USAID
FROM THE AMERICAN PEOPLE

CEPPS
Strengthening Democracy through Partnership

Violence Against Women in Elections Online: A Social Media Analysis Tool

September 2019

USAID
FROM THE AMERICAN PEOPLE

International Foundation for Electoral Systems

NDI

Violence Against Women in Elections Online: A Social Media Analysis Tool

Attention: CEPPS Administrative Director
Consortium for Elections and Political Process Strengthening
1225 Eye Street NW, Suite 800 Washington, DC 20005
jlavery@cepps.org

Disclaimer: This publication was made possible through the support provided by the United States Agency for International Development (USAID). The opinions expressed hererin are those of the authors and do not necessarily reflect the views of USAID.

Table of Contents

Part I. Understanding Election Violence & Gender in Online Spaces

Around the world, women frequently experience harassment and violence when they decide to exercise their civil and political rights. Information and communication technologies (ICTs) have created new vehicles for violence against women in elections (VAWIE), including violence that takes place on social media and in private messaging. These forms of violence are compounded by the anonymity and scale that online media platforms provide. VAWIE-Online is a means to silence women who publicly engage in political life through fear, shame and intimidation. This violence is different from the online violence and harassment experienced by politically active men in its underlying intent, its multiplied impact, as well as its frequency, form, and content.

The **VAWIE-Online Social Media Analysis Tool (VAWIE-Online Tool)** offers an adaptable method to measure the gendered aspects and understand the drivers of online election violence against women. The primary purpose of this tool is to identify trends and patterns of online violence around electoral periods. More specifically, the tool will allow users to identify the scope, breadth, and intensity of VAWIE-Online. Using artificial intelligence-based data analysis tools, the VAWIE-Online Tool quantifies and categorizes social media data to identify and distinguish forms of online violence. By incorporating this analysis of online violence into broader analysis of gender in elections, electoral stakeholders can better understand this issue in their respective countries and can begin to address it through their work. This guide is intended for use by civil society organizations (CSOs), election and human rights monitors and observers, and other activist and research groups seeking to analyze online violence against women in elections. It is constructed as a step-by-step tool to introduce users to social media analysis, specifically data mining and sentiment analysis. Although this tool is written to provide an introduction and overview for general users, data mining and sentiment analysis are sophisticated research approaches and users of this guide will find it helpful to work with a team that has experience in data analytics and a background in working on gender-based violence (GBV) and electoral politics.
In this guide, the term "implementer" is used to refer to the organization conducting the analysis (CSO, observer group, etc.) and the term "analyst(s)" refers to the person or people who are directly conducting data analysis. For the purposes of this tool and drawing on the International Center for Research on Women's[1] definition of technology-facilitated GBV, VAWIE-Online is defined here as one or more people using the internet and/or mobile technology in a way that violates an individual or a group's civil and political rights based on sexual or gender identity or by enforcing harmful gender norms.

This tool was piloted through five activities in Sri Lanka, Ukraine and Zimbabwe in 2018-19.[2] Examples, ideas, and lessons from these pilots are included throughout. In each country, local groups used the tool

[1] Laura Hinson et al., "Technology-Facilitated Gender-Based Violence: What Is It, and How Do We Measure It?" (Washington D.C.: International Center for Research on Women, 2018), https://www.icrw.org/wp-content/uploads/2018/07/ICRW_TFGBVMarketing_Brief_v8-Web.pdf.

[2] The Sri Lanka pilot was funded through the CEPPS Technical Leadership award. The Ukraine and Zimbabwe pilots were funded through cooperation with programs funded by USAID, DFID and Global Affairs Canada.

to look at big picture data spanning several months or years. In Ukraine and Zimbabwe, the tool was used for additional activities to monitor VAWIE-Online in real time during national elections.

What is VAWIE?

Political violence during elections and democratic processes is a common occurrence in many countries, as is GBV. VAWIE exists at the crossroads of political violence and GBV, targeting women who participate in public or political life specifically because they are women and often in distinctly gendered ways. Until recently, political and electoral violence have been viewed as gender-neutral concepts. However, these forms of violence are, in fact, highly gendered in their motives, forms, and impacts. Gender norms shape how and why women are subject to electoral violence, as well as what types of acts are carried out to curtail or influence their participation.[3] Such violence can be employed specifically to uphold gender norms or traditional female roles in society and may be motivated by a desire to repress, deter, control, or otherwise coerce the political rights of the survivors.[4] This violates women's political rights on the basis of their gender identity, threatening the integrity of the electoral process. Like all forms of GBV, VAWIE is rooted in power imbalances and social norms that structure the status of women around the world.[5] For this reason, VAWIE is an important topic for researchers and practitioners working from both GBV-advocacy perspectives and from election, conflict, and security perspectives.

Violence against women in politics (VAWP) and in elections is a violation of civil, political, and human rights. VAWIE affects women's civic and political right to participate as voters, candidates, election officials, activists, and political party leaders. It undermines free, fair, and inclusive democratic processes. Recognizing that it can occur both online and offline, the United Nations Office of the High Commission on Human Rights asserts that violence against women in politics (including in elections) "consists of any act of gender-based violence, or threat of such acts, that results in, or is likely to result in, physical, sexual or psychological harm or suffering and is directed against a woman in politics because she is a woman, or affects women disproportionately."[6]

---

[3] Caroline Hubbard and Claire DeSoi, "Votes without Violence: A Citizen Observer's Guide to Addressing Violence against Women in Elections" (Washington D.C.: National Democratic Institute, 2016), https://www.ndi.org/sites/default/files/Votes_Without_Violence_Manual.pdf.

[4] Gabrielle S. Bardall, "Violence, Politics, and Gender," Oxford Research Encyclopedia of Politics, February 26, 2018, https://doi.org/10.1093/acrefore/9780190228637.013.208.

[5] Caroline Hubbard and Claire DeSoi, "Votes without Violence: A Citizen Observer's Guide to Addressing Violence against Women in Elections" (Washington D.C.: National Democratic Institute, 2016), https://www.ndi.org/sites/default/files/Votes_Without_Violence_Manual.pdf.

[6] "Report of the Special Rapporteur on Violence against Women, Its Causes and Consequences on Violence against Women in Politics" (United Nations General Assembly, August 6, 2018), https://www.un.org/en/ga/search/view_doc.asp?symbol=A/73/301.

Consider This: Political Violence vs. Violence Against Women in Politics

Different researchers and practitioners have different theories, priorities and frameworks when studying violence against women online. These conceptual choices affect research findings and intervention strategies. Ultimately, the decision is up to each user of this guide. "Consider This" textboxes examine some of these perspectives.

Political violence is a generalized phenomenon that can affect men and women alike, regardless of their country or position. There is a global consensus that all violence against women is rooted in structural gender inequality, socio-cultural norms, and disempowerment.[7]. Because this type of violence aims to deter women from participating in the political process, it a breach of women's human, civil, and political rights. And because its aim is to exclude women's equal and active participation, it presents a serious barrier to gender equality and undermines democracy. As opposed to political violence that both women and men experience, violence against politically active women is a different and a specific issue because it targets women because they are women, it is gendered in its form, and is gendered in its impact.[8] This definition is aligned with that of the UN Special Rapporteur on Violence against Women (SRVAW) as stated in her report to the UN General Assembly in October 2018[9]:

> *Men and women can both experience violence in politics. Such acts of violence against women, however, target them because of their gender and take gender-based forms, such as sexist threats or sexual harassment and violence. Their aim is to discourage women from being politically active and exercising their human rights and to influence, restrict or prevent the political participation of individual women and women as a group.*
>
> *Such violence, including in and beyond elections, consists of any act of gender-based violence, or threat of such acts, that results in, or is likely to result in, physical, sexual or psychological harm or suffering and is directed against a woman in politics because she is a woman, or affects women disproportionately.*

Many practitioners and academic researchers have decided to take this approach and base their methodology on a focus on violence against politically active women as a unique phenomenon, rather than as an aspect of more general political violence.

---

[7] "Declaration on the Elimination of Violence Against Women" (United Nations General Assembly, December 20, 1993), https://www.un.org/documents/ga/res/48/a48r104.htm. ; "#NotTheCost: Stopping Violence Against Women in Politics: Call to Action," The National Democratic Institute, March 2016, https://www.ndi.org/not-the-cost. ; Jessica Huber and Lisa Kammerud, "Violence Against Women in Elections: A Framework for Assessment, Monitoring, and Response" (The International Foundation for Electoral Systems, September 2016), https://www.ifes.org/sites/default/files/vawie_framework.pdf.

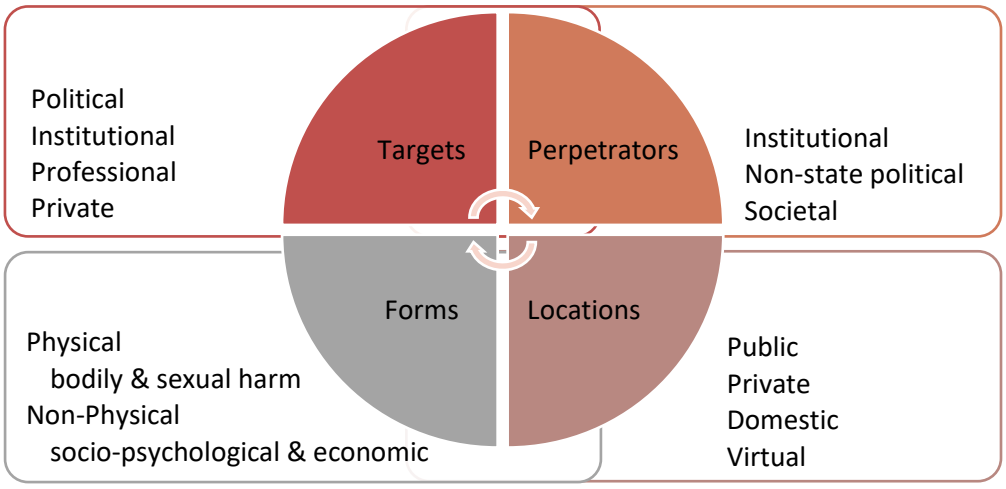
[8] Caroline Hubbard and Claire DeSoi, "Votes without Violence: A Citizen Observer's Guide to Addressing Violence against Women in Elections" (Washington D.C.: National Democratic Institute, 2016), https://www.ndi.org/sites/default/files/Votes_Without_Violence_Manual.pdf.; Gabrielle Bardall, Elin Bjarnegard, and Jennifer Piscopo, "Gender and Political Violence: Motives, Forms, and Impacts," (2017), http://mlkrook.org/pdf/Bardall.pptx.

[9] "Report of the Special Rapporteur on Violence against Women, Its Causes and Consequences on Violence against Women in Politics" (United Nations General Assembly, August 6, 2018), https://www.un.org/en/ga/search/view_doc.asp?symbol=A/73/301.

Election violence is gendered in multiple ways, including by types, survivors, perpetrators, and locations;[10] these distinctions are presented in the table below. Recognizing the many dimensions of election violence is essential to understanding the gendered dimension of this violence. Historically, assumptions and stereotypes have limited understanding of election violence to mostly male experiences (i.e., public acts of physical violence between competing party actors).

A gendered understanding recognizes that election violence happens in many spaces (including in private spaces, within domestic relationships and online—where women most frequently experience it). This gendered interpretation also recognizes that many different actors may be involved, including non-traditional perpetrators (such as family members) and survivors beyond partisan actors (particularly journalists and voters). Just as this detailed typology adds a gender lens to election violence, it also aligns election violence with international definitions of GBV.[11]

Collectively, these constitute a typology of VAWIE off and online.



[10] Election violence is distinguished by location. However, this methodology is only concerned with election violence taking place in cyber/digital spaces.
[11] UN OHCHR, "Violence against Women," United Nations Human Rights: Office of the High Commissioner. https://www.ohchr.org/EN/Issues/Women/Pages/VaW.aspx.

Important Terms & Concepts

VAWIE-Online is an umbrella term that captures a broad range of abusive, harassing, degrading and violent discourse circulating on the internet or mobile technology across a range of intensities, from sexist slurs to direct threats of physical harm.

What is VAWIE-Online?

VAWIE-Online is frequently associated with one or more of the following:

Hate speech vilifies, humiliates, or incites hatred against a group or a class of persons based on a protected attribute such as the target's sexual orientation, gender, religion, disability, color, or country of origin**.**

Disinformation is false or misleading information that is created or disseminated with the intent to cause harm or to benefit the perpetrator. The intent to cause harm may be directed toward individuals, groups, institutions, or processes

Malinformation is accurate information that is shared with the intent to cause harm or to benefit the perpetrator, often by moving private information into the public sphere.

Misinformation is false or misleading information that is shared without the intent to cause harm or realization that it is incorrect. In some cases, actors may unknowingly perpetuate the spread of disinformation by sharing content they believe to be accurate among their networks. (CEPPS/IFES 2019)

Online hate is defined as any online expression, encouragement, stirring up or incitement of hatred. (Barker and Jurasz 2018)

It is important to note that terminology in this area is still developing and not univocal (OHCHR/SRVAW A/HRC/38/47). Terms frequently used include "information and communications technology" (or ICT), "online violence," "digital violence" or "cyberviolence." Reflecting the SRVAW approach, this resource uses VAWIE-Online as a user-friendly, inclusive expression. Digital or cyberspaces may be viewed as both a location of violence as well as a facilitating tool for violence.

VAWIE-Online is a non-physical form of violence as it takes place in a virtual space. However, physical threats are very common and the intense fear and intimidation that takes place in online spaces can have very tangible, material impacts including on the survivor's physical and mental health and economic resources.

One conceptualization of VAWIE-Online reflects four forms of violence bodily harm, sexual, socio-psychological, and economic). For example, an Instagram post threatening a candidate with rape and death would be a case of both physical and sexual intimidation. Online messages that maliciously attack a person's intelligence, morality, integrity, or body image to shame, intimidate or degrade a person represent socio-psychological forms of VAWIE. Similarly, many messages threaten to rupture the social

fabric of the survivor's life, for example by suggesting the survivor should be ostracized by their religious community, rejected by their families, divorced, or even exiled from their community (thus the term socio-psychological). Economic violence also exists online, however it generally occurs in private online spaces (e.g., control of access to online bank accounts) rather than on social media.

In addition to these examples, there are many forms of violence that are specific to online spaces. Any of these forms of violence can become forms of VAWIE when they impede women's electoral participation by coercing, limiting or terminating the free exercise of their civil and political rights. These forms of violence are classified within the VAWIE framework as follows:

Forms of VAWIE-ONLINE

Categories	***Physical***		***Non-Physical***	
	Bodily Harm	**Sexual**	**Socio-Psychological**	**Economic**
Online threats pertaining to:	• Murder / attempted murder • Physical assault and injury • Battery • Maiming • Wounding	• Rape / attempted rape • Sexual assault • Intimate partner sexual assault	• Intimidation • Threats to individual or individual's family • Verbal harassment • Shaming defamation	• Denial / constant threat of denial of resources / services • Unlawful control and monitoring of the use and distribution of monies and access to services (healthcare, employment, etc.)
Cyber-specific terms	• IRL attacks ("In Real Life") • Swatting[12] • Trafficking	• Cyber-exploitation • Nonconsensual photography or leaking nudes or other	• Cross platform harassment • Deadnaming[15] • Defamation • Doxing[16] • False accusations of blasphemy	• Distributed denial of service (DDoS) • Electronically enabled financial abuse • Identity theft and online impersonation

[12] Swatting is classified here as physical harm, as swatters generally seek to elicit a law enforcement response that might include the use of weapons and possibility of being killed or hurt. Where such physical involvement does not occur, swatting may be classified as socio-psychological harm.

[15] Revealing a person's former name or gender identity. Used particularly to target lesbian, gay, bisexual, transgender, queer/questioning, intersex, asexual/ally plus (LGBTQIA+) individuals.

[16] Doxing refers to the practice of publicly sharing private information about an individual. This can also include posting personal details about an otherwise-anonymous online user.

	• Cyberstalking and stalking by proxy[13]	intimate photos[14] • The creation of pornographic deep fakes • Photoshopped sexualized images		

Why is VAWIE important in online spaces?[17]

Women frequently cite the threat of widespread, rapid public attacks on personal dignity as a factor deterring them from entering politics.[18] Although it may comprise physical, sexual, or economic acts of aggression, electoral violence most often takes the form of psychological attacks. Socio-psychological violence is by far the most pervasive form of electoral violence experienced by women and the most widespread form of online violence. Indeed, in sample data, the proportion of intimidation and psychological acts of violence experienced by women was nearly three times the same proportion among men (a ratio of 28:10).[19] Psychological violence is an "informal means of control [and] includes systematic ridicule, ostracism, shame, sarcasm, criticism, disapproval, exclusion and discrimination."[20] Coupled with threats of physical and sexual violence, these forms of violence degrade, demoralize, and shame the individuals at which they are targeted. Psychological forms of violence are frequently orchestrated through the instruments of social media.

> *"[T]he Internet is being used in a broader environment of widespread and systemic structural discrimination and gender-based violence against women and girls, which frame their access to and use of the Internet and other ICT. Emerging forms of ICT have facilitated new types of gender-based violence and gender inequality in access to technologies, which hinder women's and girls' full enjoyment of their human rights and their ability to achieve gender equality."*
>
> Special Rapporteur on Violence Against Women and Girls, A/HRC/38/47

[13] Cyberstalking is considered more dangerous than other forms of cyberbullying because it generally involves a credible threat to the survivor's safety. Where there is no physical threat, it may be classified as socio-psychological violence.
[14] "Revenge porn" includes the posting of explicit images of an individual without their consent. Former partners are regularly the perpetrators of this form of violence.
[17] The full text of this section is drawn from Bardall (2017) in Vitis & Segrave and used here with permission.
[18] Authors' interviews.
[19] Bardall , Gabrielle. "Breaking the Mold: Understanding Gender and Electoral Violence," *IFES*, December 13, 2011, https://www.ifes.org/publications/breaking-mold-understanding-gender-and-electoral-violence
[20] Ibid., 8

Online violence and abuse against women creates a hostile environment with the aim of shaming, intimidating or degrading women. Not all forms are crimes, but all impact the human rights of women[21] In a recent poll commissioned by Amnesty International in eight countries, nearly a quarter of women surveyed had experienced online abuse or harassment. A report released by the United Nations Broadband Commission called violence against women online a "problem of pandemic proportion." The report found that 73 percent of women online have been exposed to or experienced some type of cyberviolence.[22] Among the 86 countries included in the survey for the report, only 26 percent of law enforcement agencies have taken action against such violence. ICTs —especially social media—are frequently used as tools of gender-specific electoral and political violence. There is overwhelming, global evidence of ICTs being used to perpetrate a broad range of violent acts against women during elections and in public life, especially acts that inflict fear and psychological harm.

ICTs may be used directly as a tool of intimidation by threatening or inciting physical violence against women candidates, voters, or representatives. Such cyber harassment or intimidation includes sending abusive, threatening or obscene emails, explicit threats of physical and/or sexual violence and encouraging strangers to physically harm the survivor, which in some cases results in actual physical assault. Acts of VAWIE-Online may also involve spreading reputation-harming lies, electronic sabotage in the form of extensive spam and damaging viruses, impersonating the survivor online and sending abusive emails or fraudulent spam, blog posts, tweets, and other online communications in the survivor's name or subscribing survivors to unwanted email lists, resulting in hundreds of unwanted messages daily. Such attacks can be perpetrated by both strangers and individuals known to the survivor, as well as by proxy stalkers and "cyber-mobs."

Five Ways Social Media Amplifies VAWIE-Online

1. Nature of messaging facilitates psychological violence
2. Images contribute to attacks on morality through sexualization
3. Speed of information and scope of diffusion magnify impact
4. Difficult to prevent or provide preventative justice
5. Legal and moral impunity

Certain qualities of social media make these technologies uniquely suited to inflicting psychological violence on women in the course of their exercising of civic and political rights. In particular, morality-based attacks (e.g., accusations of prostitution, homosexuality, failed parental duty, etc.) often carry much greater social costs for women than for men because of the existence of double standards around what constitutes moral behavior for male and female politicians. Social media amplifies these imbalances and exacerbates attacks on women in public life in several ways.

[21] Pinto, Shiromi." What is online violence and abuse against women?" *Amnesty International*, https://www.amnesty.org/en/latest/campaigns/2017/11/what-is-online-violence-and-abuse-against-women/
[22] "Combatting Online Violence Against Women & Girls: A Worldwide Wake-up Call*," Broadband Commission Working Group on Gender*, September 2015, https://en.unesco.org/sites/default/files/highlightdocumentenglish.pdf

First, social media is well adapted to amplify the types of violence that women most often experience in electoral participation. Research shows that socio-psychological violence is the most widespread and damaging form of violence against women in politics.[23] The nature of messaging in social media facilitates ridicule, shaming, and other psychological forms of violence against women in elections and in politics. Social media creates fora where broad audiences can engage in rumor-mongering, misogynistic shaming, abuse, threats, and intimidation through multiple communication channels (re-sharing/re-tweeting, "liking," commenting, cross-platform posting, etc.).

Social media also facilitates attacks on women's ethics and morality through the ubiquitous presence of images and, increasingly, videos. The use of photoshopped, stereotypical or demeaning images and photos to sexualize, emotionalize, and trivialize women poses a strong disincentive for women considering running for office and may even pose a direct threat to their personal safety. Third, the nearly uncontrollable speed with which information travels through social media networks and the scope of its diffusion magnify the violent impact. Fourth, available redress for this type of attack—including community censure, website moderation, and legal intervention—frequently takes effect only after the damage to the survivor has been done. Finally, violence perpetrated through social media benefits from a significant degree of legal and moral impunity. Technology companies and governments may struggle with the tension between stopping hate speech and promoting freedom of speech or healthy democratic debate. Combined with the perceived—and often actual—anonymity of users, this framing complicates prosecution and emboldens perpetrators.

Understanding Ecosystems of VAWIE-Online

In tracking and measuring VAWIE-Online, it is important to situate the analysis and understand the potential and the limitations of analysis.

Cyberspace is multidimensional. Individuals share content both publicly and privately, retaining different degrees of control over content and what is said about them. Commentary and reporting on the internet appear in both more "formal" spaces such as recognized commercial media and web-published journalism as well as less regulated public fora such as chatrooms, Twitter, YouTube, Instagram, blogs and commentary spaces on media sites. All of these different forms of media draw distinct audiences; they have different purposes and, as such, have different impacts.

Each area of virtual space poses different risks regarding the presence of VAWIE and requires different approaches to deal with VAWIE. VAWIE-Online can exist across all these different spaces and is rarely limited to any single online space. Instead, it exists across an ecosystem of online spaces. For example, a woman targeted by VAWIE-Online might receive several threatening or degrading tweets, followed by a threatening text message to her phone and to her Viber account. Manipulated images of her might next

[23] Gabrielle Bardall, "Breaking the Mold: Understanding Gender and Electoral Violence," White Paper, White Paper Series (The International Foundation for Electoral Systems, December 2011), https://www.ifes.org/sites/default/files/gender_and_electoral_violence_2011.pdf.; Gabrielle Bardall, "Voices, Votes and Violence: Essays on Selected Dynamics of Electoral Authoritarianism" (PhD Dissertation, Université de Montréal, 2016), https://papyrus.bib.umontreal.ca/xmlui/handle/1866/18513.

appear on YouTube or Facebook and a blogger might publish her street address and phone number. Other internet users may find this information and decide to show up at her physical address to threaten her or her family. In cases like this, which are common, many different cyber platforms and tools are used to perpetrate VAWIE-Online against a single individual.

While these categories pertain to all forms of online abuse, the electoral context adds some particular dimensions. Electoral politics raises the profile of individuals competing in or speaking-out about politics. Cross-platform targeting can become more intense. With many more individuals engaged in a single act of harassment than might be experienced by a private individual, greater extremes are more likely to emerge. These extremes may include enflaming anger to more acute levels, attracting fringe individuals who may carry over violence into real world spaces, or creating more "noise" around any given VAWIE-Online incident as people from all parts of the opinion spectrum weigh in. It is therefore important to understand and classify these different online spaces. One way of thinking about this distribution of virtual spaces is presented in the table below:

		Content Access	
		Private	**Public**
Content Management	**Private**	Private conversations between individual users. Access and management limited to direct users (Facebook Private Messages, WhatsApp, Skype, Telegram, etc.)	Private user makes public posts on another user's social media profile or on their own profile (e.g., making a public wall post on a Facebook friend's Facebook wall, or posting a status on one's on Facebook profile, or posting a tweet from one's own Twitter account)
	Public	Formally published public media and journalism (commercial media, journalistic sites, greater degree of regulation)	Weakly regulated or unregulated public fora (Twitter, Instagram, YouTube, blogs, other forms of "chatter")

The **VAWIE-Online Tool** only looks at part of this ecosystem; specifically, it only looks at information available in online public spaces. Identifying violence and harassment in online public spaces is very important, because it often represents some of the most extreme positions (including highly partisan rhetoric), which prompt very damaging attacks and terrorizes targets. Violence and harassment expressed in public spaces often occurs at very high volumes compared to violence in private cyberspaces (for example, hundreds of thousands of individual Twitter and public Facebook posts can contain VAWIE-Online content during any given election period). High volume means that individual

incidents cannot be recorded manually but require computer analytics to be identified and recorded, which is what data mining and analysis enable us to do.

Data scraping and analysis software only analyzes content in the public domain. Because users have greater control over content that is posted in social media spaces they manage (for example, user posts to the public Facebook page of a candidate), there is an inherent skew in the data. Targets of abusive behavior delete or report such posts quickly, after which point the data cannot be retrieved. Thus, the present study captures primarily publicly available information (either from formal media sources or social media platforms) and some residual private posts that are make public that are not self-censured. It does not capture any private data.

However, users of the VAWIE-Online Tool should always remember that public data is usually only the tip of the VAWIE-Online iceberg. Data analysis can give us a clear but limited sample of what is happening. It is the first port-of-call for research on VAWIE online because public domain data can be collected with protection for the privacy rights and security of survivors. The tool is a useful instrument, but it is also an incomplete representation of the full spectrum of VAWIE-Online. It does not capture violence happening in online private spaces and even the most robust analysis will only capture a part of existing violence, as we will discuss below.

What are data mining and sentiment analysis? Why we are using them to track VAWIE-Online?

Data mining refers to the process in which an analyst takes a large amount of raw data and filters it to find patterns. The VAWIE-Online Tool mines social media data. In social media data mining, anything produced on the internet in a given period can be considered raw data. The analyst first **scrapes** the data from its source, such as Twitter or Facebook. Scraping takes information from webpages and forums and extracts information so that it can be easily compared and categorized. It would be logistically impossible to go on every parliamentarian's public Facebook page and read every post. Scraping allows us to take web information and pull the text in a way that is easier to analyze. For example, a Twitter scrape can take the message text, date and time, re-tweet count, username, and geographic location of a series of tweets and place them into a single spreadsheet. This spreadsheet would contain a more manageable and searchable series of posts and words of which we can then make sense.

Technology Changes Quickly

This tool is designed to help understand the principles and strategy to analyze VAWIE-Online. It is not a how-to guide for any specific software package and it is not limited to analyzing violence happening only on the social media that exist today. Because software changes so quickly, understanding the principles will allow this resource to adapt with changing technology.

From this data, the implementer will mine, or filter, for particular individuals and words. Unlike simply searching the web using a search engine, social media data mining filters millions of tweets, Facebook posts, blog posts, and news stories using phrases. If the implementer or organization has existing data, such as parliamentary records, this raw data can also be mined for useful phrases and words.

Social media data mining takes vast amounts of online data and filters it by particular words and subjects or by factors such as date, user, and popularity. The implementer can identify some or all of the following information: when a phrase, word, or person is discussed more or less, some information about the author of a post, how many times a post has been shared, and the overall sentiment of the post.

Sentiment analysis is a tool utilized to classify emotion or sentiment. It was initially developed as a means of classifying large amounts of information about a product or brand. A company can utilize sentiment analysis to quickly classify a series of comments or reviews into positive, negative, and neutral.

Over the last two decades, as more and more people around the world have been able to access the internet, sentiment analysis has expanded as well. Businesses remain interested in understanding the public's reception of their product. But others, such as political consultants and public relations specialists, have utilized sentiment analysis to uncover trends in attitudes. In addition to more traditional tools such as polling, sentiment analysis can take large amounts of online data and monitor positive, neutral, and negative feelings about a public figure or a policy. Sentiment analysis takes mined data, such as all online posts that mention the name of a leading political official in the current election period and classifies the data by sentiment.

Sentiment analysis can identify the public's reaction to a candidate or public official. Negative sentiment describes online posts or articles that are, overall, negative toward an individual. Among negative

sentiments, these posts may convey disgust, fear, anger, or sadness. Positive sentiment, by contrast, describes online posts or articles that are positive, or convey happiness, approval, or support. Neutral sentiment does not clearly convey either positive or negative emotions. Positive and negative sentiment does not entirely predict whether a post is violent. For example, "I'm so happy the candidate was finally punched and kicked, as she deserved" would be classified as positive by automated sentiment analysis systems, as it expresses happiness, even though the content of the post conveys violence.
The VAWIE-Online Tool can utilize sentiment analysis, but only as one method in a much broader analysis. Because sentiment analysis only looks at emotion alone—and not at violence explicitly—it is not useful to understand trends of abuse. As such, the tool proposes a broader analysis**,** in which it is possible to utilize sentiment analysis as a way to narrow down a sample or learn about the emotions involved in a post. Implementers should complement sentiment analysis by also using keyword categorization and machine learning in order to obtain a more nuanced understanding of the data.

Social Media Analysis of VAWIE-Online helps answer these questions:

- Does online violence and harassment impact women disproportionately?
- Do these gendered differences change depending on the form of violence (bodily harm, socio-psychological, sexual)?
- Do forms of violence change over time?
- What do these forms tell us about the drivers and causes of gendered political violence?
- Does online GBV increase as the election gets closer or does it ebb and flow throughout?
- Are some individuals targeted with violence more than others?
- How do national level politicians compare with local politicians?
- Are non-candidates, such as journalists and election commissioners, also subject to online GBV?
- What can data tell us about different digital platforms and different types of perpetrators?
- What is the geography of online abuse?
- How quickly do abusive and violent posts spread across the internet?

It is important to remember that: **social media analysis is only useful in conjunction with other technological tools.**

The VAWIE-Online Tool seeks to compare levels and types of online violence across genders and various political and electoral stakeholders (such as candidates, journalists elected officials, civil society actors, etc.) who are potential targets of violence. Using the data collected, the tool aims to note trends over time and analyze the differences between various categories of violence. These findings can assist in identifying drivers of online electoral violence.

In order to answer these questions, the implementer will need to set the parameters of the study. We will walk through how to establish the individuals, lexicon, and social media platforms for analysis in Part II.

Part II – Conducting a VAWIE-Online Social Media Analysis

Step 1: Set Up the Study

The VAWIE-Online Tool is intended to be implemented during electoral cycles, ideally covering the formal beginning of the election period (for example, opening of voter or candidate registration) through closure (acceptance or results, seating of new parliament, etc.). The exact start and end date of an election may not be entirely clear and varies according to context. It is significantly easier to remove data than to add historical data. So, it is best to use a broader time frame rather than a more restrictive one when setting up the analysis. The tool is adaptable for different time frames; it can be used for a specific electoral period, for historic analyses covering many years or as a rapid-response tool to evaluate VAWIE-Online on a day-to-day basis.

The VAWIE-Online Tool may be used for local, regional and/or national elections in any country where social media is actively used and is free from state control or shutdown. It may also be adapted to use outside of electoral cycles to monitor violence against women in politics more broadly. Given social media's global reach and the existence of alternate channels (proxy sites, etc.) in all but the most repressive regimes, the tool can be adapted and used virtually anywhere in the world. When setting up a study, define and record the specific timeframe it will apply (to as well as the geographic focus (nationwide/regional/local).

The VAWIE-Online Tool collects and analyzes online data through analytic platforms; there are significant differences between the software tools available and these differences impact when the analysis can be run. We will address the differences between these tools, as well as their respective benefits and disadvantages. All tools currently require some setup at the start of the analysis, as well as some maintenance throughout the election period. Many implementers will choose to run the analysis multiple times during an election, in order to gain insight into violence in real-time and take action to inhibit its spread. Criteria for the selection of a software tool are included in Step 2.1.

VAWIE-Online Methodology Steps

Step 1: Set up the study
Step 1.1: Identify the potential targets to monitor
Step 1.2: Define the lexicon to monitor
Step 1.3: Choose social media to monitor
Step 1.4: Select the research questions

Step 2: Run the analysis
Step 2.1: Select a software tool
Step 2.2: Run the VAWIE-Online analysis

Step 3: Analyze the results

The VAWIE-Online Tool collects data focused on the **survivor, type of violence, frequency** of violent incidents, and the **perpetrators** (to the degree possible).[24] The gender specific nature of attacks are determined through keyword review for gendered references and through qualitative review of results. Since all forms of violence against women reflect both violence perpetrated on a woman because of their sex, as well as violence that disproportionately impacts women, the VAWIE-Online Tool methodology advises the collection of data on both men and

[24] Many perpetrators post from anonymous accounts and/or do not publish identity markers such as name, age, gender, etc. So, it can sometimes be a challenge to determine who the perpetrators are.

women for each variable (see Step 1.1 and "Consider This" textbox). This is done to gather comparable information and to measure the amplitude of the issue. Where an equal sample is not available, calculated adjustments are to be applied.

The VAWIE-Online Tool prescribes a structured method for collecting data in order to measure and analyze these four dimensions: **survivor**, **type of violence**, **frequency**, and **perpetrators**. The steps to follow this methodology are outlined below. Because comprehensive data cannot be collected in any context, the methodology emphasizes steps to take to ensure a strong representative sample. Implementers should document each component, including the choices of variables and any deviations or modifications from the methodology, as well as obstacles encountered or country-specific adaptations. If the analysis is run multiple times during an election cycle and actions are taken based on mid-course findings, implementers should note those actions and look for impacts on reducing violence. Adhering to a common methodology and recording variations allows the VAWIE-Online Tool to develop resources that can be used for cross-national comparisons, and further enhance understanding of the gendered nature of online election violence.

The first phase of the VAWIE-Online Tool is setting the parameters for data collection. The second phase involves scraping the data—or collecting a sample of relevant data from the internet—and the third involves analyzing the data. The investigator(s) should take their time through the first step. Data can be analyzed in a variety of ways, but the analysis will only be valid so long as steps have been taken to collect good and representative data.

The worksheet included in Annex 3 can be used to complete each of the following steps.

Step 1.1: Identify the Potential Targets to Monitor

There are three stages in defining potential targets to monitor, starting with a broad concept and then narrowing into specific individual information.

Identifying Categories of Potential Targets to Monitor

The VAWIE-Online Tool monitors the social media accounts of various individuals and organizations ("subjects") during a given electoral cycle. The subjects whose social media accounts are monitored through the analysis come from the four categories below. In addition to individuals, subjects can include groups, organizations and institutions (for example, political parties, EMBs, and CSOs that have organizational social media accounts).[25]

[25] Although the targets of election violence are individual people, the VAWIE-Online Tool includes tracking organizations and institutions because VAWIE messaging is often posted in these locations.

Political	Institutional	Professional non-state/non-political	Private non-state/non-political
• Candidates from both opposition and incumbent parties • Elected officials from both opposition and incumbent parties • Political aspirants (i.e., seeking nomination) • Staffers • Party members and supporters • Political appointees in government (cabinet ministers, etc.)	• Electoral management body (EMB) permanent staff • Poll workers • Police and security forces • State administrators and civil servants	• Journalists • Civic educators • Civic activists • Community leaders • Newspapers • Civil society organizations (CSOs) • Gender-based violence (GBV) groups	• Private citizens (i.e. private citizens who are politically active on Twitter, including bloggers, professors, entertainment celebrities, retired states-people, etc.)

The process of selecting individuals begins by identifying the appropriate categories and subcategories of interest. It should *not* begin by thinking directly about specific individuals to monitor. Implementers should not select all individuals and groups that are regularly targets of violence. Instead, they should consider the political/electoral context and the demographics of the country from which to identify a representative selection of categories. This may include a full population (i.e., all presidential candidates or accredited observation organizations) or a representative sample (i.e., 30 percent of registered candidates or the top five leading newspapers).

The number of subjects selected, and the number of subcategories covered, will vary in each country where the VAWIE-Online Tool is used and is scalable according to resources. It is recommended to list the selected categories in order of priority to the study, recognizing that adaptation often occurs as data emerges. Not all projects will have enough resources to cover all (sub)categories, so the preliminary analysis should be performed by project implementers to decide which categories are the most pertinent. Implementers are urged to identify a diverse group from among the categories and to identify multiple subjects from within single subcategories (i.e., identifying multiple candidates, including several from opposition and incumbent parties, several rural and urban districts, etc.). In some cases, implementers may prefer to focus on only one or two unique (sub)categories.

Questions to consider during this initial stage are:

1. What are the benefits/drawbacks of gathering data on each of the four categories (political actors, institutional actors, professional and private non-state/non-political actors) in our

country context? Do we want to gather comparative information by documenting a sample from each of them? Should we select only a few or just one of these?

2. What subcategories are most relevant from the categories we have chosen? Are any subcategories not pertinent in our country case or not logistically viable (for example, security actors or civil servants may not have social media presence; social media accounts for political aspirants, individual party supporters and/or poll workers may be difficult to identify; there may be no female leaders in certain religious communities; etc.)? Are there subcategories that are particularly relevant for this election (national, regional, etc.) and should therefore be included? Are any subcategories missing that should be added?

3. Within these subcategories, what do we need to consider to ensure a balanced and representative sample? For example, if we intend to monitor political candidates, what are the main political parties? Are there regional differences in the country that should be considered? Religious or ethnic divisions? Rural/urban divides? Significant wealth gaps that are reflected in the candidates? To take a different example, in identifying professional non-state/non-political actors, we should ask: what are the leading civil society groups in the country? Do religious leaders and or civic actors regularly speak out about democracy and elections in this country? What are the leading state-owned and independent newspapers, TV channels, and/or radio stations in the country?

Select Potential Targets

Next, the implementer should define the selection of potential targets more specifically. At this stage, the specific details can be mapped out, including the numbers of individuals and organizations to monitor based on the categories identified in the previous step (this is the research model). For categories where it was chosen to monitor the full category population (i.e., all candidates in a given race or all accredited observation groups), the implementer will simply record the names of the individuals or organizations in the next stage. For categories where a representative sample will be monitored, the implementer should apply the following rules[26] to the selection:

1. Always apply gender parity rule to the sample (50 percent men, 50 percent women).[27] Apply an alternating approach in selecting male and female subjects for each subgroup (i.e., for every female candidate, there should be one male candidate, preferably as similar in demographic profile as possible in terms of age, rural/urban, ethnicity, religion, etc.). We recommend collecting data on both men and women for three reasons. First, collecting data on both sexes enables comparison and understanding of the proportions of violence experienced by women. For example, the pilot in Zimbabwe discovered that women faced up to 300 percent more online violence than men and that the nature of the violence women experienced was much more intense and threatening than content related to men. Second, it may be beneficial to

[26] In formal research terms, these rules reflect purposive proportional quota sampling.
[27] Or adapt proportionally if working in contexts with non-traditional gender identity categories.

collect data on both sexes because it can reveal deeper gendered patterns in VAWIE-Online. For example, in Sri Lanka and in Ukraine we found that politically active men were regularly attacked on the basis of sexuality as well, but in these cases the attacks were a way to reinforce the image of politics as being dominated by heteronormative men and exclusive and hostile to women or any people with "effeminate" qualities. Third, collecting data on both sexes can make the research findings useful to a broader range of users and multiply its impact. Many researchers and advocates today are looking for comparative data, while others want only data about women. Collecting both accommodates the broader range of interests and needs at effectively no additional cost and minimal extra effort.

2. Define other relevant demographic factors (e.g., age, race, ethnicity, religious, disability status, wealth, rural/urban, etc.) and seek to reflect this diversity in the selection (see text box on intersectionality).
3. Ensure political balance (for pertinent categories) by including incumbent and leading opposition parties, as well as smaller or fringe parties.

Based on these rules, select the individuals and organizations to monitor. These individuals are potential targets of online election violence. Implementers should avoid taking their perceived risk level into consideration when selecting them in order to maintain objective and unbiased results. This is recommended because the VAWIE-Online Tool seeks to create an objective picture of the existence *or absence* of online violence and tries to avoid possibly inflating results by simply monitoring those individuals who are deemed most likely to be survivors of violence, for any reason (e.g., past history of inciting or being subjected to violence, association with a violent region, etc.). Although high-risk individuals should certainly be included, the VAWIE-Online Tool should not focus exclusively or disproportionately on them. It is recommended to work with a focus group.

If there are high risk individuals that it would be beneficial to monitor—that do not otherwise fit into the analysis—the implementer should note those names for use as individual cases. Later, in the scraping and analysis stages, it will be possible to analyze these names separately.

Consider This: Gathering Data on Men and Women

While the present methodology proposes that researchers devise studies with a sample comprised equally of men and women to disaggregate the impact of violence, the methodological approach of analyzing data on men's experience of violence to validate women's experience of violence is a contested one.[28] Many groups working to understand the phenomenon of the gender-based violence in politics regularly conduct research exclusively on women.[29] This approach focuses on the impact of the violence experienced by women rooted in the theory of gendered power imbalances, not in an analysis of comparative data between men and women. Researchers that use this theoretical approach take the view that a comparative focus on men's experience of gender-based violence is not germane if the overall goal of a study is to understand women's experience of gender-based violence in order to address it better. If research does include an analysis of the violence men face, in order to understand and gather the data on the perpetrators, survivors, types of violence, and language used, a methodologically sound and contextually specific approach to gathering the data ***as it relates to men*** is needed. A 2007 UN Expert Group Report on Indicators to Measure Violence Against Women highlighted the potential pitfall of trying to gather and analyze data on gender-based violence experienced by men and women using the same tools by noting: "While violence against men is also an important issue requiring attention, this violence takes different forms and is not rooted in power imbalances and structural relationships of inequality between women and men. Thus, the broader issue of interpersonal violence, which has male and female survivors, who may also be vulnerable by way of age, disability or social exclusion, requires a separate approach and different methodology to measure it."[30]

Because it, too, is a broader issue, the same may be said of "political violence." There may also be other methodological issues to consider. The World Health Organization's multi-country study on women's health and domestic violence against women originally intended to survey men and women for a comparative sample on intimate partner violence. However, on the advice of the study's steering committee, it was decided to include men only in the qualitative, formative component of the study and not in the quantitative survey. It was determined that there were safety issues to be considered in interviewing men and women at the same time and the cost to carry out an equivalent number of interviews of men and women would be too high.[31] Both of these issues are relevant to the conduct of any simultaneous examination of the violence experienced by women and men in politics. Sentiment

---

[28] Elin Bjarnegård, "Making Gender Visible in Election Violence: Strategies for Data Collection," *Politics & Gender* 14, no. 4 (December 2018): 690–95, https://doi.org/10.1017/S1743923X18000624.

[29] "Sexism, Harassment and Violence against Women in Parliaments in Europe," Issues Brief (Inter-Parliamentary Union, October 2018), https://www.ipu.org/resources/publications/reports/2018-10/sexism-harassment-and-violence-against-women-in-parliaments-in-europe.

[30] "Indicators to Measure Violence against Women," Report of the Expert Group Meeting (Geneva, Switzerland: United Nations, October 2007), https://www.un.org/womenwatch/daw/egm/IndicatorsVAW/IndicatorsVAW_EGM_report.pdf.

[31] Claudia García-Moreno and World Health Organization, WHO Multi-Country Study on Women's Health and Domestic Violence against Women Initial Results on Prevalence, Health Outcomes and Women's Responses (Geneva: World Health Organization, 2005), https://www.who.int/reproductivehealth/publications/violence/24159358X/en/.

analysis and data mining research do not require interviews, however, and collecting data on men as well as women is generally resource neutral. As such, decision making around one's methodological approach to data collection may be driven by factors other than funding.

Example of selected targets worksheet from Ukraine

General characteristics of the sample

1.1. Elected Officials (national level)	**Verkhovna Rada (Parliament)**	**104**	Men	**52**
			Women	**52**
2. Political Appointees	**Executive body (Cabinet of Ministers)**	**24**	Men	**20**
			Women	**4**
3. Electoral management body	**Central Election Commission**	**15**	Men	**10**
			Women	**5**
4. Professional non-state	**Journalists**	**30**	Men	**15**
			Women	**15**
5. Private non-state	**Entertainment celebrities**	**10**	Men	**5**
			Women	**5**
	Total	**183**	Men	**102**
			Women	**81**

Intersectionality and VAWIE-Online

Violence against Women in Elections focuses primarily on differences between men and women, which can overlook the differences *between* women as well as the unique obstacles facing individuals at the intersection of multiple forms of discrimination.

Individuals at the intersection of multiple forms of discrimination, such as women who are racial minorities, may face obstacles that are not only sexist and racist, but unique challenges that result from identifying both as a racial minority and as a woman. This is true for multiple identities, including the intersections of sexism, racism, ethnic and religious discrimination, ableism, classism, and discrimination based on sexual orientation and gender identity. It is also true for political ideology or affiliation. To account for this, VAWIE-Online analyses should adopt an intersectional perspective.

Intersectionality theory recognizes that axes of discrimination—such as racism, sexism, and homophobia—are not separate and distinct from one another but instead, mutually reinforcing.[1] For example, scholars and practitioners, focusing on individual country cases, have found that race and religion impact the quantity and quality of online abuse directed at female candidates and politicians (Amnesty International, 2017; Kuperberg 2019).

Adopting an intersectional analysis, or even considering how intersectionality might deepen an analysis of gendered election violence, presents many new questions, including: What forms of discrimination can and should be included in an analysis? How should this change our methods or inform our concluding analysis? There are no definitive answers to these questions. However, analysts should be aware that conducting an intersectional or intersectionality-inspired analysis requires more than a simply adding individuals to an investigation. Intersectionality can and should inform the initial research plan, methods, and analysis.

Some strategies to integrate intersectionality into a VAWIE-Online analysis include:

- An intersectional analysis requires an understanding of the historical and social context of a community. In setting up the VAWIE-Online Tool, analysts should utilize interviews and focus groups to identify one or several salient identities and forms of discrimination in the region of analysis.
- Salient discriminations not only differ across different contexts, but also over time. Because it is nearly impossible to reflect all forms of discrimination in a specific context, choose those that are the most pertinent to the context by using qualitative approaches (talking to experts or using focus groups).
- The choice of identity-based discriminations does not need to be limited to groups that have strong political representation. Transphobic or ableist insults, for example, may be particularly severe when targeting transgender public figures or public figures with disabilities but may also be used to discredit or harm public figures who do not identify as part of those groups.
- VAWIE-Online algorithms can incorporate keywords that reflect prominent forms of discrimination as well as ensure that individuals who represent those groups, where possible, are included in an analysis. The methodology can be applied to measure whether individuals or groups in the sample are more likely to be targeted with keywords or phrases pertaining to forms of discrimination.
- Finally, at the analysis stage, analysts can explore the rhetoric or context of posts to better understand how multiple forms of discrimination affect individuals and groups. In addition to larger-scale data, analysts should look in-depth at individual posts to understand qualitatively how users are employing discriminatory rhetoric and tropes.

(text box contributed by Rebecca Kuperberg, Rutgers University)

Securely Record Subjects' Key Information

Although the VAWIE-Online methodology described here only uses publicly available information, it is very important to first define how to protect personally identifiable information (PII) in conducting social media analysis for both survivors and alleged perpetrators. PII is any data that could potentially identify a specific individual. Any information that can be used to distinguish one person from another and can be used for de-anonymizing anonymous data can be considered PII. Laws around PII vary in different countries and are rapidly evolving. All users of this guide must first check the legal requirements before collecting data on individuals.

Furthermore, because legal protections are often lacking for PII, it is the responsibility of the implementing group to define their own guidelines for protecting PII that may go above and beyond legal baselines. For example, implementing groups should decide how they will maintain and protect electronic data files (secure storage, password protection, erasing data after a time, etc.). Implementers should also decide how they will protect PII during the reporting stage (e.g., only report aggregate data findings, anonymize screen shots and texts used in reports, etc.).

After determining requirements and voluntary standards and practices in protecting PII, the next step is: For each individual identified, the following information should be researched and recorded (categories may be adapted and expanded; only categories marked with an asterisk are required):

- Full name*
- Age
- Sex*
- Nicknames (important for candidates that are well-known according to a nickname)
- Ethnic and religious identity (if known)
- Disability status (if known)
- Marital status
- Location (constituency for candidates or EMB staff, province/state for others, or rural/urban binomial)
- Political affiliation (where relevant)*
- Social media information (Twitter handle, Facebook account name, etc.)
- Category* and sub-category* (see example in Annex 4)

For each organization/group identified, research and record the following information:

- Full name*
- Legal address
- Geographic presence (nationwide, regional, both)
- Political affiliation (where relevant)
- Social media information (Twitter handle, Facebook account name, etc.)
- Category* and sub-category* (see example in Annex 4)

Information should be recorded and saved in a database, and implementers should take steps to ensure that the privacy and security of the tracked individuals' information is protected.

Step 1.2: Define the Lexicon to Monitor

Example of lexicon-based word clouds of VAWIE-Online (Zimbabwe and Ukraine)

Lexicons are lists of tag words and phrases that reflect a given type of violence or harassment. The VAWIE-Online Tool analyzes types of violence according to distinct categories because this enables a better understanding of the nature and intensity of violence that is occurring. Differentiating categories of violence also provides some insight into the drivers motivating different types of violence.

Each implementer should develop a list of tag words and phrases (i.e., a lexicon) categorized by typology of violence (see Annex 3). The main typologies of VAWIE are bodily harm, sexual, and socio-psychological. Because economic forms of VAWIE-Online are infrequent on social media and because they are generally harder to reliably identify, we do not recommend including them unless there is a strong reason.

Tips for Typology: How to Create an Effective Lexicon

Implementers may need to adapt the global categories to local context. Some common adaptations include:

- Use the "sexual" category to distinguish threats of physical sexual harm (rape, assault, etc.) from other socio-psychological threats that use sexualized language (i.e., calling someone a whore, a cunt, etc.)
- Create sub-categories to organize content better (see examples on the next page).

For example, words and phrases reflecting bodily harm would include mentions of physical threats or abuse, including threats of physical harm to the spouse, children, or other family of the survivor. Sexual violence includes direct threats of sexual harm. Socio-psychological phrases will likely be the most diverse and may include attacks on a person's intelligence or moral virtue, their standing as a spouse or member of a religious faith, threats to shun or isolate a person, threats to forcibly control or repress electoral participation, as well as intentional efforts to shame, harass, instill fear or defame an individual in relation to their electoral participation. If choosing to include economic forms of violence (not advised), these forms may include words and phrases that threaten to deny resources or services, harm a business, or unlawfully control and monitor of the use and distribution of monies and access to services (healthcare, employment, etc.).

Some words may seem to apply to multiple categories. However, in order to compare between categories, words and phrases should only be collected for one category.

Monitored words and phrases should be based on common, popular vocabulary in the country context. Local language variations should be taken into account. The words can include a variety of verbs, nouns and adjectives. An initial list should be defined at the outset. However, it will be updated over the course of monitoring as the electoral context evolves and after an initial analysis of the data.

The best way to initially develop this list is to consult with GBV service providers, female candidates, and regular social media users, who have insight into common terms and phrases used online. These will likely include vulgar and explicit language. In conjunction with consultation, the implementer can also look through the scraped, online data for patterns of violent keywords. If working in multiple languages, it will be useful to create separate lists for each language. It is also useful to refer to existing hate speech lexicons, including hatebase.org or to contact CEPPS/IFES for examples.

Implementation Requirement:

VAWIE-Online lexicons and analyses contain very disturbing and offensive content, which can be traumatizing or act as a trigger for the men and women that interact with it. Ensure the well-being of all people working on a VAWIE-Online implementation (focus group discussion participants, researchers, translators, etc.) by preparing them for exposure to this content, offering opportunities to address emotional reactions and respecting their right to withdraw.

Because the VAWIE-Online Tool requires a gender-balanced sample, it is very important to include words/phrases that are specific to men as well as to women. For example: if women-specific derogatory terms are selected (such as “bitch,” “whore,” etc.) it is critical to also identify similar words/phrases that are specific to men (e.g., “fag,” “dick,” etc.). Failure to do this will result in skewed findings and undermine the methodology.

Finally, it may be helpful to create sub-types to organize large lexicons. For example, physical violence can be further distinguished by direct and indirect threats made to the primary target or to a proxy of the primary target (such as a child or spouse). Socio-psychological forms of violence vary dramatically between different countries and it can be helpful to create sub-types here as well, for example threats relating to traditional gender roles, threats relating to body-shaming, homophobic threats and slurs, attacks on intelligence and competency, etc. In some contexts, highly specific cultural sub-types will emerge. For example, in Zimbabwe we found that accusations of witchcraft were connected with VAWIE-Online, while in Ukraine accusations of Kremlin-connections were part of VAWIE-Online. These are clearly culturally specific.

Universal Categories – VAWIE Types (on/offline)	Example Tailored Categories from Zimbabwe – (to adapt according to cultural context)
PHYSICAL	Direct harm to principal target Implied harm to principal target Direct harm to a proxy (family member, associate, etc.) Implied harm to a proxy (family member, associate, etc.)
SEXUAL	Threatened rape or other forms of *physical* sexual violence to principal target or proxy Marital-related (marital infidelity, respect for martial duty, motherhood, etc.)
SOCIO-PSYCHOLOGICAL	Competence/Intelligence Witchcraft/Sorcery Appearance Bestial comparisons Sexual orientation/gender identity

Consider This: Building Lexicons of Harassment for Women and Men

The concept of violence against politically active women, including those who are actively seeking to participate in an electoral process, holds that when a woman experiences violence it has a disproportionately negative impact on her because of her subordinate status in society.[32] This may prevent her from participating in politics in her own voice and conscience, free from fear or threat of reprisal.[33] This impact is also true of the violence that politically active women face online. This toolkit develops lexicons of harassing or harmful language, using the typology of physical, sexual, socio-psychological and economic violence to categorize the online violence. Other typologies are organized differently. For example, the National Democratic Institute (NDI) builds lexicons in categories of **insults and hate speech, embarrassment and reputational risk, physical threats**, and **sexualized distortion.**[34] Any instance of these acts represents a direct barrier to women's free speech, undermining democracy in all its key elements, and has a chilling effect on the ambitions of young women and new entrants to politics. For example, when it comes to online violence in the form of embarrassment and reputational risk, words used to describe male and female sexual behaviors are not rooted in the same intent, and do not result in the exact same impact. In the United States, for example, when a man is called a "dog" and a woman is called a "whore," they both indicate high levels of sexual activity, but the former is viewed much more positively than the latter, even though it is equating a human's behavior to that of an animal's. It is extremely difficult to find words that are intended to have an equal negative impact when applied to men as to women. In fact, in this instance, it is possible that measuring words as if their impact is equal simply because they describe a similar phenomenon (in the illustration above, sexual activity), *may lead to* a skewed outcome in terms of the results. While men and women may experience equal *quantities* of online violence or be attacked in the same *mode* (such as the online dissemination of similarly sexualized videos of party candidates), this does not mean that the online violence has the same impact either on the women's confidence in their own ability to be political leaders, or in the public's perceptions of them as political leaders relative to men.[35] As suggested by this toolkit, machine learning-based sentiment analysis may partially correct the potential for skewed results. However, the software is only as capable as the humans building it: providing the machine with a constant diet of equally negative words or phrases as these emerge and evolve in a particular language, requires significant knowledge of political context, gender norms and language.

---

[32] The National Democratic Institute, *#NotTheCost: Stopping Violence Against Women in Politics. A Call to Action,* (March 2016)

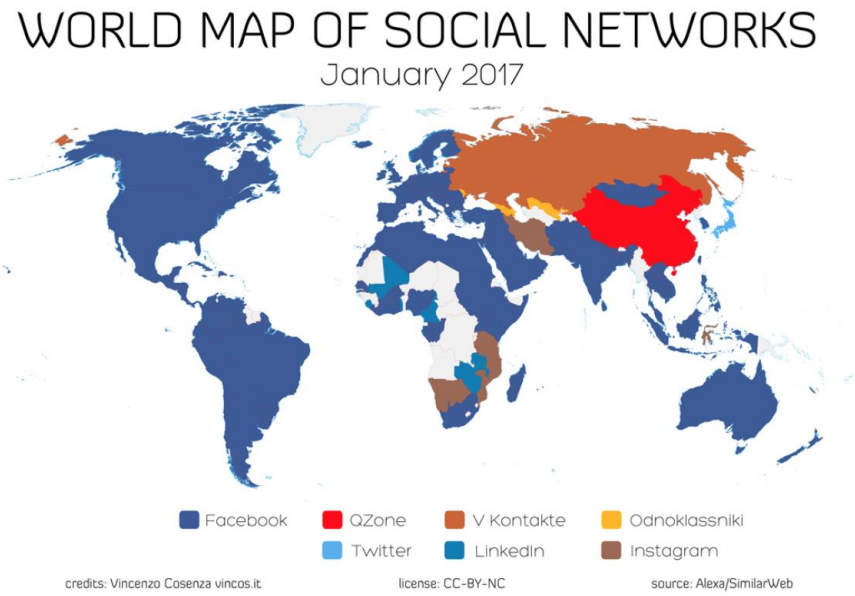
[33] Ibid

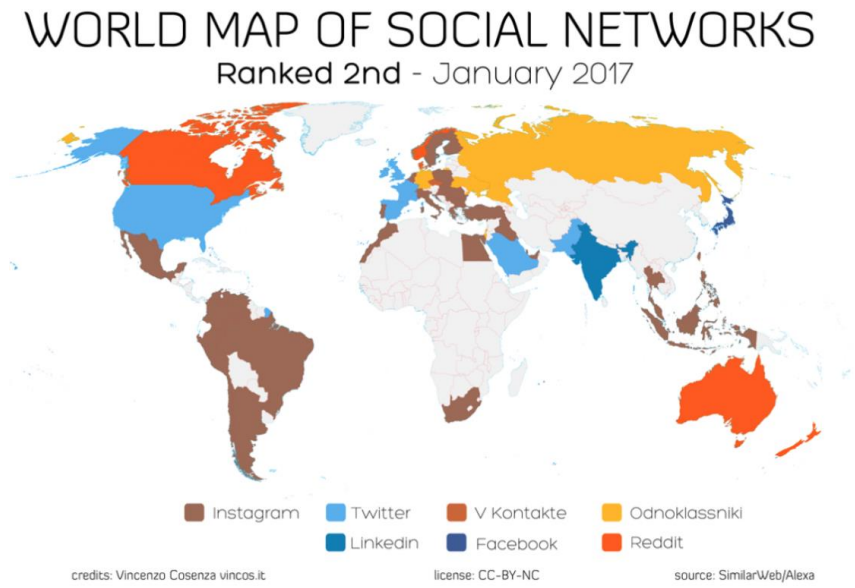
[34] National Democratic Institute report of a project to examine the impact of online violence against politically active women on young women's online political discourse. Forthcoming 2019.

[35] Jankowicz, Nina. "How Disinformation Became a New Threat to Women." (.coda, Washington, D.C., 2017) https://codastory.com/disinformation/how-disinformation-became-a-new-threat-to-women

Step 1.3: Choose Social Media to Monitor

Social media use varies from country to country. For example, Facebook is the leading social media platform in 119 out of 149 countries, with 1.8 billion monthly users worldwide,[36] meaning there are 30 other countries with more popular social media platforms. Likewise, the second most popular social media platforms vary significantly between countries.[37]





[36] Rosamond Hutt, "The World's Most Popular Social Networks, Mapped," World Economic Forum, March 20, 2017, https://www.weforum.org/agenda/2017/03/most-popular-social-networks-mapped/.

[37] Image source: Rosamond Hutt, "The World's Most Popular Social Networks, Mapped," World Economic Forum, March 20, 2017, https://www.weforum.org/agenda/2017/03/most-popular-social-networks-mapped/.

Social media use also varies within demographics in any given country. Many forms of social media attract higher rates of use according to the age, socioeconomic class, sex, and ethnicity of the user.

In adapting the VAWIE-Online Tool, the implementer must assess the social media environment in the country context. The Tool can monitor multiple social media platforms[38] and the implementer can choose multiple platforms that reflect this diversity. In assessing the social media environment, the implementer should avoid basing decisions on personal social media use and try to use formal research sources (such as http://vincos.it/world-map-of-social-networks/) as well as consult with producers of content (i.e., which platforms do candidates, EMBs, media publish on?) and consumers of content (which platforms do people of different age groups, ethnicities, and sexes use most?).
Note that social media platform uses changes regularly and quickly over time. Do not rely on past years' analyses, but seek to use the most up-to-date sources. The targeted online platforms will have an impact on the software product chosen in Step 2.1.

This assessment will determine which social media would be best to monitor. It is important to note that, depending on the software that will be used, it may be possible to monitor all sources available. Although it is critical to review other sources of information on social media reach, monitoring all available social media platforms and online media through this analysis can be a means to offer an assessment of the popularity and prevalence of social media platforms used for online harassment and violence.

Step 1.4: Select Research Questions and Understand Possibilities

At this point, there should be a list of names and organizations to monitor, as well as a tailored lexicon. These pieces will allow for a thorough and rigorous model to be developed. In the following steps, these pieces will be used to run the VAWIE-Online Tool analysis. Before discussing how to build and run the model, we're going to broadly cover what to do with the information that's already been collected. There are various questions that an implementer might want to ask about VAWIE-Online. These might include: Who is targeted by online violence? Are men and women targeted equally? Do certain categories of violence influence men or women more? What do we know about perpetrators of online violence? Where are perpetrators coming from: in country or out?

The information collected enables the implementer to answer these questions and others.
Without parameters, the amount of information on the internet would be overwhelming and not particularly useful. Using the names, words, and phrases that have been developed—the parameters for the model—and applying these to the software tool that will be used, the implementer will move from general concepts into a specified algorithm.

[38] For restrictions on types of platforms that can be monitored, refer to the end of Part II for a discussion on limitations and constraints, and to Step 2.1 on software parameters.

The algorithm should be run a number of times and be modified by different populations (political actors, institutional, etc.), different forms of violence (bodily harm, socio-psychological, etc.), and gender (men, women, non-binary individuals) to focus on each variation (i.e., male political actors targeted with bodily harm, female institutional actors targeted with economic violence, etc.). Though it may seem unnecessarily specific, running separate, tailored algorithms will allow the implementer to reduce irrelevant posts in order to compare across gender, across populations of actors, and across forms of violence. Unnecessary and useless posts are referred to as "noise." By restricting the amount of "noise" in the sample, it will be possible to identify trends and patterns of violence.

Not only will it be possible to determine whether men or women receive more overall violence within the sample, but whether male or female political candidates receive more socio-psychological violence in the sample. By running smaller analyses, the implementer can ask and answer more specific questions; importantly, it will be possible to combine smaller analyses to collect larger amounts of data and to answer broader questions.

In the next section, the parameters that have been created will be applied to a particular software tool. Though the technical component of the analysis can be complex and time consuming, establishing clear, well-researched, and representative parameters is the first step to getting valid results.

Step 2: Run the Analysis

Step 2.1: Select a Software Tool

In order to use the information that has been collected to analyze for VAWIE-Online, it will be necessary to work with a data analytics software program. The structure and approach outlined above will allow the implementer to conduct a VAWIE-Online analysis using their choice of software. The choice of software tool will depend on: (a) the social media platforms that the implementer would like to monitor; (b) the implementer's technological proficiency (and access to technological resources); and (c) financial resources.

There is no single, perfect software to run the VAWIE-Online Tool. In developing this guide, one established software program was tested (Crimson Hexagon)[39] and several self-run programs were also tested for various components of the analysis (R, Python, WEKA, #TAGS, NVivo NCapture). Some software tools have more drawbacks than others. We have worked with Crimson Hexagon because it is relatively easy to use and allows for a broad range of questions to be answered and combines several actions into a single platform (data scraping, mining, sentiment analysis). We recommend, based on existing work and research, that an established program that performs multiple functions within a single platform is used. Even with established programs, however, steps should be taken to improve the

[39] Only one established software program was tested due to cost. Crimson Hexagon was selected at the outset of the study based on a market comparison of available programs and cost-functionality options.

validity of the results. We will also cover several alternate platforms and describe some of their advantages and limitations.

Specific details on running self-run programs can be found in a later section under this step and in Annex 5.

Comparison of Software Tools

	Established Programs (e.g., Crimson Hexagon)	**Self-Run Programs (e.g., NVivo, QDA Miner, etc.)**
Cost	High (approximately $10,000 USD)	Low to Medium (generally less than $3,000 USD
Technological proficiency required	Low to medium	High
Time and effort required	Medium	High
Platforms available (as of December 2018)	Blogs, Facebook, Instagram, Newspapers, Twitter, YouTube, Forums, reddit, Google Plus, Tumblr, QQ, and review.	Facebook, Twitter, YouTube, LinkedIn and general web pages
Ability to use languages other than English	Yes, with some limitations	Varies
Multiple tools need to be utilized	No	Yes

Established Programs

We found established software packages far easier to use and they produced more comprehensive and reliable results then open-source or low-cost alternatives. Only one established program was tested for this guide (Crimson Hexagon), however similar benefits and drawbacks outlined below may be anticipated to apply to other established programs. The benefits of established software include:

- It is user-friendly and requires only a beginner to intermediate level of technical expertise. The tool provides resources to help the user become familiar with the platform, so the implementer can learn as they go.
- It has a dedicated help desk. If there are any issues, it is easy to email or chat with technical support and they can assist with the project.

- It consolidates many steps into one. Here, the implementer can direct the tool to scrape the information they are looking for and it performs the data scraping and mining processes at the same time. Once the data has been collected, it can be analyzed for sentiment and the posts can be categorized. The Tool sets up a structure for the implementer, so there is no need to write unique algorithms to categorize data.
- The implementer can look at past data. For Twitter in particular, Crimson Hexagon can show data from months and years ago.
- It creates images for export. The tool collates all data and creates images, which can be used in analyses and reports.
- Crimson Hexagon is constantly being updated with more tools and programs, some of which will enhance its ability to collect VAWIE-Online data.

But, just as with any tool, Crimson Hexagon also has drawbacks:

- In non-supported languages (English, Spanish, and Arabic are among the supported languages), the tool has greater limitations.
- Crimson Hexagon and similar established programs are costly.
- It is difficult to incorporate images from the internet into the analysis without significant added cost
- There are limitations on data exports.
- The tool is powerful and, as such, incorrect training can have a dramatic influence on the results. The implementer must be very careful to train and categorize posts precisely and to keep records to ensure consistency across the data.

Crimson Hexagon is not the only tool available that can run data analytics. Furthermore, new tools are being developed every day. If there is a tool that works better for the language being used or can perform more functions, the implementer should use it. In Step 2.2, Part G, we describe steps to maximize the accuracy of the data. Regardless of the tool being used, implementers should ensure that the data is as accurate, valid, and reliable as possible. If the tool generates inaccurate, invalid, or unreliable data,[40] it is encouraged to seek out another tool.

It is imperative that the software is one with proactively secure means of data storage. Crimson Hexagon has privacy statements on their website that transparently show how they use the data of their clients and indicates that they "put in place safeguards to protect your data." They include that they share data with selected third parties, including their "business partners, suppliers, and sub-contractors, for the performance of any contract we enter into with them."[41] Understanding how the software company handles the data—and how securely the data is protected—is an important part of the decision regarding what program is used.

[40] Social media data studies rarely achieve complete validity. However, we recommend that implementers aim for 80 percent accuracy. If this level of accuracy cannot be achieved, even after amending the algorithms, we recommend that another tool is utilized.

[41] The privacy policies of Crimson Hexagon can be found here:
https://www.brandwatch.com/legal/author-privacy-policy/
https://www.brandwatch.com/legal/user-privacy-policy/

Self-Run Programs

When implementing the VAWE-Online Tool, the use of self-run programs is not recommended for groups that do not have access to experienced data analysts. At the time of writing, there was no single tool that is both user-friendly and allows the implementer to achieve the same data output as an established program. However, given the cost limitations of established programs, and the possibility that language-specific or future self-run programs will be more comprehensive, we are providing a brief guide to some of these programs here. If an analysis is conducted using self-run programs, it will require much greater technical skill and time while offering a more limited range of results.

Using self-run programs, an implementer can scrape data and mine data for particular words. An implementer can also conduct a sentiment analysis—in particular languages—and run statistical analyses. However, it is not recommended that an implementer conduct a sentiment analysis alone, without also classifying posts as violent according to category and subcategory. As such, an implementer using a self-run program will need to data mine the text and then hand code either all posts or a representative sample of posts.

Implementers using a self-run program will need to data mine the text and then manually code either all posts or a representative sample of posts.

In general, self-run programs have several drawbacks. First, it is much more difficult to look at historical data. For this reason, if an implementer decides to use a self-run program, we recommend that the study is set up well in advance of the election cycle in order to run a sample and make adjustments prior to the start of the election period. Second, while it is possible to automate some parts of the process, the implementer will have to continually check in or re-run aspects of the analysis. Third, different platforms, such as Twitter and Facebook, will almost certainly require different tools. Further, it may be found that data scraping, mining, and analysis—even using the same social media platform—cannot be conducted in the same platform. For each step, the data will have to be reformatted so that the subsequent software tool can read and analyze it. Fourth, self-run programs require significant technical expertise.

In general, the cheaper the tool, the more technical expertise is required. Some of the free software tools, for example, require expertise at the level of a developer or computer programmer. Finally, most self-run programs do not have built-in, machine learning technology. As such, while the implementer can look up keywords or determine sentiment of a series of posts, they will likely need to incorporate a significant amount of hand-coding (one-by-one categorization) into the analysis.

Despite these limitations, self-run platforms offer an affordable alternative to established programs and allow for some greater flexibility in establishing algorithms. Though implementers are more limited in what they can achieve using these tools, they have control over more components of the process. As a downside, however, this means that at every part of the process, there is greater room for user error. We will now briefly address a few recommended self-run programs. Keep in mind that these tools are subject to change over time and that new tools may become available.

#TAGS

#TAGS is a free program that works through Google Docs. It allows the user to scrape data from Twitter using keywords and Boolean search operators ("AND" and "OR"), and is explained more fully in the section below. It has a self-run function, so the implementer can set up the tool to run every hour and will not need to continually run the analysis. It is also relatively straightforward and requires some, though not an extensive, amount of technical skill. There is a help page and several forums online where implementers can find assistance. #TAGS is able to pull very similar data as Crimson Hexagon. However, this tool is only useful for Twitter and cannot be used for other platforms. It also only scrapes data; it does not data mine and therefore data needs to be exported into a separate tool.

NVivo NCapture

NViVo is not a free program but is offered at a reduced price for some institutions. NVivo NCapture runs as a plug-in to one's internet browser. Once the plug-in is downloaded, the implementer can visit a public Facebook page of interest and click the plug-in. From there, the tool will capture all comments and posts made on the page and export them into NVivo for analysis. NViVo can be utilized for Facebook, Twitter, YouTube and blogs, though it is more effective for particular types of sites and particular forms of data. NViVo can perform some data mining and allows the implementer to categorize posts, though its machine learning capacities are limited. It is also limited as an implementer has to pull one page at a time and it can be difficult to control time-based parameters. NVivo is an unwieldy program, prone to crashing. This program should not be used on a computer that does not have a large amount of available storage and relatively upgraded specs.

Other Tools

There are a number of other tools available such as such as tidytext, which can be used for data mining and sentiment analysis in R**,** and VADER, which can be used for sentiment analysis in Python. Once formatted, data pulled from #TAGS can be imported into R or Python for analysis. WEKA is a free data analytics and machine learning tool, but it requires a substantial amount of technical expertise. More information on self-run programs can be found in Annex 5.

For any program that is used, it is critical that the implementer ensures that the privacy and dignity will be protected for all individuals that the program will find in the analysis.

Step 2.2: Run VAWIE-Online Analysis

Running Crimson Hexagon

All of the parameters of the study have now been established: names, lexicon, and social media platforms. Following these parameters, the implementer should now construct algorithms to run the analysis. An algorithm is a set of rules created using words and phrases. In this case, the algorithm will instruct the computer and the software platform that will obtain the data.

The implementer should follow the parameters that have been established in the preceding steps, without excluding or modifying any parts. If there are any changes or adjustments, record them and apply them consistently across all of the subjects affected by the change.

As mentioned, it would be impossible to read through all of the posts on every politician's Facebook page(s) and copy them one-by-one for analysis. However, even having all Facebook, Twitter, and blog posts over a certain period doesn't mean that analysis will be easy. Not all social media posts are violent. And, because some of the politicians that the monitors are following may be very popular, the implementer might have to wade through millions of posts in order to find the forms of violence they are interested in. In order to find and categorize violent posts, we encourage implementers to follow the following methodology.

The analysis will involve scraping data from the internet, mining this data for violent posts, and categorizing these violent posts. In an established program such as Crimson Hexagon, these steps will be consolidated and made easier. Using other software tools, these steps will be different and will require technical knowledge at each step. Thus, depending on which software tool is utilized, these steps will be slightly different.

Three steps of analysis

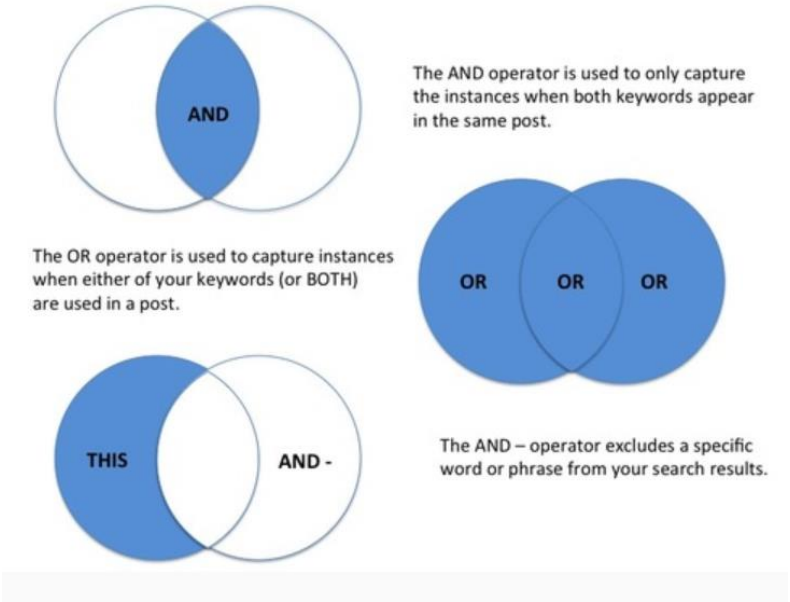
1. Scraping data from the internet
2. Mining the data for violent posts
3. Categorizing the violent posts

a. Write the General Algorithm

No matter which is used, the implementer will need to take all the parameters they have established and turn them into an algorithm, or a code that the computer can use. If looking at less than 20 candidates—which is not recommended for a countrywide study—this step can be skipped.

Because the implementer is likely looking at a large number of individuals, groups, and organizations, it would take a very long time to look up each of these individuals one-by-one. Instead, the categories of individuals should be separated—political, institutional, professional, and private—and these four categories should be further separated by male and female. The implementer should be able to utilize the spreadsheet from Step 1.1 and easily separate that list into those eight categories.

Algorithms are sensitive to language and spacing. Many run using Boolean logic, which hinges on two possibilities: true and false.



For the purposes of this methodology, Boolean logic separates "and" commands (this AND that) from "or" commands (this OR that). For example, when looking for posts that mention both Barack Obama and Hillary Clinton, we write: "Barack Obama" AND "Hillary Clinton." In this case, only posts that mention both of these phrases will be included: a post that discusses Barack Obama but not Hillary Clinton, will be viewed by our tools as "false," and will not be included in the set. Alternatively, if we want any post that mentions either Obama or Clinton, we will substitute the "AND" for an "OR": "Barack Obama" OR "Hillary Clinton." If we want a post that mentions Obama but explicitly *does not* mention Clinton, we write: "Barack Obama" AND- "Hillary Clinton."

For this step, we are looking for any post that has any of the names that were collected in Step 1.1. If we utilize "AND" commands, our software will only collect posts that contain multiple names. Instead, we need to use "OR" commands. That way, any post that contains one name OR another name will be included.

Utilizing the list of names and organizations, the implementer should create eight separate algorithms (male and female for each of the four groups of targets (political, institutional, professional, and private) writing all the names with OR in between them. Also note that if searching for phrases, they should be put in inverted commas, but if searching for one-word, inverted commas should not be used. For example, for phrases the implementer will type "Barack Obama" OR "Hillary Clinton" and for one-word logic they will type Obama OR Clinton. Placing parentheses around the lists instructs the tool to do one action; this also makes it easier to organize subsequent algorithms. Each list with parentheses around it containing all of the Boolean logic counts as one single algorithm, no matter how many AND/ OR connectors are between each word or string of words. Here are hypothetical examples:
("Hillary Clinton" OR Clinton OR "Beyoncé Knowles" OR Knowles)
(Obama OR "Barack Obama" OR "Michael Jordan" OR "Donald Trump" OR Trump)

Here, we've separated males and females. When the algorithms are completed, the implementer should have a list for each group, and within each group, two lists separated by gender. It is possible to have a list of male candidates and female candidates, male political journalists and female political journalists, and male and female civic activists and election administrators. If individuals are commonly referred to by a nickname, these can be included too ("Barack Obama" OR Obama) ("Beyoncé Knowles" OR "Queen Bey").

This step should be done as early as possible. In Crimson Hexagon, for example, the program will only begin to collect Facebook data when it's requested. Thus, this algorithm needs to be established and run at the start of, if not before, the period of analysis. We recommend that the analysis is started early, so that any problems any problems can be found before the election period begins. This way, the analysis can be edited before the start of the election period, resulting in better data. Crimson Hexagon is particularly user friendly in that, if there are errors in the Boolean logic of the algorithm, the software will send an error notification (similar to the ones that pop up when formulas are written incorrectly in excel) that prompts the user to correct the logic.

The implementer can use the name and nicknames of the public figure or the username of the public figure. If using an established software, the name of a person should be used, not their social media username, as it will be useful across Twitter, Facebook, blogs, and other platforms.

During this step the implementer should write out the lexicon (from Step 1.2) into Boolean logic. This will involve taking each word or phrase and writing OR in between them (e.g., murder OR kill OR die OR beat OR bury). Note that the same set of keywords used for men and for women should be written into the Boolean strings regardless of the sex of the targets they will monitor. This because even if a target is a particular sex, they might be attacked on the basis of the sex of significant others who surround them. Wives are attacked because of their husbands, daughters for their fathers, sisters for their fathers, and so on. We are not going to use this lexicon algorithm yet, but we will use it in a later step.

When further training and refining the algorithm, the Boolean operator "AND-" will be used to remove any posts that contain keywords that are likely to create noise or bring in irrelevant conversation. The AND- will not return the results even if the conditions for the AND and OR connectors have been satisfied. For example, suppose the logic is as follows: (Obama OR "Barack Obama" OR "Michael Jordan" OR "Donald Trump" OR Trump) AND (murder OR kill OR die OR beat OR bury) AND- (lotto OR "code generator" OR "play lotto"). With this, even if a post contains Trump and murder, as long as it has the word lotto, such results will not be included in the found posts. Use of AND- connectors should be used after realizing the critical keywords to use to suppress noise. Avoid using more general words that may end up being more stringent. A good keyword for AND- is the one that is peculiar to the noise content.

b. Scrape Raw Data

This step is used by all software tools, but how it is conducted will differ by software.
Using the general algorithm of names and organizations, the implementer will now scrape—or collect—data from the social media platforms they are using.

In all cases, scraping data should only be done in compliance with the laws of the country that is being analyzed as well as the policies of the social media platform. For example, when presenting the data, the implementer should make sure to avoid using usernames of perpetrators of violence and abuse, even if the data is public. Rules and regulations change regularly. It is the responsibility of the implementer to be informed of applicable laws and to apply them.

Laws on social media analysis are underdeveloped in many countries and therefore the implementer may choose to go above and beyond statutes to adapt voluntary ethics guidelines on the use and management of data collected online. It is critically important to protect the privacy and dignity of all individuals (including alleged perpetrators).

If the implementer is using an established software tool, such as Crimson Hexagon, they will simply need to create a new monitor and insert the names algorithm into the monitor. The software will then ask them to set the limits of the analysis: identify the dates, determine which social media platforms they are interested in, etc. With the exception of the algorithm, the established software will guide the implementer through the rest of the analysis parameters.

c. Review Data and Amend Lexicon Algorithm

This raw data is only limited by the number of names or organizations that have been selected. It is not specifically drawing out violent posts. As such, it is likely that a lot of raw data will emerge, much of which will not be useful. Step d covers how to limit the irrelevant posts pulled by the software. But first, it will be necessary to identify some of the violent keywords present in the data.

It is possible to skip this step and just use the lexicon algorithm that has already been established. However, this is not advisable because there is a risk of missing many violent posts that don't conform to original predictions of the algorithm. For example, upon running the algorithm the implementer may find that a curse word is regularly misspelled. Thus, though the lexicon algorithm included the general phrase or word, there could be a significant amount of data missing. The way to course correct would be to add both the correctly spelled version and the commonly misspelled version so that the data are more fully captured.

This step can be overwhelming. However, it is recommended that the implementer take a small sample of posts (from different categories) and just read through them. Make a note of words or phrases that continue to come up in violent or abusive posts. It is important to set a limit—whether a time or post limit—on this step. Ideally, the implementer should read through posts until new words and phrases begin to repeat themselves, then add any new words and phrases to the lexicon algorithm.

d. Run Combined Name and Lexicon Algorithm

Given all the work we have done so far this step may seem especially difficult, but it won't be. This is the main step we use to get our VAWIE-Online data. Here, we are asking the software to identify posts that not only have the names we have identified, but also violent keywords and phrases. We are combining the amended lexicon algorithm from Step c and the names algorithm from Step a. Running the combined algorithm involves nearly the same steps as running the general algorithm. Here, the implementer will take a names algorithm (such as political male) and run it with the lexicon algorithm. For example:

- ("Barack Obama" OR "Joe Biden" OR "Donald Trump") AND (punch OR kick)

The implementer should use the "AND" in between the two algorithms so that the posts that are pulled out include a name as well as one of the lexicon words.
For implementers using an established software analysis, such as Crimson Hexagon, the algorithms should be run at the beginning of the study period or before. As mentioned, Facebook data is only collected when prompted; the implementer should begin collecting early on as it is not possible to look at past Facebook data. In Crimson Hexagon, Facebook data availability is dependent on the type of monitor that is being used. The FB Social Account Monitor will only allow for the implementer to access data from the day they started the monitor, and this would require them to do this early as recommended here. However, for use in the Targeted Content in a BUZZ or Opinion Monitor the historical data is available up to one year prior to the date of download. Note that changing regulations may restrict or increase availability of data over time. If the implementer is conducting a self-run software analysis, it may be necessary to continuously run these algorithms.

The implementer may find that the algorithm is still producing a lot of "noise" or unnecessary posts. If, for example, there are colloquial phrases that have violent keywords in them—but are not violent themselves—the implementer may choose to filter out for these phrases. In English, "you're killing me!" often means that someone is saying something funny. Because this phrase has the word "kill" in it, it could overwhelm the analysis, making it hard to find actual examples of violent posts with the word "kill."

In Boolean logic, it is possible to use AND- (AND followed by a dash or minus sign in front of it) to signify that the algorithm should take out any posts with a particular word or phrase (Capitalization of this operator is required and there should not be a space between the hyphen and the keyword to be excluded from results). For example, the implementer could write: ("Hillary Clinton" AND kill AND-"you're killing me"). In other cases, some words may generate so much noise that they must be removed entirely, even though they are considered important words for the study. This is acceptable: The VAWIE-Online Tool gives a broad (but not exhaustive) view of the presence of online violence and abuse, and the goal is to achieve the highest level of confidence in the data, not the highest number of individual social media posts.

Further, it may be found that one or several individuals receive so many posts that they overwhelm the sample and do not allow space to focus on other names. If this is the case, we recommend that the implementer removes these individuals from their names algorithm and run them in their own algorithm; they should be kept in the broader study but removed from the general algorithm to allow for other trends and patterns to be visible. This way, it will be easier to learn more about all of the individuals in the analysis.

VAWIE-Online gives a broad (but not exhaustive) view of the presence of online violence and abuse and the goal is to achieve the highest level of confidence in the data, not the highest number of individual social media posts.

The processes of reducing noise and tailoring the algorithm may continue throughout the analysis. In a broad look, as in Step c, there may not be a pattern of noise that is as clear as when posts are being trained. Also, there may be new violent keywords, misspellings, and themes that come up while training posts. The implementer can and should continue to update the algorithms to get the most useful information. However, they should make notes of what they are changing and when.

It should be noted that if self-run analysis is conducted, the implementer will not able to go back in time to look at posts from months and months ago. As such, they should go through these steps early, so that the model can be refined before the study period officially begins. If there are significant changes that are made, we recommend that the analysis is run as it was first written and to make changes in a separate algorithm so that the model can be perfected while still using consistent algorithms throughout.

e. Mine the data

If the analysis is being run through an established program, data mining is part of the step above and the implementer can skip this step.

f. Train posts

At this point, the implementer should have all of the posts that contain a name in their algorithm as well as a word or phrase from the lexicon algorithm. However, we are not done yet. Despite all the attempts to clean out unnecessary posts, there will still be many in the analysis. The implementer will have a list of violent posts but will not know how these posts are categorized. In this step, the implementer will train the posts, meaning they will instruct the software to recognize posts as belonging in one category or another.

Just as it is logistically impossible to read all political Facebook posts, it is not logistically practical to "hand code" or sort each post manually. Training allows the implementer to give the software formal and informal rules that it can use to categorize posts. It should be noted that if the implementer is not working in a language supported by the software being used, this may be less accurate. Further, if using a self-run program, machine learning may not be possible and, where it is, it may not produce valid results. As such, the posts will need to be hand coded and classified. A tool such as NVivo or Excel can be used to track these classifications.

It will be necessary to either scale down the analysis by limiting the number of names or organizations or select a random number of posts to hand code. The number of posts will depend on the amount of time the implementer has and the size of their team, but they should aim for at least 7,500 posts. Again, these posts need to be randomly selected from the various categories that have been collected. The implementer should not select 7,500 posts from the same category. It will not necessarily be able to generalize to the broader sample, but the implementer can make inferences and conduct reliability tests.

If using an established software, "training posts" refers to the process by which the implementer instructs the software to categorize posts (using an algorithm) for them. A crucial category will be "irrelevant." This is how the implementer will tell the software, or indicate for their self-run study, that this post is not relevant for analysis. In this case, an irrelevant post would not be a violent or abusive post.

Within the relevant posts, the posts can be classified as belonging to one of the three categories of violence: bodily harm, sexual and socio-psychological (a fourth category, economic violence, may be considered in distinct cases as discussed above). Not all posts will be clearly relevant or clearly belong to one category. Rules should be developed that will allow posts to be categorized so that the categories are consistent over time and so that the implementer can report their categorization method. Clear notes should be kept of how posts were categorized to contribute to consistency. This is also a place where colleagues can be called in to provide a second opinion. While there are parts of this process in which machine error can bias or invalidate results, this is a key place at which human error can contribute to issues with data output. If multiple coders are going through the data—or even if only one person is coding and classifying posts over time—very clear and explicit rules and notes should be kept. These notes should be shared with an additional person that tests the reliability of the sample, explained in the next step.

g. Test the Accuracy of the Algorithm

Once the posts have been trained, ideally just after an initial training has been performed, the implementer should run basic tests to determine the accuracy of their algorithm and training. There is no universal test as each country context will be different. Further, as noted in the limitations section, it will not be possible to ensure that the sample is complete or thorough. However, it is important to make sure that the data is internally reliable and valid.

Internal reliability is achieved when there is consistency across data. If the implementer finds that posts in one category are correctly categorized but posts in another category are not, they have not achieved internal reliability. Validity is achieved when the implementer is measuring what they want to be measuring. To do this, they should select a sample of posts (if they are using an established platform or machine learning, not the ones that were trained by the implementer themselves) and make sure that they are examples of violence and in the correct category. If it is found, after reviewing a number of posts across categories, that these issues are not present or are insignificant in number, the implementer can confidently move forward with the analysis. If, however, issues are found with either

internal reliability or validity, the algorithms should be adjusted, more posts should be trained, or a combination of the above until there is confidence in the data. There will always be outliers; ensuring that there is perfect data is not feasible. But, if the implementer finds that 70 percent or higher of their posts are correctly assigned, the model is acceptable.[42] Once the model is completed, they should go through and determine what percentage of their posts are correctly assigned and report this number. If the sample was hand-coded, the reliability of the data can be tested by selecting a small amount of additional, random data and seeing whether the proportions that were determined in the primary sample of data match those in a new sample. It should not be necessary to test the validity of the sample, because the implementer coded all of the posts themselves.

Below in Step j are our recommendations for analysis that will best ensure a valid and reliable sample.

h. Export the Results

Depending on which platform is being used, the data may be saved on a website or in a program. Because the results will be big files, programs may have limitations, and it will be important to ensure that there is documentation of the data, so the implementer should always export or save the results in a separate location. We recommend utilizing Dropbox or Google docs, which are (up to a point) free data storage servers and offer strong security. Further, Dropbox or Google docs can be used to share files with colleagues. Colleagues should be reminded not to make any changes to the data files, as seemingly small changes can have significant consequences on the data.

The results should be exported and the data should be backed up several times a week at a minimum. When an analysis is finished, the data should also be exported at that time. If the program being used is prone to crashing, such as NVivo, or for computers with limited RAM, the data should be backed up and exported even more frequently. This is another step of the process in which the implementer should ensure that the data is being securely handled and that privacy is protected.

i. Repeat for All Algorithms

Each group of individuals and organizations constitutes one set of data, which is then categorized into relevant/irrelevant and violence categories. At this point, the steps above should be repeated for each set of names and organizations. The first set is the most difficult and as the analyses are repeated, and the implementer's comfort with the platforms increases, each subsequent analysis will be easier.

Please note that each analysis will need to be run before the start of the analysis period to work out problems and start the process of collecting data. For self-run software, it will be necessary to periodically run the algorithm again, even up to once a day. The implementer should ensure that there is enough pre-analysis time to construct and perfect their algorithms.

[42] This is lower than statistical modelling. Social media studies, by design, have significantly lower validity than other studies. Achieving a 70 percent benchmark is thus acceptable, but should always be reported. For cases where the implementer is unable to achieve adequate levels of confidence, it is not recommended to use quantitative data for public reporting on frequency, but instead to conduct a qualitative manual analysis of select content to discuss factors such as the nature and impact of threats and abuse.

Though these steps, and the documentation of these steps, is time-consuming, the more accurately and purposefully the analysis is conducted, the easier it will be to interpret and utilize the results. We will continue with this process in Step 3.

j. Recommendations

There is no perfect tool or system for conducting VAWIE-Online analysis, as each has its limitations. However, there are significant differences between tools. To limit the disadvantages of any single tool, below are recommendations for how to construct the VAWIE-Online analysis, in order of preference:

1. Utilize an established program, like Crimson Hexagon. The implementer should conduct their analysis with either (a) a rigorous test of validity and reliability, making changes as needed or (b) a secondary hand-coding component. Using the latter method, the program would be used to filter relevant posts and then export the relevant posts to hand-code them, ensuring that those that remain are truly measuring abuse and/or violence. This would require that the implementer create their own graphics rather than rely on the program's graphics and analysis.
2. During the creation of the algorithm, and once the analysis is concluded, conduct a test of validity. If it isn't possible to make the algorithm more precise (as recommended), the implementer should at least aim for the 70 percent benchmark and report the validity of the sample in their reports. If this is not achievable, the statistical data should not be reported, but the implementer may choose to manually examine some content to produce qualitative analysis.
3. Utilize self-run programs to scrape and/or mine data and then hand-code a large number of a random selection of posts.
4. Utilize self-run programs to scrape, mine, and analyze the data, without hand-coding. If the data is less than 80 percent accurate, the implementer should not utilize this data and instead should rely on hand-coding.

Running Self-Run Programs

a. Write General Algorithm

This is generally the same as same process as with Crimson Hexagon. A key difference for self-run programs is that it will only be possible to pull the data from recent posts, usually within the last several days (one of the reasons established programs are costly is because they provide access to large amounts of data). Also, when conducting a self-run analysis, the implementer can utilize the different ways in which a public figure presents themselves online (i.e., Twitter username, public Facebook page, nicknames, etc.) more easily.

b. Scrape Raw Data

This process will look different if the implementer is utilizing a self-run program. If using #TAGS or another program which pulls data from Twitter, it will be beneficial to set up an application on Twitter, which will provide individualized passwords that permit the implementer to pull Twitter data. To do so,

visit Twitter's developer page.[43] This is free to access and use. The implementer will need to answer a few questions in order to gain access to developer permissions (as of July 2018). Once they are granted developer permissions, they will be given four passcodes, which can be used to gain access to large amounts of Twitter data. No matter which software is chosen, there are resources available that can help with this and future steps. Established programs have large help sections and often, dedicated programmers that can be emailed with specific questions.

c. Review Data and Amend Lexicon Algorithm

See instructions in section above (same process as for Crimson Hexagon).

d. Run Combined Name and Lexicon Algorithm

If the implementer is using a self-run program, this step can be skipped.

e. Mine the Data

If a self-run program is used for analysis, the implementer will need to look through the raw data for violent words and phrases. The programs that are used to scrape data—namely #TAGS and NVivo NCapture web scraper—cannot also analyze the data. To do this, the raw data must be imported into a program that can conduct a data mining analysis.[44] For NVivo NCapture, the data will be prepared for importation into NVivo for analysis. NVivo has sentiment analysis capabilities, word search capabilities, tagging (or categorizing) capabilities, and machine learning. Though this will require a bit more technical work than in an established platform, it is largely user-friendly, though will be easier to navigate for English-language speakers. NVivo is available in English, Chinese (simplified Chinese), French, German, Japanese, Portuguese and Spanish.

For Twitter data, #TAGS will compile the raw data into a Google sheet, which is easily downloaded as an Excel file. From here, it will be possible to reformat the data, import the data into another program to mine, and categorize. An Excel sheet has the capability to search for words and phrases but will not allow the implementer to effectively mine, conduct a sentiment analysis, or utilize machine learning to train posts. For that, it will be necessary to import the file into a separate program, like R or Python. R and Python are almost like blank webpages with programs inside of them. Depending on what analysis will be run, the implementer will need to launch different programs within R and Python. To use these programs, they will need to write a code or language that is specific to the program. There are many examples of code online, which can be used to start the process. Through R, Python, and NVivo, images and word clouds can be created that are helpful for the presentation of results.

f. Train Posts

Most self-run programs either do not have the ability to train posts or have limited capacity. Thus, data cleaning will have to occur manually.

---

[43] Here is the link to Twitter's developer page: https://developer.twitter.com/en/apply/user

[44] #TAGS is able to combine data scraping and mining but cannot handle as large an algorithmic request as Crimson Hexagon. As such, it may still be necessary to export #TAGS data for mining.

g. Test the Accuracy of the Algorithm

Unlike Crimson Hexagon, self-run programs require much more manual involvement for this step. Algorithm testing will not be automated.

h. Repeat for All Algorithms

Step 3: Analyze results

If the analysis has been successfully run, the data will already be categorized by group of people/organizations as well as by type of violence. Now is the time for the implementer to consider the types of questions they would like to answer and to use the data that they have to answer it. The questions in Part I can be referenced as well as the implementer's own questions. It should be noted that if a self-run program is used, there will be greater limitations in the questions that can be asked and the conclusions that can be made; however, by using a sufficient number of randomly selected posts and hand-coding them, the implementer will be able to make some inferences based on their selection of data.

It is also important that the social media analysis is connected with the lived experiences of individuals. As stated at the outset, this self-analysis tool only enables us to capture the tip of the iceberg of VAWIE-Online; much (or most) of this violence occurs in the private spaces of the internet that are invisible to researchers. Therefore, it is recommended that the key findings are confirmed through focus groups discussions. It may also be beneficial to conduct one-on-one interviews with individuals who have had particularly bad experiences with VAWIE-Online. Likewise, talking to the people that have experienced this can help the implementer understand how the trends that we can trace in the public domain are connected and interwoven with other incidents happening in private online space and/or in the real world.

Possible Analyses

Frequency of different types of violent sentiment (large-scale quantitative analysis)

Compare the different categories of violence with each other. The implementer can compare them as they impact men and women or as they impact only women, and can also look at differences over time—the analysis may indicate peaks and valleys in dates, possibly based on political or social events.

Frequency of different types of targets of violence (large-scale quantitative analysis)

Compare the different groups of targets of violence (political, media, etc.) with each other. Again, it is possible to collapse the male and female categories, compare men against women in each category, or compare women across all categories. The implementer can also look at differences over time, as the analysis will indicate peaks and valleys of frequencies as well.

Identify impact of a post (small-scale quantitative and qualitative analysis)

Not only can we determine whether a post is violent or abusive, we can also look to see how often it has been retweeted and infer how many people have seen the post. When we are considering the impact of VAWIE-Online, this is crucial. However, at the same time, a retweeted post may inflate our findings. If it appears that on a particular day there are twice as many counts of violence as on surrounding days, and this increase is the result of only one single tweet or post that has been shared, this should be noted. This does not negate the importance of that post but should note that this increase has not been caused by many negative posts when only one post has been particularly impactful.

Identity of perpetrator (large- and small-scale quantitative and qualitative analysis)

Information about perpetrators may be relatively limited and the implementer may decide to establish internal ethics guidelines to only examine collective data (not individual perpetrator information). If this is the case, they can look for profiles that come up repeatedly and control for them (either analyze separately or remove from analysis). It may also be beneficial to look to see if users come from outside the country or within the country. The implementer can limit their analysis to in-country users—for those who report their location—or separate out external perpetrators to note the impact of expats, migrants, and international actors. Finally, if an account has no followers, it may be a bot, or a robot account. In a self-run analysis, it is possible to control for these accounts. While questions of bots persist on issues of VAWIE Online, the impact of violent rhetoric may be significant whether posted by a human or bot account. Implementers can look for individual accounts that get many re-tweets, as these may have a larger impact, regardless of the perpetrator.

Identify specific patterns of online abuse (small-scale qualitative analysis)

The lexicon that has been established can be used to identify particular patterns of violent rhetoric. Going deeper than the broad categories, implementers can look for instances of online violence that include slut-shaming, doxing, swatting, and deadnaming. If, for example, one of the individuals being tracked is a transgender candidate, a separate analysis may be run for their former name (deadnaming) or for inaccurate pronouns. Qualitatively, the implementer will be able to note the way in which VAWIE-Online exhibits these forms of violence.

Speed of the spreading of online abuse (large- and small-scale quantitative and qualitative analysis)

Pilot tests of the self-run analysis tool suggest that the speed and intensity with which online violence spreads vary significantly between men and women, as well as between women. Women may “go viral” more quickly because they are constrained by more social limitations that can be manipulated for hostile purposes online. The speed with which hateful messages spread can cause excessive negative impacts.

Distinguishing degrees of intensity between incivility and abuse (small-scale qualitative analysis)

Not every negative tweet is violent. There is a scale of intensity within the content of VAWIE-Online data. For example, "You look like a pig" is of lower intensity than "I want to kill you." Intensity cannot be computed in sentiment analysis so it may be useful to explore this range manually, using frameworks suggested by authors such as Southern and Harmer.[45] Also, implementers may want to use VAWIE-Online data to examine how frequency can amplify intensity—in other words, a single tweet of "you look like a pig" may be low intensity on its own, but if it is shared 10,000 times and amplified with images and other insults, it could conceivably cause very serious harm to its targeted survivor. Also "low intensity" content is often connected to escalating, "high intensity" abuse and even real-world violence. This is also something that manual analysis of VAWIE-Online data can examine.

Considerations in Presenting Data and Findings

Social media analytics are very powerful tools that can be exciting to work with. However, there are also many pitfalls, limitations, and constraints. It is important for the analysts to exercise caution and common sense, to apply strong research standards including checking and double-checking the quality of the data, and to adapt and apply ethical guidelines. Social media analytics generate vast amounts of information and it is the responsibility of the analysts to determine what inferences can and cannot be drawn from the findings. Because social media analytics give a general impression (not an absolute measurement) of online abuse, it is especially important to take care in how statistical findings are reported. The following section discusses some of these common challenges and how to address them.

Common Pitfalls

There are many possibilities from this research, but also potential limitations. During five pilot tests, several common pitfalls appeared. Some of these limitations can be corrected by purposeful methodology and reporting. Others, as will be discussed in the next section, are endemic to the analysis of VAWIE-Online.

[45] Rosalyn Victoria Southern and Emily Harmer, "A Gendered Analysis of Uncivil and Abusive Tweets Sent to MPs in the UK," (August 30, 2018), https://convention2.allacademic.com/one/apsa/apsa18/index.php?cmd=Online+Program+View+Paper&selected_paper_id=1384545&PHPSESSID=1r36jkscvrrd45o2ttpefsmdq2.

Proportionality

If the sample does not include an equal number of men and women, it will be necessary to ensure that the numbers are proportional in reporting their data. If the implementer is comfortable with statistics, a t-test can be used. Another simple method is to weight the data that the implementer wants to compare. For example, if the sample has four men but two women, the implementer will need to double the results for women in order to have proportionally comparative data. Because of this, data in the sample is not equivalent to universal data or data in the country.

Grounding AI tools in human feedback loops

Although AI is improving every day, there is no replacement for a human touch. VAWIE-Online Analysis should connect with the community throughout the course of a study including by:

- Working with local data analysts and using local languages
- Conducting pre-analysis focus group discussion (FGD) to tailor model and define lexicons
- Conducting post-analysis FDG to triangulate findings, identify connections between private and offline violence, analyze impacts and responses

As mentioned, the implementer will not be able to write "women in this country receive four times as much violence due to their gender." Instead, it would be appropriate to write: "Women in our sample received four times as much sexual violence as men. Sexualized tropes may provide an explanation for this finding." Similarly, if there are only 50 women candidates in a sample and 150 men, and the data show that 80 percent of the content women receive has a negative sentiment while 40 percent of the content men receive has a negative sentiment. Instead of writing "women receive twice as much negative content online compared to their male counterparts," it would be more useful and nuanced to write, "despite being largely invisible online and making up a significantly smaller proportion of online users—likely due to phenomenon such as 'the silencing effect'—women candidates receive twice as much negative content than their male counterparts."

Causation

Following from above, implementers should be careful about the way they present their results and the implications for causation and motivation. From qualitative data, such as interviews, we can better understand the likely causal mechanisms and motivations behind violence. From large-scale data, we can infer information about the drivers of violence and how different categories impact individuals differently.

Images

Most of software tools can produce graphics and images. Implementers should be sure to use these responsibly. They should explain what is in the images and make sure that dates are consistent, particularly if using separate self-run programs. If the data was amended, implementers should make

sure that the up-to-date data, particularly the more valid and reliable data, is represented in the final images.

Clear Methodology

Implementers should make a point to identify and clarify their methods. How was the lexicon determined? Were changes made? How valid were the findings? These methods make it easier for future studies to learn from this work and to interpret these findings. In addition, a clear methodology may assist with clearer policy outcomes, which will be addressed in Part III.

Double Negatives and Sentiment

As mentioned, sentiment analysis is a limited tool. Determining whether posts are positive, negative, or neutral can direct us toward violence, but is not a replacement for classifying posts as violent and organizing them by category or subcategory. For example, "I can't wait to see her thrown to the ground" would likely be classified as positive, as it expresses excitement, even though the post author is advocating violence.

Typos

In some cases, perpetrators accidentally misspell words, making it difficult to find accounts of violence. In others, perpetrators purposefully misspell or vary words to evade detection by social media algorithms. Regardless of the intention, it can be difficult to capture violent posts utilizing a specific set of words when misspellings and typos are common. For example, "I hope she perishes" is a violation of most social media rules, while "I hope she parishes"[46] is not. Though there will always be violence that remains uncaptured, implementers can use steps to capture some misspellings, nicknames, and new violent words.

Other Considerations – Potential Limitations and Constraints

The suggestions above can reduce potential issues with data scraping and analysis. However, there will inevitably be some data limitations; no study of online behavior can capture all posts. Being aware of some of these broader limitations can assist implementers in managing expectations and being aware of what they cannot do.

Publicly Available Data

Implementers can only scrape for public data. This means that private messages are not included in the analysis. Further, even if access to private messages was obtained, there may be ethical and legal concerns in including these posts in an analysis. Any posts that have been removed for violating terms of service will not be incorporated into a data scrape. This means that any conclusions will inevitably be skewed toward more positive data as any direct threats should be removed from the program before they can be captured. If it is noted that there is a lack of direct, physical threats in the data sample, this could be part of the explanation.

[46] Parish is an English term for a district, usually a religious district of a Church. The latter sentence does not make grammatical or content-based sense as written but is pronounced the same way as the former, violent sentence.

Photos and Images

Self-run programs are better at capturing images than other software, but none are geared toward photographs and images. A Crimson Hexagon image recognition feature was released shortly before publication, for additional cost. If pictures are a large part of VAWIE-Online in the country context, an additional analysis focusing on images may be important. Image analysis was rapidly developing during the period of research and new software is expected to become increasingly available and affordable.

Perpetrator Identities

It can be difficult to determine the identities of perpetrators. Identifiers such as age, gender, location, etc. are not always available or verifiable online. Many users do not provide this information and when they do, it may not be reliable. Very rarely, a user will leave their "geotag," or geographical tag, on. In this case, the location of the user can be pinpointed. If an identity is verified or the user has a relationship to the survivor, this data is particularly important to include as it allows the implementer to make unique inferences about the drivers of violence. It is important to note that perpetrators also have privacy rights, not just their targets. Data on perpetrators should only published in an anonymized, aggregate format and should not directly reference individual online users.

Distinguishing Abuse, Violence, and Other Damaging Posts

Misinformation about female candidates, which is rife online, may not be classified as abuse or harassment but still may be very damaging and worthy of study. This can include creating fake profiles and distributing false information or images about individuals. Further, not all phrases will have violent keywords. Step D is particularly important in this respect and will allow for a more precise algorithm.

Difficulty in Establishing a Motive

It is difficult to establish motivations of any form of violence, let alone violence for which the perpetrator is nearly always anonymous. Instead, we can make inferences based on the information that is available and use categories and targets of violence to infer motivations on a sample-wide level.

Software has Limitations

As mentioned, there is no perfect software tool. Some of the above constraints and limitations are inherent in the available software tools. There are pros and cons to all tools. Utilizing a precise and highly documented methodology can assist in mitigating some of the limitations of software tools, allow for improvement over time, and enable others to read, review, and reconstruct the analysis.

Part III. Using the Data – Stakeholder Responses

Violence during elections has serious and potentially irrevocable implications for democracy. Violence can render a candidate unable to stand for office, encourage a politician to leave office or quit politics, discourage potential candidates or aspirants from running in future elections, alter voters' perceptions of candidates, and keep individuals from voting, volunteering for campaigns, or otherwise engaging in the political process. Online violence can lead to these consequences just as in-person violence does. But, the spread of social media and online tools allows us to better understand online violence all around the world. Numerous stakeholders can play a role in tracking online violence, in order to respond and stop it from taking place.

Election Management Bodies

Electoral institutions are in a unique position to track incidents and intervene. Election management bodies (EMBs) hold the list of candidates in an election; this information may not be available to others. EMBs can use that information to pull through data and mine it for incidents of violence. In addition, EMBs can work with candidates to collect sources of information. As a central organization for all candidates, EMBs can offer a place for candidates to share any violence they experience, whether online or in-person. Working with executive agencies and legislative bodies, EMBs can work to draft, pass and enforce legislation to protect individuals from online violence.

Lawmakers

Online violence against women is overlooked in most legal frameworks, but can be addressed through targeted legal reforms. According to scholars Barker and Jurasz, "Prevention of and accountability for online violence against women needs to be prioritised [sic] within multiple legal frameworks—preferably within those frameworks dealing with issues concerning online communications and the governance of online space. Online violence against women needs to be recognised [sic] as a form of gender-based abuse of women and girls as well as a factor standing in the way of their full and equal participation in public and online spaces. Gender should be incorporated as a protected characteristic in hate crime legislation […] creat[ing] legal avenues for addressing misogyny as a hate crime but also allow[ing] compliance with existing equality legislation at national and international levels."[47]

Law Enforcement and Security Sector

The law enforcement and security sector can crucially protect candidates experiencing violence. Analyzing online violence can serve as a tool to track and act on threats to candidates before an act of physical violence takes place. Though not every instance of online VAWIE constitutes a credible threat of violence, some do. Security-sector actors can help candidates manage their online privacy (teaching candidates not to share their home locations, for example) and step in to offer digital and physical

[47] Olga Jurasz and Kim Barker, "Online Violence Against Women: The Limits & Possibilities of Law" (Stirling Law School and Open University Law School, 2018), http://oro.open.ac.uk/53637/1/OVAW%20-%20The%20Limits%20%26%20Possibilities%20of%20Law%20%282018%29.pdf.

protection in the wake of online violence. Importantly, security services will often enforce legislation against online violence.

Political Parties

Like EMBs, political parties can serve as important sites of data collection, asking their members to share incidents of violence. Parties can pass internal policies to prevent and condemn violence. This is important, as online violence may be instigated by politicians themselves, including within their own parties. Political parties can lobby for legislation against online violence, civil or criminal consequences for perpetrators, restitution for targets of violence, and increased oversight of the online space. Online violence is an issue likely affecting all political parties. It is in the interest of all parties, therefore, to work toward solutions.

Media

The media can publicize stories of online abuse, particularly with the support of the target of abuse. Members of the media may also find themselves subject to harassment and violence. It is important not only to track and report on online harassment against formal politicians and candidates, but against other electoral stakeholders as well. The media can keep up pressure on legislative bodies to write and pass legislation and can use source data to back up claims of violence. Journalists can be trained to recognize common forms of VAWIE-Online, including those involving artificial intelligence.

Social Media Companies and Internet Governing Bodies

Social media companies have a unique advantage—they can collect, remove, and report to authorities any information that is transmitted on their platform. Social media companies should work with political leaders to ensure that the goals of internet freedom and safety from harm are not mutually incompatible. In terms of solutions, political legislation can provide important leverage and encourage a long-term response from social media companies. Nonetheless, social media companies should be motivated to make their platforms as safe as possible for all users. Companies have the responsibility and ability to enact the most wide-reaching changes to combat online violence and do not need to wait for legislation to do so. Companies can update and enforce codes of conduct, shut down offending accounts, prevent perpetrators from opening new accounts under different names, and utilize in-country individuals to monitor posts for contextually inappropriate language. Internet governing bodies will enforce legislation and sanction noncompliance.

GBV Service Providers

VAWIE-Online can be understood as a subcategory of gender-based violence. The same patriarchal roots feed both gender-based violence broadly in society and online violence against women in politics. GBV service providers have knowledge on how to best support survivors of violence as well as to design gender-sensitive and survivor-driven policies. This expertise should be solicited by other key stakeholders.

Advocates and Civic Educators

VAWIE-Online is poorly understood or entirely overlooked in many countries in the world. The findings from VAWIE-Online analyses can provide a base of evidence for civic educators and advocates to develop campaigns, offer facts and examples and raise awareness about the prevalence of this problem, its impacts, and advocate for responses.

Real-time Response

The social media analytics tool can be used for real-time reporting, which can feed responses by various actors, including law enforcement, EMBs, advocacy groups, and responsible journalists. To offer help at an individual level, the tool can also be paired with initiatives that support GBV survivor services and assistance.

RESPONSE EXAMPLE: Ending Impunity for VAWIE-Online

As part of a larger program in Zimbabwe, IFES worked to end impunity for online and VAWIE during the 2018 elections. Through support from Global Affairs Canada and USAID, local women's legal associations and civil society advocacy groups were trained to recognize all forms of VAWIE, including VAWIE-Online identified with this tool. With support from IFES, local organizations helped survivors bring their cases into the justice system and prosecute perpetrators. National Police were educated on the topic and committed to fight it. In the course of a short period, the program resulted in over 100 cases documented, dozens of charges pressed, multiple arrests and at least two successful prosecutions.

While this type of approach is important, there are many legal gaps around online violence which make it more difficult to prosecute. Understanding and advocating for legal reform to combat VAWIE-Online is a major challenge today.

RESPONSE EXAMPLE: VAWIE-Online for Rapid Response

VAWIE-Online Sentiment Analysis can reveal long-term trends and patterns, but it can also be used in real time to support active responses. For example, in Zimbabwe and the Ukraine, the VAWIE-Online analysis teams examined data on a daily basis during the immediate period around the elections and reported key findings to stakeholders, including security actors, media, election observers and rights advocates.

RESPONSE EXAMPLE: Bystander Intervention

The first line of action is to stop VAWIE-Online when and where it occurs. Using the approach of "bystander intervention," IFES works with the non-profit organization Hollaback! to adapt their award-winning platform HeartMob to help stop VAWIE-Online. HeartMob responds to online harassment by activating a supportive community, helping reduce trauma for people being harassed online by giving them the immediate support they need. Learn more about HeartMob at www.iheartmob.org and www.ifes.org/vawe

Part IV. Conclusion

Social media has transformed politics, opening space for dialogue and creating new pathways for citizen involvement and education. Political engagement has also been plagued by the misuse and abuse of new information communication technologies (ICTs), from rising incivility, disinformation, hate speech and direct violence. Until now, the significant impact of these trends on women's political engagement has yet to be thoroughly investigated.

VAWIE-Online is a threat to democracy as it hinders the right of women to safely and freely participate in elections. With the help of resources such as the VAWIE-Online Tool, steps can be taken in addressing online violence against women. We need to understand the scope, shape and impact of this new form of violence in order to prevent and end it. The VAWIE-Online Tool can be used by actors from across a range of professions who are concerned by hateful and violent speech online and are motivated to end it. Although new technologies are being used to inflict harm, they also offer vast opportunities to combat it. New technology such as AI helps us understand the challenges we face and come together to speak out against injustices and inequality. Harnessing these tools for development and democracy protects the freedom to speak freely online without fear.

Annex 1: Online Violence Terms

Cross-platform Harassment

When a harasser, or group of harassers, deliberately sabotages or invades multiple online spaces for the purposes of harassing a target. Cross-platform harassment is very effective because users are currently unable to report the scope and context of the harassment when they contact platforms, each of which will only consider the harassment happening on their own sites.

Cyber-exploitation, Nonconsensual Photography or "Leaking Nudes"

The distribution of sexually graphic images without the consent of the subject of the images. The abuser obtains images or videos in the course of a prior relationship, or hacks into the survivor's computer, social media accounts or phone. Women make up more than 95 percent of reported survivors. The unauthorized sharing of sexualized images is still not illegal in the majority of U.S. states. Twenty-two states now have laws on the books and proposed national legislation is being drafted

Deadnaming

A form of direct harassment in which a target's former name is revealed against their wishes for the purposes of harm. This technique is most commonly used to out members of the LGTBQIA community who may have changed their birth names for any variety of reasons, including to avoid professional discrimination and physical danger.

Defamation

Coordinated attempts at defamation take place when a person, or, sometimes, organized groups, deliberately flood social media and review sites with negative and defamatory information.

DOS

DOS stands for "denial-of-service," which is an attack that makes a website or network resource unavailable to its users.

Doxing

The unauthorized retrieving and publishing, often by hacking, of a person's personal information, including, but not limited to, full names, addresses, phone numbers, emails, spouse and child names, and/or financial details. "Dox" is a slang version of "documents" or .doc. Causing fear, stress and panic is the objective of doxing, even when perpetrators think or say that their objective is "harmless."

Electronically Enabled Financial Abuse

The use of the internet and other forms of technology to exert financial pressure on a target, usually a woman involved in intimate partner abuse. This might include, for example, denying access to online accounts, manipulating credit information to create negative scores and identity theft.

False Accusations of Blasphemy

Women face online threats globally, but they run a unique risk in conservative religious countries where blasphemy is against the law and where honor killings are a serious threat. Accusing someone of blasphemy can, itself, become an act of violence.

Flaming

A flood of vitriolic and hostile messages including threats, insults, slurs and profanity.

Gender-based Slurs and Harassment

Name-calling is common online. Gendered harassment, however, involves the use of words, insults, profanity and often images to communicate hostility toward girls and women because they are women. Typically, harassers resort to words such as “bitch,” “slut,” “whore,” or “cunt” and include commentary on women’s physical appearances.

Google Bombing

The deliberate optimization of malicious information and web sites online so that when people search for a target they immediately see defamatory content. In 2012, for example, Bettina Wulff, the wife of Germany’s then president, sued Google because the company’s autocomplete search function perpetuated rumors that she was once a prostitute.

Grooming and Predation

Online grooming is when a person using social media to deliberate cultivates an emotional connection with a child in order to sexually abuse or exploit that child.

Hate Speech

Hate speech has no uniform legal definition. Online, this means that every social media platform has its own unique definition. As a baseline, however, hate speech is language or imagery that denigrates, insults, threatens, or targets individual or groups of people based on their identity—gender, race, color, religion, national origin, sexual orientation, disability, or other traits. There is no hate speech exception to the First Amendment. Hate speech usually has specific, discriminatory harms rooted in history and usually employs words, action and the use of images to deliberately shame, annoy, scare, embarrass, humiliate, denigrate, or threaten another person.

Most legal definitions of harassment take into consideration the intent of the harasser. This, however, fails to translate usefully in the case of cyber-harassment. In the case of technology-enabled harassment and abuse, intent can be difficult to prove and defuse. For example, most laws do not currently consider third-party communications to be harassing. Whereas sending someone a threatening message for the purposes of extortion is illegal, there is no law that addresses the non-consensual sharing of sexual images to someone other than the subject of the photograph illegal or hateful.

Identity Theft and Online Impersonation

As defined by the U.S. Department of Justice, identity theft includes “crimes in which someone wrongfully obtains and uses another person’s personal data in some way that involves fraud or

deception, typically for economic gain." The law applies to any person or entity who impersonates another person on the internet with the "intent to obtain a benefit or injure or defraud another." Many states distinguish this from impersonation, in which a person creates an account, website or ad using a person's name and address with the intention of harming another person.

In 2013, for example, a jury convicted 32-year-old Michael Johnson of more than 80 counts related to his having impersonated his ex-wife online. He had purchased online ads and connected with would-be johns, posing as his wife. He posted rape fantasies inviting men to kick down her door and rape her. In addition to sharing prices for sex with her, he also included prices for sex with her three daughters as well as for the toddler boy that the couple had together. The abuse continued when he contacted one of the daughters' schools, posting a message to the school's website in her name, reading, "I will have sex with the teachers in return for passing grades." As many as 50 men a day showed up at the woman's home. She eventually moved her family to another state. Other similar impersonation cases involving false fantasies of violent gang rape are commonly used as part of ongoing intimate violence. The difference between these two tactics is that identity theft benefits the perpetrator, while impersonation results in a distinct harm to another person.

In Real-Life Attacks

In real-life (IRL) attacks describe incidents where online abuse either moves into the "real" world or is already part of an ongoing stalking or intimate partner violence interaction. IRL trolling can also mean simply trying to instill fear by letting a target know that the abuser knows their address or place of employment.

Mob Attacks/CyberMobs

Hostile mobs include hundreds, sometimes thousands, of people, systematically harassing a target. #Slanegirl, a hashtag that was used for the trending global public shaming of a teenage girl filmed performing fellatio, is one example.

Rape Videos

Videos of rapes in progress that are subsequently used to shame or extort or are sold as nonconsensual pornography. These images are sometimes used to populate online spaces created expressly for sharing them, cyber-cesspools whose sole purpose is to deprive people of dignity by humiliating and harassing them. In India, rape videos are part of what law enforcement has described as a thriving revenge-porn economy. They are used to blackmail, shame and extort. The U.S. and U.K. have seen multiple publicized cases of teenage girls, whose rapes were filmed and shared, commit suicide.

Retaliation Against Supporters of Survivors

Online abusers will often threaten to or engage in harassing their target's family members, friends, employers or community of supporters.

Sexual Objectification

Harassers frequently objectify their targets, including through the use of manipulated photographs and sexually explicit descriptions of their bodies. Photographs of girls and women are often used without their consent and manipulated so that they appear in pornographic scenes or are used in memes.

Shock and Grief Trolling

Targeting vulnerable people by using the names and images of lost ones to create memes, websites, fake Twitter accounts or Facebook pages. Feminist writer Lindy West has described how harassers set up Twitter accounts using a stolen photograph of her recently deceased father. The name on the account was a play on his name and a reference to his death. “Embarrassed father of an idiot,” the bio read. It cited his location as, “Dirt hole in Seattle”.

Spying and Sexual Surveillance

Most people think of spying and surveillance in terms of governments spying on citizens, however, women are frequently illegally surveilled. This happens in their apartments, changing rooms, department stores, supermarket bathrooms, on public stairways and subway platforms, in sports arenas and locker rooms, in police stations, and in classrooms while they teach. The euphemism “Peeping Tom” is particularly insufficient given the nature, scale and the internet amplifies the power of stolen images and recordings to be used in harmful ways.

Stalking and Stalking by Proxy

Justice Department records reveal that 70 percent of those stalked online are women and more than 80 percent of cyber-stalking defendants are male.

Sexting/Abusive Sexting

Sexting is the consensual electronic sharing of naked or sexual photographs. This is different, however, from the nonconsensual sharing of the same images. While sexting is often demonized as dangerous, the danger and infraction is resident in the violation of privacy and consent that accompanies the sharing of images without the subject’s consent. For example, while teenage boys and girls sext at the same rates, boys are between two and three times more likely to share images that they are sent.

Slut-Shaming

A form of gender-based bullying often targeting teenage girls. Slut-shaming, stalking, the use of nonconsensual photography, and sexual surveillance frequently overlap, amplifying impact on targets. Amanda Todd, Rehtaeh Parsons, Audrie Potts, Felicia Garcia, Tyler Clementi, Rachel Ehmke, Steubenville’s Jane Doe and Jada are people who were targeted by combinations of these tactics.

Swatting

Deliberately tricking authorities into responding to a false emergency at a specific address. The term comes from “SWAT” (Special Weapons and Tactics), a branch of the U.S. police that uses militarized techniques, equipment and firearms to breach targeted sites. Harassers will report a serious threat or emergency, eliciting a law enforcement response that might include the use of weapons and the possibility of someone being hurt or killed.

Threats

Rape and death threats frequently coincide with sexist, racist commentary. While online threats may not pass current legal tests for what constitutes a “true threat,” they do generate anxiety and alter the course of a person’s life.

Trafficking

While not traditionally thought of as a form of online harassment and abuse, trafficking involves multiple types of electronically enabled abuse. Social media is used by traffickers to sell people whose photographs they share, without their consent, often including photographs of their abuse of women as an example to others. Seventy-six percent of trafficked persons are girls and women and the internet is a major sales platform.

Unsolicited Pornography

Sending unsolicited pornography, violent rape porn gifs or photographs in which a target's photograph has been sexualized. For example, in 2003 United Nation's Development Fund for Women (UNIFEM) website was stolen online by a pornographer who populated the site with violent sexual imagery. More recently, editors at Jezebel, an online magazine, reported that an individual or individuals were posting gifs of violent pornography in the comments and discussion section of stories daily. Writers at Jezebel, almost all women, were required to review comments sections daily. Women politicians, writers, athletes, celebrities and more have their photographs electronically manipulated for the purposes of creating non-consensual pornography and of degrading them publicly.

Source: Women's Media Center: http://www.womensmediacenter.com/speech-project/online-abuse-101

Annex 2: Data Analysis Key Terms

Algorithm: A process or set of rules to be followed by a particular software in order to generate data.

Artificial Intelligence: The ability of a computer program or a machine to perform tasks that typically require human intelligence. This may include speech recognition, visual perception, and translation between languages.

Boolean Logic: The idea that all values are either true or false. This logic is centered around the following Boolean operators: "Or," "And," and "Not." These Boolean operators can be used as connectors between key terms in order to define the parameters for an algorithm.

Data Mining: The process of examining large databases in order to find patterns correlations and anomalies that can allow for the generation of analysis and the prediction of outcomes.

Data Scraping: The process of importing information from a website into a spreadsheet or a local file saved on one's computer.

Machine Learning: A branch of artificial intelligence that stems from the idea that software systems can learn from data, identify patterns and make decisions with minimal human intervention.

Opinion Monitor: Built by crimson hexagon to measure the general broader sentiment (positive, negative or neutral) of posts present on social media.

Sentiment Analysis: The process of systematically identifying and categorizing opinions expressed in a piece of text in order to determine whether these opinions are positive, negative or neutral.

Validity: The extent to which data findings are reliable and generalizable.

Annex 3: VAWIE-Online Social Media Analysis Tool – Worksheet for Setting Up a Study

1. Identifying Categories of Potential Targets to Monitor

These three steps will guide the selection of potential targets. The suggested categories should be tailored to local context.

Identify Categories of Potential Targets

Tick the boxes for the categories appropriate to your study.

I. POLITICAL		II. INSTITUTIONAL	
Category	*check here to include*	*Category*	*check here to include*
Candidates from both opposition and incumbent parties		EMB permanent staff	
Elected officials from both opposition and incumbent parties		Poll workers	
Political aspirants (i.e., seeking nomination)		Police and Security Forces	
Political staffers		State administrators civil servants	
Party members and supporters		Other? _____	
Political appointees in government (cabinet ministers, etc.)			
Political parties and subsidiaries women' wings, youth wing, etc.			
Other? _____			
III. PROFESSIONAL NON-STATE/NON-POLITICAL		**IV. PRIVATE NON-STATE/NON-POLITICAL**	
Category	*check here to include*	*Category*	*check here to include*
Journalists		Private citizens (entertainment celebrities, retired states-people, bloggers, professors, etc.)	
Civic educators		Other? _____	
Civic activists			
Community leaders			
Civil society organization leaders			
Leaders or members of gender-based violence assistance groups			
Other? _____			

Specify the Model for Selecting Potential Targets

Identify the target numbers of individuals that would ideally be monitored. Seek a gender balance where possible. In some cases, gender balance may not be possible and results will need to be adjusted proportionally afterwards (for example, if there are very few women cabinet members).

I. POLITICAL			
Category	*Political Affiliation*	*Sex*	*Target number to include*
Candidates from both opposition and incumbent parties	Ruling party	Women	#
		Men	#
	Opposition party 1	Women	#
		Men	#
	Opposition party 2	Women	#
		Men	#
Elected officials from both opposition and incumbent parties	Ruling party	Women	#
		Men	#
	Opposition party 1	Women	#
		Men	#
	Opposition party 2	Women	#
		Men	#
Political aspirants (i.e., seeking nomination)	Ruling party	Women	#
		Men	#
	Opposition party 1	Women	#
		Men	#
	Opposition party 2	Women	#
		Men	#
Political staffers	Ruling party	Women	#
		Men	#
	Opposition party 1	Women	#
		Men	#
	Opposition party 2	Women	#
		Men	#
Party members and supporters, youth/women's wing leaders	Ruling party	Women	#
		Men	#
	Opposition party 1	Women	#
		Men	#
	Opposition party 2	Women	#
		Men	#
Political appointees in government (cabinet ministers, etc.)		Women	#
		Men	#
Other? _____		Women	#
		Men	#

II. INSTITUTIONAL		
Category	*Sex*	*Target number to include*
EMB Permanent Staff	Women	#
	Men	#
Police and Security Forces	Women	#
	Men	#
State Administrators/Civil Servants	Women	#
	Men	#
Poll Workers	Women	#
	Men	#
Other? _____	Women	#
	Men	#

III. PROFESSIONAL NON-STATE/NON-POLITICAL		
Category	*Sex*	*Target number to include*
Journalists	Women	#
	Men	#
Civic educators/Activists/CSO Leaders	Women	#
	Men	#
GBV Service Providers/Activists	Women	#
	Men	#
Community or Religious Leaders	Women	#
	Men	#
Other? _____	Women	#
	Men	#

IV. PRIVATE NON-STATE/NON-POLITICAL		
Category	*Sex*	*Target number to include*
Private Citizens (celebrities, retired states-people, bloggers, professors, etc.)	Women	#
	Men	#

Securely record subjects' key information

In this step, record the information for the individuals identified in the previous step.

I. POLITICAL												
Category	*Full name*	*Nickname*	*Social Media Handle*	*Sex*	*Political Affiliation*	*Marital Status*	*Constituency or Region*	*Religious minority?*	*Ethnic minority?*	*Racial minority?*	*Disabled person?*	*LGBTQ?*
Candidates from both opposition and incumbent parties												
Elected officials from both opposition and incumbent parties												
Political aspirants (i.e., seeking nomination)												

Political staffers
Party members and supporters, youth/women's wing leaders
Political appointees in government (cabinet ministers, etc.)
Other? _____ ____

II. INSTITUTIONAL										
Category	*Full name*	*Nickname*	*Social Media Handle*	*Sex*	*Marital Status*	*Religious minority?*	*Ethnic minority?*	*Racial minority?*	*Disabled person?*	*LGBTQ?*
EMB Permanent Staff										
Police and Security Forces										
State administrators/Civil Servants										
Poll workers										
Other? _____										

II. INSTITUTIONAL										
Category	*Full name*	*Nickname*	*Social Media Handle*	*Sex*	*Marital Status*	*Religious minority?*	*Ethnic minority?*	*Racial minority?*	*Disabled person?*	*LGBTQ?*
Journalists										
Civic educators/Activists/CSO Leaders										
GBV Service Providers/Activists										
Community or Religious Leaders										
Other? _____										

IV. PRIVATE NON-STATE/NON-POLITICAL										
Category	*Full name*	*Nickname*	*Social Media Handle*	*Sex*	*Marital Status*	*Religious minority?*	*Ethnic minority?*	*Racial minority?*	*Disabled person?*	*LGBTQ?*

Private citizens (entertainment celebrities, retired states-people, bloggers, professors, etc.)

2. Define the Lexicon for Monitoring

Define a general lexicon or words and phrases according to the three general categories. Test this in the software to eliminate words or phrases that generate "noise" or to identify other words or phrases to add into the lexicon.

Lexicon	**Bodily Harm**	**Sexual**	**Socio-Psychological**
Words	Murder Kill Die Beat Bury Destroy Hit Punch Slap Attack Perish Off	Rape Fuck Screw Ravish Violate Bang Pound Screw Molest Spoil	Idiot Moron Stupid Bimbo Asshole Fool Cunt Bitch Fag Dick Tits Whore
Phrases	Knock off Six feet under Grassy knoll	Sexually assault Do her Do him	Crazy Woman Mad man Shut up and make me a sandwich

Review the tested lexicon to try to identify subcategories

Within the general typologies of VAWE-Online (socio-psychological, sexual, physical etc.) usually there will be reoccurring themes in the social media posts that can be further sorted into specific subcategories of violence, which helps in capturing a more detailed level of analysis.

Example of Subcategories from the VAWE-Online Ukraine Assessment

Typology	Subcategories
Socio-psychological	• Xenophobia • Antisemitism • Bad reputation, biography • Lack of intelligence • Sarcasm, ridicule • Discreditation of family • Lack of competence
Sexual	• Prostitution • Sexual harassment • Sexual blackmail • Homosexuality
Physical	• Direct harm to principal target • Implied harm to principal target • Direct harm to a proxy

	• Implied harm to a proxy
Economic	• Accusations of corruption • Threats of denial of resources, control, confiscation

3. Define which Social Media to Monitor

Which public social media and online fora will be monitored? *These can be monitored via data mining and sentiment analysis.*	-Facebook *(public pages)* -Instagram *(public pages)* -Medium -VK *(public pages)*	-Reddit -YouTube -Baidu Tieba -TikTok/Douyin	-Tumblr -Pinterest -Sina Weibo -Qzone
Which private online spaces might also be relevant? *These can be analyzed through follow-up interviews and focus groups.*	-WhatsApp -WeChat -QQ -Viber -Snapchat -Facebook *(Messenger and private pages)*	-Telegram -Line -Instagram *(direct messages and private pages)* -Odnoklassniki *(private pages)*	-Tik Tok/Douyin *(private pages)* -VK *(direct messages and private pages)*

Annex 4: Database examples

Individual Database	Hillary Clinton	Beyoncé	Michael Jordan	Mahatma Gandhi	Candidate 1	Candidate 2
Full Name*	Hillary Rodham Clinton	Beyoncé-Knowles Carter	Michael Jeffrey Jordan	Mohandas Karamchand Gandhi	Candidate 1	Candidate 2
Nicknames	Hillary HRC	Beyoncé	Michael Jordan MJ	Bapu Gandhi Mahatma Gandi	Candy, Candidato	Candi, Candidata
Age	71	37	55	149	47	56
Sex*	Female	Female	Male	Male	Male	Female
Ethnic Identity	White	Black/African American	Black/African American	Indian	Indigenous- Maya	Arab
Religious Identity	Christian	Christian	Christian	Hindu	Catholic	Muslim
Disability Status	No	No	No	No	Yes	No
Location	National-level	N/A	N/A	National-level	District 17, rural, local-level	National-level; urban; reserved seat
Political Affiliation*	Democratic Party	Democratic Party	N/A	Indian National Congress	Leftist party	Secular, centrist party
Social Media Information	Facebook Instagram: hillaryclinton Twitter: @HillaryClinton	Twitter: @Beyonce Instagram: @Beyonce	N/A	N/A	Twitter: -- Facebook: -	Twitter: -- Facebook:-
Category*						
Subcategory						

Organization Database	**World Wildlife Fund**	**Organization 1**	**Organization 2**
Full Name*	World Wildlife Fund (WWF)	Organization 1	Organization 2
Legal Address	1250 24th Street N.W. Washington, DC 20037 USA	1 Gender Equality Street, Addis Ababa, Ethiopia	2 Social Justice Way, Port-Au-Prince, Haiti
Geographic Presence	International	National	Ouest department
Political Affiliation	N/A	EPRDF	N/A
Social Media Information	Facebook Instagram: @wwf Twitter: @WWF	Facebook Instagram Twitter	Facebook Instagram Twitter
Category*			
Subcategory			

Annex 5: Self-Run Platforms

These represent some, not all, options for self-run platforms. This information was accurate according to publicly available content as of July 31, 2019.

Self-run platforms	#TAGS	Discover-Text	NVivo	NVivo Capture	R and Python	Rapid Miner*	WEKA
Access:	Website Through Google Docs	Website	Website	Website Used with Google Chrome and Internet Explorer	R: website Python: website	Website	Website
Help Resources	Support forum	Assistance webpage	Webpage as well as a technical team; trainings	Webpage as well as NViVo support	R: site and online forums Python: site and online forums	Webpage: training and support forums	Various online forums
Social media Platforms	Twitter	N/A: used to mine and analyze existing data	N/A: used to mine and categorize existing data	Facebook; Blogs; YouTube; Webpages; Twitter posts	Can be used for Twitter collection; also used to mine and analyze	Twitter; used to mine and categorize existing data	N/A: used to mine and categorize existing data
English-language focused	No	No, but not tested for all languages	No	No	No	No, but not all languages have full functionality	No
Cost	Free	$24-$2,000 USD per month	$1,119 USD for NGOs (one-time fee)	Free, with NVivo	Free	$2,500-$10,000 per year	Free

Technical Expertise Required	Low-Moderate	Moderate	Moderate	Low	High	Low-Moderate	Very High
Sub-programs			NVivo NCapture		VADER, tidytext, and others	*RapidMiner is an "established program"	

International Foundation
for Electoral Systems

IFES | 2011 Crystal Drive | 10th Floor | Arlington, VA 22202 | www.IFES.org