el hostigamiento dirigido a mujeres políticamente activas. Esta violencia cibernético contra las mujeres en la política (VAW-P) crea un ambiente hostil que hace que las mujeres se autocensuren, guarden silencio o se retiren del discurso político cibernético. Es una de las formas más visibles de violencia que obstaculiza el compromiso político activo de las mujeres, pero sigue estando poco documentada y poco investigada.

Cuando los ataques contra mujeres políticamente activas se canalizan en línea, el alcance expansivo de las plataformas de las redes sociales magnifica los efectos del abuso psicológico al hacer que esos efectos parezcan anónimos, sin fronteras y sostenidos, socavando el sentido de seguridad personal de las mujeres en formas no experimentadas por los hombres. Muchos de los actores estatales y no estatales que perpetran este VAW-P en línea se están movilizando a través de redes transnacionales. El uso indebido, por parte de estados, organizaciones e individuos, de las mismas libertades que se supone que el espacio de información debe permitir, se ha convertido en una de las mayores amenazas para su integridad. No todas las formas de VAW-P en línea son delitos, pero todas tienen un impacto en los derechos humanos, cívicos y políticos de mujeres y niñas.[1]

En respuesta a este nuevo y creciente desafío a la integridad de la democracia, el NDI emprendió un programa para comprender y crear conciencia sobre la naturaleza y el impacto de la violencia en línea en el compromiso político y la ambición de las mujeres. NDI indica que los resultados de este programa informarán los cambios en las políticas de la plataforma de medios sociales y los marcos legales y políticos nacionales y mundiales.

En Indonesia, Colombia, y Kenia, el NDI trabajó con varias mujeres en la política y con organizaciones de tecnología cívica y de derechos de las mujeres para desarrollar una manera de examinar los desafíos específicos de cada país y contexto que enfrentan las mujeres mientras intentaban participar en el discurso político en línea. Los resultados de estos tres estudios de caso de país permitieron a NDI desarrollar una tipología sólida de VAW-P en línea, identificando sus formas dominantes como insultos y discursos de odio, vergüenza y riesgo de reputación, amenazas físicas y distorsión sexual. Trabajando con sus socios en el país, NDI desarrolló léxicos tanto del lenguaje de acoso basado en el género como del lenguaje político del momento, para examinar la violencia en línea experimentada por las mujeres políticamente activas. Estos léxicos, cada uno de ellos desarrollado en idiomas locales (bahasa para Indonesia, español colombiano y una combinación de swahili e inglés en Kenia), se utilizaron para llevar a cabo el datos raspado[2] de un grupo de muestra de cuentas de Twitter dentro de la población objetivo de mujeres y hombres en edad universitaria que participaron en la investigación. Se seleccionó a Twitter como la plataforma de redes sociales para el análisis de datos cuantitativos porque la mayoría de las interacciones de Twitter son públicas, lo que permite a NDI analizar grandes conjuntos de datos de forma retrospectiva. El plazo para el rastreo de Twitter se estableció dentro de un período de seis meses de un evento político importante - una elección, un referéndum, un escándalo político o una crisis - en cada país. La población objetivo para

el análisis era mujeres y hombres jóvenes en edad universitaria. Este análisis cuantitativo de Twitter se complementó con un análisis cualitativo de las discusiones del taller y respuestas de encuestas administradas a las mismas poblaciones.

El NDI descubrió hallazgos clave de cada país, y en conjunto, estos resultados confirmaron que el VAW-P en línea tuvo un "efecto escalofriante" en el compromiso de las mujeres con las redes sociales políticamente activas. Las respuestas a la violencia en línea incluyen pausar, disminuir o detener la actividad de Twitter. Este hallazgo, junto con otros resultados del programa, enfatiza la necesidad de estudios adicionales para ayudar a comprender y mitigar el impacto del acoso en línea en la participación política de las mujeres, y proporcionar la evidencia para los esfuerzos de promoción. La investigación futura, la promoción y el desarrollo de políticas deben garantizar que las mujeres puedan participar de manera significativa en dos espacios clave: el ámbito político y el mundo en línea - y las áreas donde se cruzan. El compromiso político de las mujeres y la integridad del espacio de información deben protegerse para que se pueda lograr una gobernabilidad democrática inclusiva, resistente y sostenible.

EXECUTIVE SUMMARY

Swahili

Majukwaa ya kisiasa yanavyozidi kuhamia mtandaoni, mtandao wa bure, ulio wazi na shirikishi ambapo raia wote wanaweza kushiriki katika mijadala ni jambo mtambuka kwa demokrasia ya kisasa. Uwanja huu wa siasa za mtandaoni unaweza kuwa nguzo muhimu sana katika kuhakikisha wanawake na makundi mengine yaliyotengwa kuondokana na vikwazo vya kuwa mstari wa mbele kisiasa, na hivyo kukuza hali ya ushiriki na ushirikishwaji katika demokrasia jumuishi. Hata hivyo, ulimwengu huu wa mtandao pia umekuwa ni jukwaa litumikalo kutoa habari potofu, maneno ya chuki na unyanyasaji dhidi ya wanawake walioko mstari wa mbele katika harakati za kisiasa. Mashambulizi haya ya mtandaoni dhidi ya wanawake hutengeneza mazingira yasiyorafiki kwa wanawake na kupelekea wanawake kuwa kimya, au kujiondoa kabisa katika mijadala ya kisiasa. Mbali na kuwa aina mojawapo ya ukatili dhidi ya wanawake inayorudisha nyuma ushiriki wao katika siasa, ukatili huu bado haujafanyiwa tafiti za kutosha na hivyo kupelekea taarifa zake kuwa hafifu.

Wakati mashambulizi dhidi ya wanawake walioko mstari wa mbele kisiasa yakielekezwa mtandaoni huku mawasiliano kupitia mitandao ya kijamii ikiendelea kupanuka kwa kasi, unyanyasaji huu umebaki kuwa wa kisaikolojia kwa kufanya madhara yake yaonekane hayana mashiko, hayana mipaka, lakini ni maswala endelevu na hivyo kuwafanya wanawake wajihisi hawako salama tofauti na ilivyo kwa wanaume. Wahusika wengi walioko serikalini na wasiokuwa serikalini wanaofanya hasama dhidi ya wanawake mitandaoni wameendelea kushamiri kupitia mitandao mbalimbali. Kutokana na matumizi mabaya ya tunu ya uhuru wa habari mitandaoni, nchi, mashirika na watu binafsi wameifanya njia hii ya habari kutokuaminika. Japo kuwa sio kila aina ya mashambulizi haya mitandaoni ni makosa kisheria, bado yote kwa pamoja yanachangia kuathiri haki za kibinadamu, kiraia na za kisiasa kwa wanawake na wasichana.

Ili kukabiliana na changamoto hii mpya inayokua kwa kasi na kuhatarisha ukuaji wa demokrasia ya kisiasa, NDI ilianzisha mpango wa kutambua na kuongeza ufahamu juu ya asili na athari za mashambulizi ya mtandaoni na jinsi yanavyoathiri ushiriki na utayari wa wanawake katika shughuli za kisiasa. NDI inaamini kwamba matokeo ya mpango huu yatapelekea mabadiliko ya sera ya jukwaa la mitandao ya kijamii na mabadiliko ya kisheria katika taratibu za uendeshaji wa siasa kitaifa na kimataifa.

Nchi za Indonesia, Colombia na Kenya, ambapo NDI imefanya kazi na baadhi ya wanawake katika maswala ya siasa na utumiaji wa teknolojia katika siasa kupitia mashirika ya kutetea haki za wanawake kwa kuandaa njia za utafiti wa nchi hizo ikiwemo changamoto zinazo wakabili wanawake kwenye mazingira hayo wakiwa wanajihusisha na mijadala ya siasa kupitia mitandao ya kijamii. Matokeo ya tafiti hizi kwenye nchi tatu zimewezesha shirika la kimataifa la NDI kuandaa usanifu thabiti wa ukatili wa wanawake katika siasa unaofanyika kupitia mitandao ya kijamii, na kuweza kutambua aina zake kuu ambazo ni matusi na hotuba za chuki, uzalilishwaji na hatari ya kuchafuliwa hadhi, vitisho vya kimwili na kupotoshwa kingono. Kwa kushirikiana na wadau mbalimbali nchini, NDI imeweza kuandaa majumuisho ya maneno mbalimbali (Lexicons) wa aina za lugha za unyanyasaji wa kijinsia na lugha ya kisiasa itumikavyo kwa sasa, ili kutathmini ukatili unaowakabili wanawake wanaojihusisha na siasa mitandaoni. Majumuisho haya ya maneno, yameandaliwa katika lugha rasmi za nchi husika (Lugha ya Bahasa kwa Indonesia, Kihispania cha Kikolombia, na mchanganyiko wa Kiswahili na Kiingereza kwa Kenya), yakatumika kufanya uhakiki wa akounti za twitter ndani ya kundi lengwa la vijana wa kike na wakuime walioko vyuoni

walioshiriki katika utafiti. Twitter ilichaguliwa kama mtandao wa kijamii wa kutathimini viwango vya taarifa kwa sababu mawasiliano mengi katika mtandao huu ni ya wazi kwa jamii, hivyo kuiwezesha NDI kutathimini taarifa nyingi kwa muda na hata za wakati uliopita.

Muda wa miezi sita wa uhakiki wa twitter, ulitengwa ndani ya kipindi muhimu kisiasa-ikiwemo uchaguzi, kura za maoni, kashfa za kisiasa au migogoro kwa kila nchi. Uchambuzi huu wa Twitter, umeongezwa nguvu na tathmini ya semina ya majadiliano na maoni yaliyokusanywa kutoka tafiti zilizoendeshwa kwa watu husika.

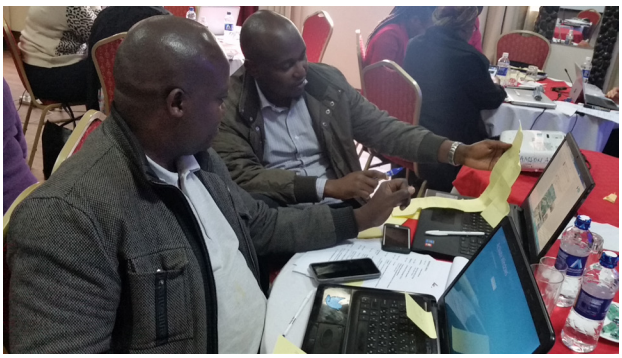
Katika tafiti zilizofanyika kwa kila nchi nakuwekwa pamoja, NDI iliweza kuthibitisha kwa matoekeo kwamba kuna hali ya kukandamiza wanawake wanaojihusisha na siasa kupitia mitandao ya kijamii. Matokeo ya ukatili unaofanyika husababisha wao kusimamisha, kupunguza harakati zao mitandaoni au vyote kwa pamoja na kuachana kabisa na twitter. utafiti pia uliweza kuambatanisha matokeo ya miradi mingine ambayo inahamasisha uhitaji wa kuongeza ujuzi wa kuweza kukabiliana na madhara yatokanayo na unyanyasaji wa wanawake wanaojihusisha na siasa kupitia mitandao ya kijamii na kupata taarifa za harakati hizo. Tafiti, harakati na uboreshaji wa sera zinapaswa kuhakikisha wanawake wanaweza kushiriki kikamilifu kwenye maeneo mawili; itikadi za kisiasa na uhuru mitandaoni. Hivyo, ushirikishwaji wa wanawake katika siasa na uhuru wa majukwaa ya mawasiliano unabidi kulindwa ili kuifikia demokrasia endelevu.

INTRODUCTION

As political discourse increasingly shifts online, a free, open, and inclusive internet where all citizens can engage in dialogue is critical to modern democracy. Online platforms are a space where citizens gather information and voice their opinions, and can be particularly important in helping women overcome barriers to becoming politically-active. However, in too many places, citizens' ability to engage in political discourse online is under threat from the prevalence of online violence, whether perpetrated by individuals or organizations. Depending on its type, this online violence causes politically-active citizens - both women and men - to withdraw from public debate. The particular focus of this report is the growth in online activity by individuals and organizations that specifically seek to silence or exclude the voices of politically-active women. This activity can be anonymous, borderless, sustained, and permanent. The perception of impunity emboldens perpetrators and raises women's sense of insecurity and violation. The resulting limitation of both the number of women able to participate in political discourse and the range of issues discussed poses fundamental challenges to democracy, progress towards gender equality and women's empowerment, as well as to the integrity of the information space.

Online violence against politically-active women is one form of the global problem of violence against women in politics (VAW-P), which encompasses all forms of aggression, coercion, and intimidation of politically-active women simply because they are women.[3] Online VAW-P cuts across all sectors of politics - civil society, political parties, elections, and governance - posing one of the main challenges to women's full and equal participation in public life. In this way, online communication technologies have a significant impact on the reach and shape of violence against women in politics, creating new threats and obstacles to achieving gender equality in politics and decision-making. Overall, this type of violence and abuse creates a hostile online environment with the aim of shaming, intimidating, or degrading women as individuals and as a group. Not all forms of online violence are crimes, but all impact on women's human, civic, and political rights. Online VAW-P represents a direct barrier to women's free speech, undermining democracy in all its key elements, and with negative effects on the ambitions of young women and new entrants to politics.

In response to this new and growing challenge, NDI, with its technical partner, Dr. Derek Ruths of Charitable Analytics International (CAI), undertook a program to examine the nature and impact of online VAW-P on politically-active women by conducting case studies in three countries where NDI works. The project employed Twitter data analysis and a survey about social media conversations among college-aged young women and men, in Indonesia, Colombia, and Kenya. The goal of the program was to understand and raise awareness about the nature and impact of online violence on women's political engagement and ambition, with the ultimate aim of promoting changes to social media platform policies, and to national, regional, and global legal and political frameworks. This report presents the findings of the survey and Twitter data analysis from that program.



Workshop participants in Kenya, June 2018

METHODOLOGY

As part of its program of work on VAW-P, NDI developed a methodology to examine the nature and impact of online VAW-P, focusing specifically on examining whether it has a chilling impact on the political engagement of young women. The analysis employed a mixed methods approach comprising: (1) workshops involving in-country civil society organizations engaged with women's rights and the digital space; (2) surveys with female and male college-aged young people in each country about their social media practice and experiences; and (3) quantitative Twitter data analysis of selected accounts. NDI worked with local field offices and civil society organizations in three case-study countries - Indonesia, Colombia, and Kenya. The target population for the quantitative Twitter data was college-aged women given their higher usage of social media to share and discuss political issues, as compared to older age groups.

SURVEYS

Alongside the development of the typology and local lexicons through the in-country workshops, NDI conducted surveys with college-aged young people on university campuses in each of the case study countries. The survey was made available in paper format and in an online Google Form, and the sample size for the questionnaires in each country was 1,000, collected from a target of at least two universities. In all cases, efforts were made to achieve an approximate gender balance of responses between young women and young men. The surveys gathered information on participants' social media usage, and their experiences with online violence, both on Twitter and other social media platforms. The surveys focused particularly on how the participants individually responded to online violence, and whether and how this abuse impacted their online engagement. The surveys were also used to identify specific Twitter handles of college-aged, politically-active women in each country, to be used in the Twitter data analysis portion of the study. NDI piloted the survey in Indonesia with professional survey enumerators to collect 1,000 surveys, from approximately 600 women and 400 men. Based on feedback from this pilot, the survey was refined and finalized for the Colombia and Kenya case studies.

WORKSHOPS

In each country, NDI and its partners ran three-day workshops to gain a deeper understanding of the local issues surrounding online violence against women in general, before focusing on its particular manifestation in politics. Workshop activities included awareness raising and capacity-building sessions, as well as participant-led exercises that explored the lived experience of online violence in that country's context. Importantly, the workshops developed lexicons of words and phrases in local languages that both characterized gender-based online violence targeted at women, as well as reflected the language and idioms of the political moment. During the workshops, participants and

NDI tested the relevance and accuracy of these lexicons by using them to search Twitter in real time, to identify example Tweets, and to categorize those Tweets under various headings.

In order to reduce the possibility of false-positives in the analysis of online VAW-P, participants at each workshop applied several lenses to examine and better understand the context and use of the words and phrases within the lexicons. First, the words were gathered in two categories: gender-based violent language and political language. For the political language, NDI and partners sought to determine whether the language was nationally bounded or more widely used within the region. For the gender-based violent language, workshop participants were asked to consider whether the words always had an inherent negative connotation and/or whether any of these words had been "reclaimed" by the intended target group or individual. For example, the Kenya workshop determined that the Swahili term "momo," which is defined as a woman who nags, is always used in a negative context. However, the term "boss lady" was originally used in a derogatory context, but has recently been reclaimed by women (the intended targets of this insult) as a term of pride.

These unique workshops allowed NDI to reflect the local expressions of gendered linguistic and political context, in order to develop a robust typology of online VAW-P. This typology evolved from NDI's original typology of violence against women in politics as presented in the 2017 *#NotTheCost: Violence Against Women in Politics Program Guidance.*[4] It is important to note that there is a clear link – known as the gateway effect – between persistent and aggressive harassment online and the actual physical assaults that some women face.[5] However, for this research, in order to reflect the virtual character of the digital space, what was visible in the physical world had to be adapted to what could be examined and tracked in the online world. The table below maps NDI's original violence against women in politics typology (left) with that of online VAW-P (right). In addition to developing local lexicons according to this typology, the participatory nature of the workshops also had three immediate impacts in each country: (1) an increased understanding among the participants of the nature and anti-democratic impact of online VAW-P; (2) enhanced capacity of civic tech and women's rights partners in coding social media language in order to measure and monitor online VAW-P in contextually- and linguistically-specific ways; and, (3) the creation of in-country cross-sectoral forums for collaboration on internet-related issues including online violence, policy, and governance.

MAPPING TYPOLOGIES

VAW-P TYPOLOGY	ONLINE VAW-P TYPOLOGY
Psychological	Insults and hate speech
Psychological / Threats & coercion / Economic	Embarrassment and reputational risk
Physical	Physical threats
Sexual	Sexualized distortion

TWITTER DATA ANALYSIS

The typology and lexicons developed during the in-country workshops (discussed above) were then utilized to conduct a robust analysis of online violence using data from Twitter. The Twitter data analysis focused on Twitter activity in a six-month window surrounding a significant political event in each country: Indonesia's June 2018 local elections; Colombia's March 2018 parliamentary election; and the general election in Kenya in August 2017. In this analysis, NDI used tweeting frequency as a measure of online engagement, and tested whether there was a relationship between receiving higher amounts of violence and a person's level of engagement on Twitter.

The lexicons developed in the local languages (Bahasa, Spanish, and English/Swahili respectively) were then used by NDI's technical lead, Dr. Derek Ruths, to develop a method for identifying and measuring violent tweets online. This method consisted of two stages: (1) identifying potentially violent tweets targeting women in the country; and (2) manually coding those tweets to determine the kind of online violence. In the first stage, the team used a program to gather potentially violent tweets that explicitly mentioned a Twitter account belonging to a local college-aged woman. To identify potentially violent content, the program used the lexicon of common local language violent terms developed during the workshop. To identify accounts belonging to local college-aged women, NDI engaged local partners to code Twitter accounts that followed popular in-country Twitter accounts (e.g., news sites, university Twitter feeds).

In the second stage, the accounts and the potentially violent tweets from the first stage were coded by volunteers from the in-country workshop. To code the accounts, the volunteers looked at each account and indicated: (1) whether the account belonged to a person from the case-study country; (2) whether the account likely belonged to a woman; and (3) whether the account likely belonged to a college-aged person. Once the accounts were identified, in each case study country the team gathered all Twitter activity related to those accounts within the 6-month time window, and these became the sample of Tweets for analysis. To code the Tweets, the volunteers looked at each tweet and indicated: (1) whether the tweet was actually violent; and (2) what kind of violence it was (using the online VAW-P typology). This human coding was determined to be a necessary step because, although the computer algorithm could get a list of potentially violent tweets based on the lexicon, human coders are required to analyze the content and determine whether it matched the criteria of a violent tweet according to the methodology and parameters of the study. Having collected these violent tweets, a final data collection step involved gathering the tweets (via Twitter scraping) of the targeted users within the six-month timeframe of the study. In the subsequent analysis, NDI used the tweeting frequency of a user to measure online engagement, and tested whether there was a relationship between receiving higher amounts of online violence and a person's level of engagement on Twitter.

Scraping is a mechanism for extracting a large amount of data from online platforms by downloading a web page's content and using a tool to organize this large quantity of data in a digestible format for further analysis.

KEY FINDINGS

The study offered several key findings regarding the prevalence, type, and impact of online violence on politically-active women. This analysis offers insights into the problem of online VAW-P, and also highlights enduring methodological challenges to making reliable and accurate measurements of online violence.

PREVALENCE OF ONLINE VAW-P

Core considerations of the study were: (1) the prevalence of the broad phenomenon of online violence against politically-active women; and (2) the impact of this violence on their online political engagement. Results from the survey[6] and the Twitter study[7] provide two different ways of examining these issues.

Self-reported levels of online violence were higher

Figure 1 shows that self-reported and observed levels of online violence differed substantially, with self-reported levels (17-50%) being higher than observed (5-8%) levels. An important consideration that may explain this discrepancy is that the survey and the Twitter study measured different notions of online VAW-P. In the case of the survey, online VAW-P was defined however the respondent chose to interpret it when responding to the question. Individuals may have differing and broad definitions of online violence based on their lived experience of participating on social platforms including anything that made the individual feel threatened, devalued, or hurt. On the other hand, in the Twitter data analysis, NDI used its typology and lexicons of online VAW-P, to examine overtly threatening, demeaning, or hurtful language as identified and defined by workshop participants. It is possible there were many violent tweets that did not use these overt signals and therefore slipped past the detection methods used in this Twitter analysis.[8]

COUNTRY	% OF FEMALE TWITTER USERS WHO EXPERIENCED ONLINE VAW-P ACCORDING TO TWITTER ANALYSIS	% OF WOMEN REPORTING VIA SURVEY THEY HAD EXPERIENCED ONLINE VAW-P
Indonesia	5.5%	17.6%
Kenya	3.6%	22.7%
Colombia	8.3%	50.2%

Figure 1: Frequency of online violence experienced by politically-active women. The first column represents the observed findings from Twitter data and the second column represents the self-reported survey data.

Another factor that might explain the divergence between the self-reported survey and the Twitter data analytics concerns the **experience of indirect versus direct online** violence. It is possible that, in the survey, individuals reported incidences of online VAW-P in which they witnessed gender-based violence of other people that, nonetheless, made them personally feel devalued, threatened, or hurt. This is consistent with NDI's understanding of all forms of VAW-P which have an impact beyond their intended victim:

"While acts of violence against women in politics are directed at individual women, they have an intent beyond their specific target: to frighten other women who are already politically-active, to deter women who might consider engaging in politics, and to communicate to society that women should not participate in public life in any capacity."

NOT THE COST CALL TO ACTION, 2016[9]

However, due to practical issues with sampling, and concerns over the ability to define and track indirect effects of online violence, the Twitter study did not measure indirect violence levels.

High levels of per-person online violence in Indonesia

As evidenced in Figure 2, NDI observed that Twitter users in Indonesia, on average, experienced higher levels of online VAW-P than users sampled in other countries. One possible explanation for this trend, which arose from the workshops and conversations with partners, is that there are stronger 'pile-on' dynamics in Indonesia. In the Twitter context, the pile-on effect occurs when a user experiences violence on the platform and other users quickly join in, either through amplifying the original violence through "re-tweets" or compositions of their own. This is a particularly prevalent occurrence if the accounts are held by celebrities, who by definition have larger followings. The pile-on effect was brought up by participants in all three of the country workshops, however understanding the extent to which this happens in Indonesia could be an important direction for future work on understanding country-specific online violence practices.

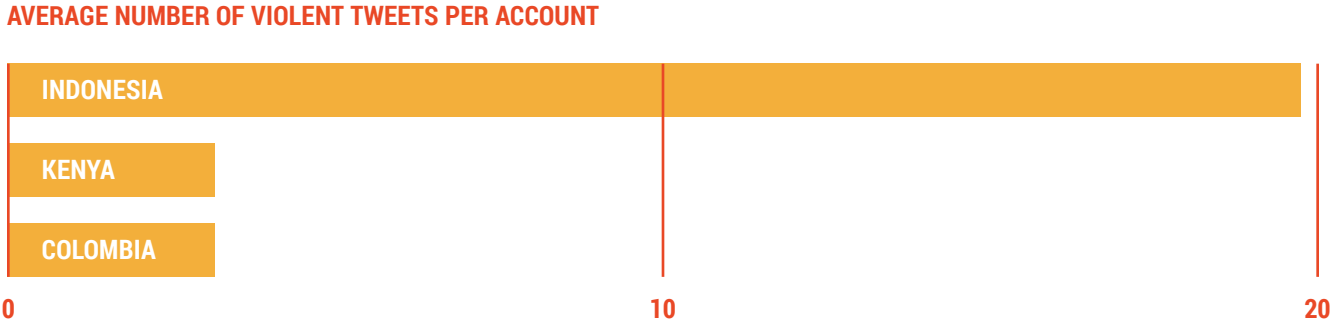


Figure 2: Frequency of online violence as measured by the average number of tweets experienced by individual users across the three case study countries.

TYPE OF ONLINE VIOLENCE

Comparing the Twitter and survey data across the countries, NDI gained three key insights on the frequency of the different types of online violence.

Agreement between survey responses and Twitter data on the frequency of types of online violence

Comparing the relative online violence rates within a given country, NDI found that, overall, the trends agree – e.g., in Indonesia, both the survey and Twitter analysis found that insult-based online violence was the most frequent type and that threat-based was the least frequent type. This is an important point of validation. The above section on prevalence illustrated that the frequency of online VAW-P was different from country to country. However, the consistent trends between Twitter and survey data within countries about the **type** of online violence suggests that this study is, in fact, measuring similar violent phenomena.

DATA ANALYTICS

TYPES	INDONESIA	KENYA	COLOMBIA
Insult	72%	71%	96%
Physical threat	0.3%	2.5%	2.5%
Reputational threat	48%	66%	11%
Sexual	11%	40%	4.9%

Figure 3: The box above reflects results from the data analysis, the one below from survey responses.

COMPARED TO SURVEY RESPONSES

TYPES	INDONESIA	KENYA	COLOMBIA
Insult	23%	50%	41%
Physical threat	4%	23%	11%
Reputational threat	16%	24%	19%
Sexual	10%	40%	57%

Frequency of insults vs. other types of online violence

The Twitter data and, to a lesser extent, the survey data both show that online violence via insult was the most common kind of online VAW-P experienced by women across all three case studies. This was an unexpected finding, as insulting content can be among the hardest to detect compared to blatantly violent or threatening content. This finding underscores the need for further attention to detecting and handling more nuanced insulting content if platforms are to effectively measure and eventually flag or remove violent content.

Additionally, the workshop in Indonesia revealed that reputational online violence - online violence meant to damage or discredit the moral reputation of a woman in her community - was prominent, as was the use of words or phrases with religious relevance. However, in the Indonesian case, sexual-based online violence was comparatively less common.

Sexual online violence in Colombia

While the survey and Twitter data on the frequency of sexual online violence were internally consistent within each of the country case studies, and similar between the Indonesian and Kenyan cases, the Colombian case merited further investigation. There was a gap between the rates of sexual online violence from the self-reported survey responses (57%) and observed Twitter data (4.9%). This could point to a disconnect between the lexicon that was used to identify sexually violent tweets. However, given that the lexicon was developed by a diverse panel of Colombians drawn from across the sectors of civil society, this seems unlikely. The issues raised in the previous section of indirect violence and the need to develop lexicons that can detect more than blatantly violent content, may also at least partially explain the results in Colombia too. This trend deserves further consideration in future studies.

IMPACT OF ONLINE VIOLENCE

One of the driving questions behind this study, which has not been addressed elsewhere in the recent literature on online VAW-P, is whether violent content has an effect on women's online engagement. NDI's research utilizes the self-reported survey and observed Twitter data to best examine how politically-active women responded to the online violence they experienced. Among possible responses in the survey, participants could indicate that they ignored, stopped, paused, or increased their social media engagement after a violent Tweet or other post. Stopping or pausing online social media activity (through decreasing the overall number of Tweets/posts, taking a break from social media engagement, or abandoning a social

media presence altogether) indicate what NDI has defined as overt chilling effects. NDI analyzed the responses to determine whether or not individual incidents of online VAW-P had a "chilling effect" on women's social media engagement.

Using the Twitter data analysis and in order to determine an individual's response to instances of online VAW-P on Twitter, the team looked at whether each targeted user decreased the number of tweets they were producing immediately after the violence occurred. The team measured this by subtracting the tweeting rate before the violent event from the tweeting rate right after the violent event. The more negative the number, the less the user tweeted after the violent event – this is shown in the second column of Figure 4. To know whether this drop in the rate of tweets per second was unusual, the team also looked at the amount a user's tweets might drop on any given day. This is given in the first column of Figure 4. By comparing the two, NDI could determine when the violent tweets (on average) had a chilling effect. When the drop due to online violence was greater than the "normal" drop in tweeting on any given day, the violent tweets likely had a chilling effect. As illustrated in Figure 4, both Kenya and Colombia had larger changes in response to online VAW-P, which illustrates the phenomenon of politically-active women slowing down, pausing, or altogether stopping their social media presence after experiencing online VAW-P. It is important to note here that there are other factors that could account for a decrease in a user's rate of online engagement – for example, the timing of the violence might come at a time when the user was likely to tweet less regardless of the violence (e.g. at the beginning or end of a weekend, at the tail-end of an election). Further research from NDI and others would provide essential data to control for these factors.

COUNTRY	AVERAGE "NORMAL" DECREASE IN RATE OF TWEETS PER SECOND	AVERAGE DECREASE IN RATE OF TWEETS PER SECOND IMMEDIATELY FOLLOWING VIOLENT TWEET
Indonesia	-0.33	-0.09
Kenya	-0.07	-0.22
Colombia	-0.09	-0.18

Figure 4: Twitter users' change in online engagement before and after a violent tweet.

This finding was echoed in the survey responses. In Indonesia (32%) and Kenya (40%), the most common response to online VAW-P was to ignore the online violence. However, 20% of survey respondents in Kenya paused their social media activity in response. This is a clear example of the "chilling effect" as described above. Across all three case studies, it was rare (2-4%) for respondents to completely stop engaging online in response to online violence.

RESPONSE	INDONESIA	KENYA	COLOMBIA
Ignored	32%	40%	28%
Responded	9%	17%	28%
Stopped	2%	4%	3%
Paused	6%	20%	10%
Increased	1%	0%	38%

Figure 5: Responses to the survey question on how students responded to online violence.

By comparing the Twitter data against the survey responses, NDI found that both sets of data provided strong evidence that in Kenya and Colombia, online violence decreased politically-active women's willingness to continue engaging in social media. In contrast, politically-active women in Indonesia seemed less inclined to decrease their social media activity in response to online violence. Fewer women in the Indonesian case reported stopping (2%) or pausing (6%) posting online in response to online VAW-P.

This finding is aligned with feedback from participants at the in-country workshops. While workshop participants in all three countries expressed the view that they expected and accepted online VAW-P as part of the online culture, participants in Indonesia were particularly emphatic and vocal about this point. It is possible that women in Indonesia were so accustomed to pervasive online violence that it did not affect their own social media use. However, further study is needed to fully test and contextualize this theory.

A workshop participant in Indonesia, February 2018

LESSONS LEARNED

Through its mixed method approach and the ability to analyze data from the three case studies to compare perceptions of online VAW-P and social media engagement, with observations of politically-active women's responses to online violence, NDI learned several lessons to inform and contextualize further work.

Lesson 1 Contextually- and linguistically-specific lexicons of online violence must be created and then evolve

Across all three case study countries, workshop participants highlighted the fluid and evolving nature of language and brainstormed ways to account for this nuance in the study methodology. For example, NDI learned from the Colombia workshop that violent language in Spanish varied across Latin America, with both Colombia-specific and words from other parts of the region being used within the country. In Indonesia, religious words or phrases were used, complicating and heightening the online violence by invoking religious messages at the same time. In Kenya, workshop participants noted that a number of violent words/phrases that were in common usage in spoken Swahili, had not yet made it into written text online on Twitter. These varied lessons point to the need for contextually- and linguistically-specific lexicons that can be continuously refreshed, modified, and implemented with human coders working alongside computer algorithms.

Lesson 2 Attention to minority communities and intersecting identities is essential

Online VAW-P is varied and contextual, as it differs from country to country and culture to culture. However, it is also the case that the expressions used and impacts of online violence can vary significantly between and among communities within the same country. For this reason, it is important to intentionally include and consider historically marginalized communities among women (e.g. women with disabilities, LGBTI women, and female members of religious and ethnic minorities) when exploring the phenomenon of online VAW-P. During the Colombia workshop, female representatives from the deaf community shared that the violence they faced was not in text, but through the uploading of violent GIFs and/or video clips in sign-language. It was explained that this delivery mechanism was particularly effective in conveying threat and insecurity because, for the majority of the members of the deaf community in Colombia, sign language is their first language, and the targeting was therefore unmistakable. Understanding that the kinds of threats and modes of online violence can differ substantially when targeting different marginalized communities indicates that further work is required to create relevant lexicons.

Lesson 3 If women's rights initiatives are action- and solution-oriented, otherwise fatigued partners are eager and enthusiastic to engage

During the planning phases of the research, there was an initial concern about the level of engagement and interest by local partners in workshops. Moreover, many participants mentioned their initial wariness in participating in "yet another workshop on women's issues." Their comments expressed a fatigue from participating in workshops, research, and other programs that they felt failed to lead to concrete action or change in their communities. However, once it was effectively communicated that workshops were focused on developing useful and actionable ways of measuring and monitoring online violence, a range of partners and actors from civil society expressed enthusiasm to participate. Participants were particularly receptive to exercises that generated concrete tools and resources (e.g., the lexicons of violent and political language, and guides on coding best practice). Future studies on online VAW-P may, therefore, want to ensure partners' buy-in through a productive focus on results in the framing and anticipated outputs of in-country engagement.

Lesson 4 Under-reporting of VAW-P in online spaces exists and merits investigation

As widely recognized in the literature on gender-based violence and VAW-P,[10] instances of online violence and violence against women on social media platforms often go unreported. For example, in this study, gaps between the qualitative workshop discussions and the quantitative data showed that online violence was not widely recognized as such by Indonesian participants. However, this is likely not unique to Indonesia. Researchers and activists should be careful not to take a lack of reported (including self-reported through qualitative and quantitative research methods) online violence as an indication that online violence is not occurring.

NDI's work in uncovering and understanding trends in VAW-P as a global phenomenon has created an evidence-basis for the prevalence of online violence against politically-active women. However, both online and offline violence, regardless of its modality and method of expression, is often normalized within cultures or expected as "the cost of doing politics."[11] It is likely that as presented in NDI's initial Call to Action to End Violence Against Women, politically-active women using digital spaces are still cautioned against speaking out against online VAW-P, and stay silent rather than risk being labeled unreliable to their political colleagues or leaders.

CONCLUSION

While some research exists on the prevalence of online violence experienced by young women,[12] this report has been among the first to identify and quantify the direct impacts of this on young women's political discourse online. The results of this study provide compelling evidence of the direct and varied impacts of online VAW-P as a direct barrier to women's free speech, and its chilling effect on women's ambition to be politically-active. Specifically, the data shows that in both Kenya and Colombia, instances of online violence during political discourse were followed by a decrease in female users' Twitter engagement. However, further research is needed to explore and contextualize this and other findings.

For example, future studies could examine why politically-active women in Indonesia did not exhibit the same aggregate decreases in engagement in response to online violence. While qualitative findings from workshop discussions highlighted cultural factors that might account for heightened resilience, it is important for future work to consider that the generalized level of online violence against all women in a country, may be normalized to the degree that online violence against politically-active women continues to be under-reported and therefore difficult to identify and examine.

This study sets out methods that can be adopted by other studies aimed at understanding the relationship between online VAW-P and women's political participation online. The study shows how a combination of participatory and context-driven qualitative and quantitative methods can provide a more thorough understanding of the nuances and impacts of online VAW-P. NDI's in-country workshops enabled a diverse range of local civic technology partners and women's rights organizations to contribute to the development of contextually-specific lexicons of violence in local languages. These workshops also provided a forum for the capacity-enhancement of these partners to code social media language in order to measure and monitor online VAW-P. In each country, participants commented that the workshop provided ideas for ways in which they could develop mechanisms to apply the methodology and concepts to their own work to tackle the persistent challenge of online VAW-P. The workshops also fostered in-country cross-sectoral forums for collaboration on internet-related issues including online violence, policy, and governance.

The study utilized survey data and large-scale collection, annotation, and analysis of Twitter data as important points for triangulating the workshop findings. Both the survey and Twitter analysis found that, according to NDI's typology of online VAW-P, insulting language was the most common kind of online violence experienced by women across all three case studies. While the frequency of online VAW-P varied from country to country, the consistent trends within the data about the type of online violence suggests that this study has, in fact, measured similar violent phenomena experienced by women engaging in politics online across all three countries. The overall findings of the research underscore the need for further attention to detecting and handling more nuanced violent content if platforms are to effectively

measure and eventually flag or remove violent content. Ultimately, NDI's highly participatory and locally-driven mixed-methods approach provided a powerful way of investigating the intrinsically hard-to-study phenomenon of online violence.

NDI's goal is that this study will lay the foundation for action against VAW-P in the many forms it appears as a suppressor of women's voices in formal and informal political spaces. The design and lessons from this study can and should be used to inform and inspire similar studies. Further research on this topic will collectively help understand and eventually mitigate the impact of online violence and women's political participation. Moreover, further study will provide the international community with the necessary evidence upon which to base advocacy efforts, including those aimed at digital platforms on the need to ensure that women can continue to be politically-active online, safely and meaningfully. Ultimately, regulation, advocacy, policy change, and the active engagement and participation of politically-active women and civil society stakeholders will be necessary to ensure progress towards gender equality and women's empowerment, and to the integrity and inclusivity of the information space.

Workshop participants in Colombia, April 2018

ANNEX

Case Study Partner Organizations/ Workshop Participants

Indonesia
Qlue Performa Indonesia
Pulse Lab Jakarta
Kalyanamitra
Jurnal Perempuan
Koalisi Perempuan Indonesia

Colombia
Asociacion de Victimas Jovenes Colombianos
Fundación Karisma
Ministerio del Interior
Conferencia Nacional de Organizaciones Afrocolombianas
Fundación Escuelas de Paz
Fundación Grupo de Acción y Apoyo a Personas Trans - G.A.A.T
Fundación Esperanza Afro
Redepaz
Mediox
Federación Colombiana de Periodistas

Kenya
Internet Society of Kenya
Nailab
Ushahidi
Safaricom
Saisa Place
Association of Media Women in Kenya (AMWIK)
Kenya Women Holdings
African Women Child Features
Health Care Assistance Kenya (HAK)
FIDA Kenya
Centre for Rights Education and Awareness (CREAW)
Youth Agenda
Twaweza Communications
Media Focus
Women Empowerment Link
Coalition on Violence Against Women
IHUB
KICTANet
Election Observation Group (ELOG)
Article 19
Development Through Media
Country Governance Forum

ENDNOTES

[1] Shiromi Pinto, "What is Online Violence and Abuse Against Women," *Amnesty International*, November 20, 2017, https://www.amnesty.org/en/latest/campaigns/2017/11/what-is-online-violence-and-abuse-against-women/.

[2] Scraping is a mechanism for extracting a large amount of data from online platforms by downloading a web page's content and using a tool to organize this large quantity of data in a digestible format for further analysis.

[3] NDI, "Not the Cost: Program Guidance for Stopping Violence Against Women in Politics," *NDI*, 2017, p. 12, https://www.ndi.org/sites/default/files/not-the-cost-program-guidance-final.pdf and;

NDI, "Not the Cost: Submission by the National Democratic Institute to the United Nations Special Rapporteur on Violence Against Women," NDI, 2018, p. 5, https://www.ndi.org/sites/default/files/NDI%20Submission.pdf.

[4] Ibid, p. 18.

[5] Ibid, p. 15.

[6] Survey data is referred to in this document as *self-reported* as it relies on participants discussing their own feelings, attitudes, and experiences.

[7] Twitter data is referred to in this document as observed as it is drawn from snapshots of actual Twitter activity which was then analyzed by NDI and local coders.

[8] Roya Pakzad and Niloufar Salehi, "Anti-Muslim Americans: Computational Propaganda in the United States," *Institute for the Future*, 2019, p. 12.

[9] NDI, "Not the Cost: A Call To Action for Stopping Violence Against Women," *NDI*, 2016, p. 13.

[10] See:

Amnesty International, "Toxic Twitter: A Toxic Place for Women," *Amnesty International*, 2018, ch. 4, https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-4/#topanchor.

Atalanta, "(Anti)Social Media: The benefits and pitfalls for digital for female politicians," *Atalanta*, 2018, p. 44, https://www.atalanta.co/news/2018/3/13/antisocial-media-the-benefits-and-pitfalls-of-digital-for-female-politicians.

Julie Ballington, "Preventing Violence Against Women in Elections: a programming guide," *UN Women*, 2017, http://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2017/

preventing-vaw-in-elections-en.pdf?la=en&vs=3742.

NDI, "No Party to Violence: Analyzing Violence Against Women in Political Parties," *NDI*, 2019. https://www.ndi.org/sites/default/files/2019%20VAW-PP%20Compendium%20Report.pdf.

[11] NDI, "Not the Cost: Program Guidance for Stopping Violence Against Women in Politics," *NDI*, 2017, p. 13, https://www.ndi.org/sites/default/files/not-the-cost-program-guidance-final.pdf.

[12] See:

Maeve Duggan, "Online Harassment 2017," *Pew Research Center*, 2017. https://www.pewinternet.org/2017/07/11/online-harassment-2017/.

Amnesty International, "Toxic Twitter: A Toxic Place for Women," *Amnesty International*, 2018, https://www.amnesty.org/en/latest/research/2018/03/online-violence-against-women-chapter-4/#topanchor.

Jan Moolman, "Global Information Society Watch 2013: Women's Rights, Gender and ICTs," *Association of Progressive Communications*, 2013, https://www.apc.org/sites/default/files/gisw13_chapters_0.pdf.

About the National Democratic Institute

The National Democratic Institute for International Affairs (NDI) is a nonprofit, nonpartisan, nongovernmental organization that responds to the aspirations of people around the world to live in democratic societies that recognize and promote human rights. Since its founding in 1983, NDI and its local partners have worked to support and strengthen democratic institutions and practices by strengthening political parties, civic organizations and parliaments, safeguarding elections and promoting citizen participation, openness and accountability in government.

With three decades of experience in 132 countries, NDI is the leading organization working to advance women's political participation around the world. The Institute supports women's aspirations for inclusive and responsive government, by ensuring that they are able to participate, compete and lead as equal and active partners in democratic change whether as citizen activists, voters, political party workers, candidates or decision-makers.

NDI
NATIONAL
DEMOCRATIC
INSTITUTE

@NDIWomen | www.ndi.org