

---

# Alliance for the Protection of Children Project

## Broadcast and Social Media Narratives on Violence Against Children in Haiti



**USAID**  
FROM THE AMERICAN PEOPLE



JULY 2017



*This publication was produced by American Institutes for Research (AIR) for review by the United States Agency for International Development. It was prepared by an AIR Researcher Team. The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States government.*

# Alliance for the Protection of Children Project

## Broadcast and Social Media Narratives on Violence Against Children in Haiti

July 2017



# Contents

	<b>Page</b>
Introduction.....	1
Project Overview .....	1
Report Summary .....	1
Broadcast Media .....	2
Data Collection Method.....	2
Findings—What Are People Talking About on Radio Broadcasts?.....	5
Findings—Who Is Influencing the Discourse on Child Violence? .....	6
Findings—Who Are These Programs Talking About?.....	8
Social Media .....	8
Data Collection Methods .....	8
Findings: How Are People Talking on Twitter?.....	11
Findings: What Are People Talking About on Twitter?.....	12
Findings: Who Is Talking? .....	14
Findings: What Are These People and Organizations Saying? .....	15
Findings: If Not Child Violence and Protection, Then What Are Organizations Saying on Facebook? .....	16
Appendix A. List of Radio Stations.....	A-1
Appendix B. Social Media Data Collection.....	B-1
Appendix C. Twitter Keyword List and Resulting Data.....	C-1
Appendix D: Twitter Relevance Model.....	D-4

## Tables

Table 1. Social Media Data Sources and Accessibility for Data Collection .....	9
--	---

## Figures

Figure 1. Total Radio Stations With Child Violence and Protection Programming.....	3
Figure 2. Role of the Journalist by Gender in Radio Broadcasts .....	7
Figure 4. Frequency of Top 10 Keywords by Number of Tweets .....	10
Figure 5. Word Cloud Generated From All Relevant Tweets .....	12
Figure 6. Distribution of Followers, Friends, and Retweets by Sex .....	14
Figure 7. Percentage of Tweets by Topic for Subgroups of Twitter Users .....	15
Figure 8. Word Cloud of Words Used in NGO's Facebook Posts .....	17
Figure 9. Process of Tweet Evaluation .....	A-4

# Introduction

## Project Overview

The Alliance for the Protection of Children (APC) is a 3-year U.S. Agency for International Development (USAID)-supported project implemented by American Institutes for Research (AIR) with a goal to leverage local assets and strengthen the protection of children exposed to all forms of abuse, neglect, exploitation, and violence. Lumos Foundation is a resource partner to the project leading efforts that mitigate abuse, neglect, exploitation, and violence in residential care institutions. The APC will support strategic objectives on the part of USAID and the Government of Haiti to strengthen the protection of vulnerable children. The APC will ensure that its work aligns with the National Child Protection Strategy and contributes to Objectives 3, 4, and 5, as well as existing laws governing the protection of children in Haiti. We will work in partnership with the Institut du Bien-Être Social et de Recherches, Université d'Etat d'Haiti (UEH), Zanmi Lasante, Combite pour la paix et le développement, and Restavek Freedom/OPREH in Year 1 to establish a rigorous evidence base on which to build Year 2 and 3 pilot interventions.

The APC will collaborate with other government agencies, ministries, strategic United Nations agencies such as UNICEF and UNHCR, and community-based and local nongovernmental organizations (NGOs) in Haiti. We will prioritize child protection through this program; specifically, we will pilot evidence-based interventions that reduce violence against children; mitigate the trafficking and forced labor of children; protect unaccompanied and separated migrant, stateless, and displaced children; integrate street children into safer learning and care spaces; prevent the separation of children from their families; and explore alternative care and protection services. The APC has four distinct phases: (1) research; (2) design of pilot interventions; (3) implementation of pilot interventions; and (4) evaluation and learning for scale-up. The Year 1 project plan covers the first and second phases.

## Report Summary

This report summarizes the collection and analysis of social and broadcast media data in Haiti. This project is part of a larger research project, the APC, designed to collect evidence on individual, family, community and system level attitudes, behaviors and responses to violence against children and identify those factors that mitigate or perpetuate cycles of violence. The scope of the overall project is focused on the design of evidence-based pilot interventions that will leverage local assets and strengthen child protection efforts in Haiti. To do this, the research team is committed to look at the problem of abuse, neglect, exploitation, and violence against children using several different research methods. This report's focus on communications is just one of several unique lenses for data collection explored as part of the overall project. To the extent that future interventions focus on affecting content of radio broadcasts, Twitter posts, and Facebook posts, the results presented in this report can be thought of as "baseline" data for monitoring changes in what people are communicating about and patterns of influence.

Specifically, mapping influencers and followers of social and broadcast media is a significant component of community-based research and gives evidence to support that the Haitian people

and organizations are engaging in discourse around the problem of child violence. This evidence can feed interventions on how and with whom to communicate as well as which people and organizations are particularly influential in their communications. These influential users of social and broadcast media are targets for potential communication strategies through their continued discourse.

This report describes the results from two prongs of research: one conducted using sophisticated data science techniques and approaches based on collected social media data and one conducted using a qualitative review of collected data from radio broadcasting. The goal of this research was to better understand the communication of the problem (what), who is communicating (who), and what venue of communication is used (how) to determine whether we can identify which communications are the most influential and effective.

After briefly discussing the methods used to collect data, this report presents the findings related to (a) which topics within the child violence and protection theme are talked about by whom, (b) which speakers have the most influence, and (c) which topics are being talked about by the most influential speakers.

## Broadcast Media

### Data Collection Method

To examine and analyze broadcast media to understand how the Haitians are discussing violence against children and to pinpoint how they are influencing the public, the AIR and UEH research teams designed the following data collection process after assessing the feasibility of collecting data from radio station programming and from walkie-talkie broadcasts. The team limited data collection exclusively to radio station programming because radio is the “most popular source of news and information in Haiti,”<sup>1</sup> and unlike walk-talkie broadcasts, radio broadcasts are readily available on the Internet or via smartphone applications.

A media consumption survey taken on behalf of USAID in 2010 showed 83% of Haitians across all departments preferred radio programming to find locally relevant information. According to Internews, which conducted the survey, 375 radio stations aired across Haiti in late 2012. Most radio stations are small and privately owned with limited funding broadcasting on local issues of immediate relevance to their target audiences.<sup>2</sup>

The research team began data collection by obtaining a list of all radio stations in the four target departments (l’Artibonite, the South East, the West, and the Southeast) from the State Office of

---

<sup>1</sup> Internews. (2012). *Haiti media and telecoms landscape guide*. Retrieved from [https://www.internews.org/sites/default/files/resources/haiti\\_media\\_guide\\_final\\_211012\\_with\\_retyped\\_index\\_19.12.12.pdf](https://www.internews.org/sites/default/files/resources/haiti_media_guide_final_211012_with_retyped_index_19.12.12.pdf)

<sup>2</sup> See footnote 1.

Communications (CONATEL). This included 99<sup>3</sup> legally registered stations. Most of these stations are located in Port-au-Prince or the West department (53 stations), 35 stations are in Artibonite, and two stations apiece in the Northeast and the Southeast, respectively. Through the course of data collection, an additional seven radio stations were found and added to the original list, for a total of 106 radio stations with 57 stations in the West, 37 in Artibonite, 10 in the North East, and 2 stations in the South East.

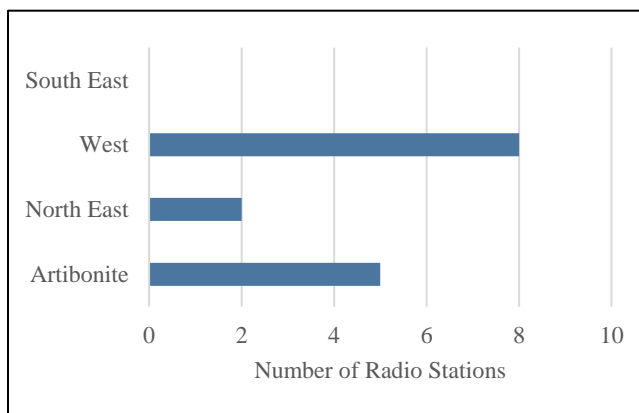
To identify programming related to child violence and protection, the UEH research team contacted and interviewed radio station program managers and other station representatives to identify aired programs that focused on issues on child protection and violence against children. The UEH team asked about station programming with themes on child health, education, schooling, discipline, and homelessness. They also asked station representatives whether they had broadcasted programming specifically about child abuse. Included in the questionnaire was information about what time or day these programs aired and whether the public was allowed to participate.

Collecting these data from each radio station was challenging. Radio stations did not often have working telephones and rural stations were difficult to find.

*[...] gestured to a dirt trail—banana leaves curled over it to form a tunnel. A few minutes later, he reached a concrete hut, from which an antenna stretched up to the cloud. “C’est la Radio Brillante!” he declared.<sup>4</sup>*

Once the UEH researchers located the radio station, the UEH researchers had to return to the radio station after making an appointment. To obtain information about radio station programming, it took the team more than four contacts (via telephone or in person visits) on average. The UEH research team used local knowledge of the environment to focus on those radio stations that were more likely to have relevant programming. Of the 106 radio stations, the UEH researchers collected information only from 30 stations, and of those, 16 stations had relevant programming.

**Figure 1. Total Radio Stations With Child Violence and Protection Programming**



<sup>3</sup> The research team assumes that this number is significantly lower than the 375 radio stations reported in 2012 for a few reasons. First, small, rural radio stations are likely not to have survived Hurricane Matthew. Second, radio broadcast infrastructure was planned to switch to a digital format, and it is likely that smaller stations could not afford the cost. Thirdly, this number includes radio stations that are not legally registered with CONATEL. Last, the social, political and financial environment for journalists is unstable, which could make it difficult for stations to have significant longevity.

<sup>4</sup> Cuesta Roca, C. (2016, April 14). Radio in Haiti is heartbeat of the community. *Miami Herald*. Retrieved from <http://www.miamiherald.com/news/nation-world/world/americas/haiti/article71931957.html>



In addition to soliciting programming data from radio stations directly, the research team asked similar questions of partner NGOs as a means to independently verify and supplement information collected from radio stations. Although the research team requested this information from three project partner NGOs, only the NGO in the West Department had enough knowledge of radio programming to provide the team with any information. The data provided were used to verify the data for one radio station and served as a supplemental, or proxy, data source for an additional **two** radio stations from the West Department outside of Port-au-Prince. Both additional stations had relevant programming.

Of the 106 radio stations, the UEH and AIR teams collected data from 32 radio stations (either directly from the station or via partner NGOs); of those, 18 stations had relevant programming.

The 18 radio stations with relevant programming had 26 programs that dealt with topics in child protection and violence. These programs aired throughout the week; most (18 of 26 programs) aired during the weekend. The programs tended to be broadcasted in the morning (11 of 26) or in the early afternoon (12 of 26) and were unlikely to air in the evening (four of 26 programs).

Given the time and resource constraints, the team recorded and listened to 16 relevant programs. They listened specifically for the following topics in order to complete a broadcast media instrument:

- **Child well-being:** What do children need to develop well? Who is responsible for providing those things? What kinds of things are contributing to child well-being? What kinds of things are barriers to well-being? Why?
- **Opportunities for children:** What opportunities do people see for children in Haiti? Outside Haiti? How do adults help children with these opportunities? What kinds of difficult choices do adults have to make in order to provide opportunities for children?
- **Children living outside of families:** Why do children end up living outside of their family? Where do they go? What is life like for them when they leave? What services are there for them when they go to a new place? How are they protected in those situations? How are they vulnerable in those situations? Who monitors them in those situations and how? Who refers them to services they might need?
- **Violence against children:** What counts as violence (or neglect or abandonment or abuse or exploitation)? Who perpetuates the violence? Why? Who are the victims? Where is it mainly happening?
- **Child protection:** Who protects children? What programs or organizations protect children? What works? What does not work? Why?
- Data collection was designed to target the most important themes and to help the team understand the content of each program. The team specifically looked for topics in youth or children's health, school education, punishment, corporal punishment, development, and discipline, as well as sex education, as well as sexual, physical, and emotional abuse. Topics related to homeless youth, trafficking and services for children were also included.

The following findings are representative of the data collected but cannot be generalized to all radio stations and programming in Haiti.

Appendix A gives a complete list of all radio stations with notations to which radio stations had station and programming data collected.

## Findings—What Are People Talking About on Radio Broadcasts?

Radio is the most popular source of news and information in Haiti.<sup>5</sup> In the aftermath of the earthquake in 2010, radio programming emerged as a source of information and to share news.<sup>67</sup>

*“The radio helps me know how to behave in society,” he said, his flip phone open on top of his snack cart, playing the weekly science show. “I don’t go to school but Brillante helps me learn.”*<sup>8</sup>

Radio stations were used following Hurricane Matthew (Nov 2016) to communicate information about World Food Programme (WFP) assistance and allowed people to call in to ask questions or to provide feedback.

*“Though each disaster is unique, we see that access to vital information in emergencies is an ever-present need and really does mean survival.”*  
*Enrica Porcari, Chair of the ETC and WFP Chief Information Officer and Director of IT*<sup>9</sup>

The AIR and UEH researchers found that in addition to communicating to the public information about natural disaster assistance, the Haitian people used radio broadcasts to engage in public conversation about child violence and protection issues, to inform people in how to help those in need avoid violence, and even to incite violence to gain media attention toward this issue. In no particular order, the following types of discourse highlighted with representative samples were found upon review of the collected radio programming.

### **Using Radio to Engage in Public Conversation About the Issue**

Multiple (?) broadcasts sought to broaden public awareness about the problem of violence against children. Some used broadcasts to draw attention to the issue, while others also promoted strategies to increase public discourse. For example :

*Program name and radio station: Protégeons les enfants, Radio nationale*

*Summary:* This program, focused on child protection against sexual abuse, aired solely to raise public awareness of the problem of violence against children. It talked about educating journalists on how to deal with stories related to child abuse and aggression.

---

<sup>5</sup> See footnote 1.

<sup>6</sup> Maloney, J. (2010, September 2). In post-quake Haiti, radio evolves into powerful mouthpiece. *Public Radio International*. Retrieved from <https://www.pri.org/stories/2010-09-02/post-quake-haiti-radio-evolves-powerful-mouthpiece>

<sup>7</sup> World Food Programme. (2017, March 16). *Haiti: Community radio stations back on the airwaves in hurricane-affected areas*. Retrieved from <https://www.wfp.org/news/news-release/haiti-community-radio-stations-back-airwaves-hurricane-affected-areas>

<sup>8</sup> See footnote 4.

<sup>9</sup> See footnote 7.

*Eviter dans une maison ou vivent de multiples gens, de faire l'amour, en présence des enfants.*

*Eviter la stigmatisation.*

*Eduquer les journalistes sur la façon dont ils doivent traiter les problèmes d'agressions que viennent de subir les enfants.*

### **Using Radio to Discuss Measures of Avoiding Violence**

A second theme identified was a focus on general actions that might be taken to avoid violence against children. While not describing specific programmatic or community interventions, these programs promoted awareness and avoidance at the personal level. For example :

*Program name and radio station: Tou piti on Radio Pacific*

*Summary:* This program spoke about the functioning of the children in daily life. It talked about the impact of domestic violence on children's lives and gave advice on how to avoid violence against girls and women. It also discussed music that incites violence against young people.

*Impact de la violence familiale dans la vie des enfants.*

*La violence n'est pas un bon exemple pour l'enfant qui grandit.*

*Eviter la violence dans la famille n'est pas une bonne image pour l'avenir des enfants.*

### **Using Radio as a Call to Action**

The final group of programs identified the issue of violence against children, and promoted calls to action. These programs differed from those promoting awareness (the first theme) and personal avoidance (the second theme) in that they focused on collective action towards combating the program, including among youth themselves, as this example illustrates :

*Program name and radio station: Pale poun jije, Radio Timoun*

*Summary:* This broadcast, designed and followed by youth, uses violent messaging to draw the listener's attention to the possibility of violations occurring to the children who listen to it. On the day of this airing, the messages were directed at the Haitian bourgeoisie who are allegedly mistreating the most vulnerable categories of children.

*Il faut que la population se mobilise pour qu'on ne continue pas a le maltraiter.*

*Soulevons-nous afin de nous débarrasser de ces malfaiteurs.*

## **Findings—Who Is Influencing the Discourse on Child Violence?**

Understanding the different types of media actors who are active in efforts against child violence is particularly important given the substantial role-based differences in approaches to information dissemination. Although the shift toward online, interactive content can be seen as

eroding<sup>10</sup> the traditional role of the journalist as the gatekeeper of information,<sup>11</sup> radio continues to play a dominant role in Haitian media. Journalists and radio broadcasting continue to have an important role as information-curators. Given this particular role—and the cross-national journalistic norms that accompany it<sup>12</sup>—we expect journalist-moderators to differ from nonjournalist media personality moderators in their approaches. As such, it is important to know the role of the moderator as well as their political affiliation (which has a more obvious direct impact on the way issues are framed<sup>13</sup>). In the media data collected, 80% of the programs had a male moderator. In more than 40% of the cases, the moderator was male and a journalist.

In terms of political affiliation, most (~70%) moderators had no reported affiliation. Three moderators were open about their political affiliation. Two moderators were supportive of the right and one moderator was supportive of the party of the former president Jean Bertrand Aristide.

The behavior patterns of the moderator signified that most programs are being used to incite some type of call for action from their listening population in addition to using radio media to inform the public.

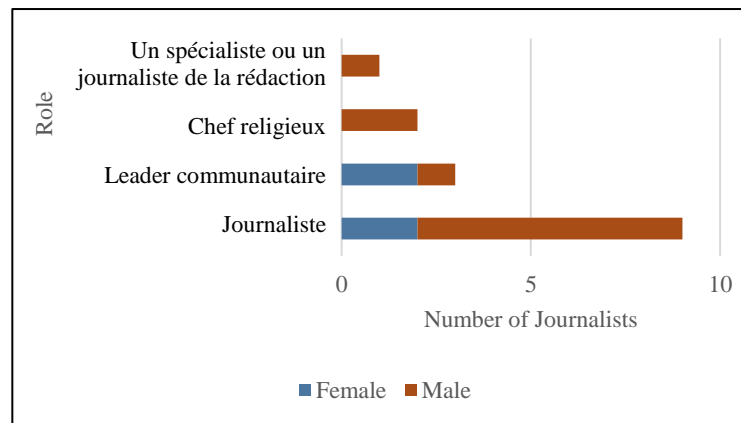
The *Woule, woule rakontem* program airing on Radio Espace FM encouraged the public to take action by helping children go to school.

*Aider les enfants à aller à l'école, parce que l'éducation c'est de la richesse.*

The *Vendredi Famille* program airing on Radio Lumière called for protection of specific groups by asking young men to protect their relationships from domestic violence.

*Jeune Homme, vous devez protéger cette jeune fille qui est votre compagne.  
Ne la frappez pas en cas de jalousie. Protéger votre relation de couple.*

**Figure 2. Role of the Journalist by Gender in Radio Broadcasts**



Note. N = 16 programs.

<sup>10</sup> Weaver, D. H., & Wilhoit, G. C. (1996). *The American journalist in the 1990s: U.S. news people at the end of an era*. Mahwah, NJ: Lawrence Erlbaum.

<sup>11</sup> White, D. M. (1950). The “gate keeper”: A case study in the selection of news. *Journalism Quarterly*, 27(3), 383–390.

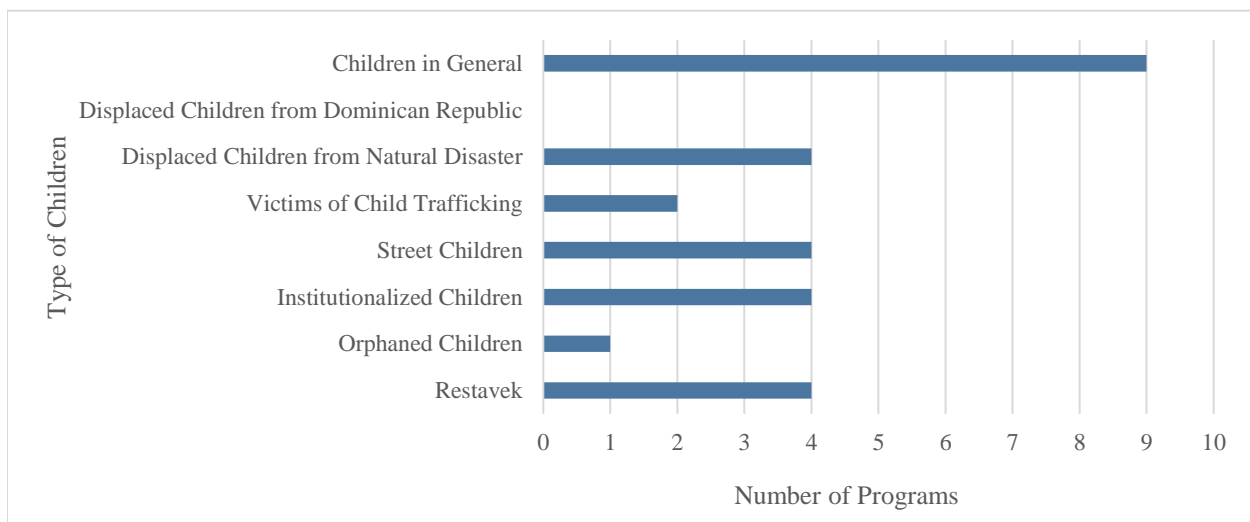
<sup>12</sup> Hanitzsch, T., Hanusch, F., Mellado, C., Anikina, M., Berganza, R., Cangoz, I., ... Yuen, E. K. W. (2011). Mapping journalism cultures across nations: A comparative study of 18 countries. *Journalism Studies*, 12(3), 273–293.

<sup>13</sup> Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of Communication*, 49(1), 103–122.

## Findings—Who Are These Programs Talking About?

During most programs (60%), the moderator and call-in participants (if there are any) spoke about children in general. These programs did not target child violence or protection specifically as the main theme for the program, but the topic was encountered in conjunction with the main themes. For programs that centered specifically on the topic of violence, protection, and children, the moderator and participants talked generally about restaveks, orphaned children, street children, and children displaced from natural disasters. It is interesting to note that although likely to be a significant problem affecting many children, no program mentioned specifically the subgroup of those children displaced from the Dominican Republic.

**Figure 3. Number of Programs that Spoke About Specific Types of Children**



## Social Media

### Data Collection Methods

To examine and analyze social media to understand how Haitian people discuss violence against children and to pinpoint who and how they are influencing the public discourse, the AIR research team began data collection by assessing which social media sources are available to the Haitian people, which sources are being used, and which sources allow for automated access to the data. Specifically, the team assessed the following data sources: Facebook, Instagram, Twitter, Wordpress, Blogger, Disqus, and WhatsApp. Table 1 describes these data sources. More detailed information is available in Appendix B.

**Table 1. Social Media Data Sources and Accessibility for Data Collection**

Source	Type of Media	Do People/Organizations in Haiti Use the Application?	Are the Data Accessible?
Facebook	Used for communicating with closed networks	Yes—Organizations and individuals use Facebook.	Yes and no—Per Facebook’s data use agreement, only public organization page data can be extracted.
Instagram	Used for posting videos and pictures	Yes—Organizations and individuals use Instagram.	Yes and no—Per Instagram’s data use agreement, only public organization page data can be extracted.
Twitter	Used for communication to the wider public	Yes—Organizations and individuals use Twitter.	Yes—Twitter allows for automated extraction of organization and individual data.
Wordpress	Used to create and post online journals	Unclear—Without collecting data, it is difficult to assess the usage of blogger by local Haitian people.	Yes—Wordpress allows for automated extraction of online blog and website data.
Blogger	Used to create and post online journals	Unclear—Without collecting data, it is difficult to assess the usage of blogger by local Haitian people.	Yes—Blogger allows for automated extraction of online blog data.
Disqus	Used on websites for public comments	Yes—Haitian websites and bloggers use Disqus for comment applications.	Yes—Disqus allows for automated extraction of website comments.
WhatsApp	Used for SMS messaging between two or more people	Yes—WhatsApp is a very prolific instant messaging application used in Haiti.	No—WhatsApp is a closed IM.

The research team focused on data collection from Twitter and Facebook. Twitter is particularly suited to this type of data collection and analysis because the Twitter platform is set up to share content publicly and allows users to spread the information wide rather than within dense friendship network like Facebook and Instagram. Twitter allows for public access to their content and user information and supports this access through application programming interfaces<sup>14</sup> (APIs), a web-based system created to filter and display data.

### **Twitter**

The team began data collection from Twitter by developing a comprehensive list of keywords to search for, collecting tweets containing those keywords, and then filtering the resulting data to find those that were most relevant. This keyword list was developed using the final keyword list from the program’s systematic review<sup>15</sup> and a brainstorming session with UEH researchers. All

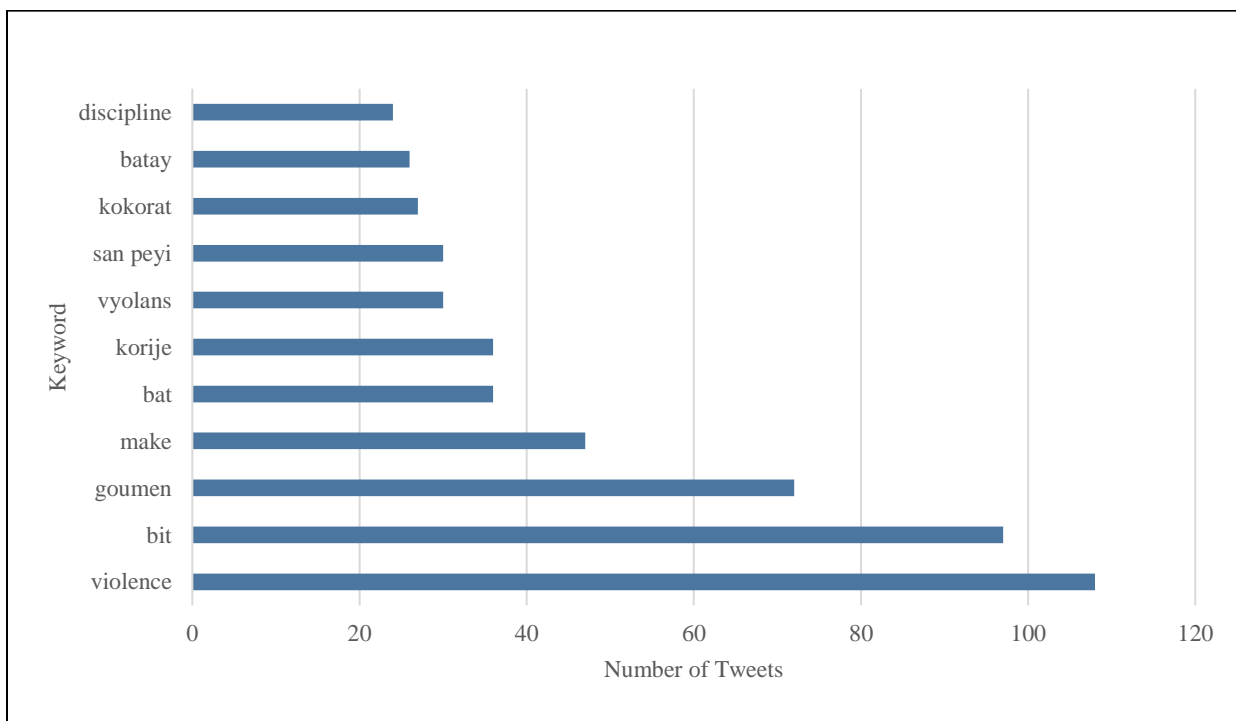
<sup>14</sup> An API is a set of subroutine definitions, protocols and tools for building application software. [https://en.wikipedia.org/wiki/Application\\_programming\\_interface](https://en.wikipedia.org/wiki/Application_programming_interface)

<sup>15</sup> Alliance for the Protection of Children Project: A Systematic Research Review and Meta-Analysis, American Institutes for Research, March 2017

keywords were translated to English, French, and Creole. For a complete list of keywords, see Appendix C.

The research team then used the Twitter search and streaming APIs to collect tweets based on the developed keyword list, a geographic filter for tweets from users who set their location as “Haiti” as well as a targeted search of users from within a 10 mile radius of Port-au-Prince, Ferrier, Ouanaminthe, Anse a Pitre, Marchand-Dessalines, Gonaives, and Saint Marc. Most data came from the Port-au-Prince area. Data were collected at intervals beginning January 30, 2017, through June 9, 2017, which yielded a total of 2,557 tweets. Figure 4 displays the top 10 keywords that generated the most tweets. Appendix C gives the total number of collected tweets by keyword.

**Figure 4. Frequency of Top 10 Keywords by Number of Tweets**



## **Facebook**

In addition to collected data from Twitter, the research team collected data from Facebook. Unlike Twitter, Facebook only allows the collection of public organization data. Data created on individual pages, regardless of their public/private status, are not allowed under the current data use regulations. Data collection from Facebook began by collecting the Facebook web links and associated page identifiers (static numeric identifier used to locate each organization page) for a directory of NGOs obtained from the Ministry of Planification originally provided for the project’s Rapid Qualitative Assessment.<sup>16</sup> For example, this list included Haitian nonprofits, such

<sup>16</sup> Alliance for the Protection of Children Project: Findings from a Rapid Qualitative Assessment on Child Protection in Haiti, American Institutes for Research, June 2017

as Action Concertée pour le Développement and Arbo de Vivo Haïti, as well as international NGOs, such as Global Health Action and Christian Aid. This list represented all potentially active NGOs with offices located in the departments of interest in Haiti. The Facebook links and page identifiers were collected manually for all 605 organizations. Of those 605 organizations,<sup>17</sup> the team identified 97 local Haitian organizations and 130 international organizations with a Facebook page. Pages were tagged as “local” if the organization was located inside Haiti.

The AIR team then used the list of Facebook page identifiers to extract all the associated post and response data associated with the identifier collecting the most recent 100 posts (if there were that many) from each organization. This analysis resulted in 1,482 posts from Haitian organizations and 2,704 posts from international organizations operating in Haiti (of which 265 mentioned Haiti specifically).

Unfortunately, upon review of the Facebook data, the AIR team did not find enough relevant Facebook posts to continue. The team manually read through all 2,704 posts, and fewer than 20 of those posts contained text related to the research topic. The Findings section for Facebook data overviews the basic themes that organizations are discussing, but the important finding from this endeavor is that this topic is not discussed on Facebook.

## Findings: How Are People Talking on Twitter?

The AIR team used a relevance model (Appendix D) to pinpoint those tweets that were relevant to the project topic. The resulting data set of relevant tweets included 45 tweets in English, 62 tweets in French, and 8 tweets in Creole, for a total of 221 tweets and 115 Twitter users. Appendix C provides samples of relevant tweets.

People and organizations on Twitter who discussed child violence mostly discussed the topic in general terms rather than mentioning specific events witnessed personally, or repeated stories from their local communities. Most tweets related to child violence can be categorized into four broad themes that center around raising awareness of general issues surrounding child violence and protection:

*32000+ kids in #Haiti live in orphanages 80% aren't orphans  
They are used to profit from well-intentioned donors.*

1. Calling attention to address the problem publicly.

*Au forum #ECWECC nouvelle structure des pouses des Chefs dEtats de la CARICOM  
pr aider les femmes les enfants*

2. Highlighting efforts being undertaken:

*@AntonioGuterres calls on Security Council members to come together to work to end  
human trafficking*

3. Encouraging others to take action.

---

<sup>17</sup> List of NGO organization offices from Haiti’s Ministry of Planification with Facebook links  
[https://docs.google.com/spreadsheets/d/127ee8W97tNim\\_VR0RZlYBbVlSh1XH2r4INOSfOkgK4/edit?usp=sharing](https://docs.google.com/spreadsheets/d/127ee8W97tNim_VR0RZlYBbVlSh1XH2r4INOSfOkgK4/edit?usp=sharing)



session de sensibilisation sur la protection de l'enfance À l'école Méthodiste de Duplan

4. Mentioning specific instances of youth violence that had been in the news.

*Nos plus sincres condolances aux familles de ces jeunes gens assassins injustement et lchement Carrefourlls mrite*

No Tweets mentioned witnessing, encouraging, or remembering child violence directly, although one highly influential post (i.e., a post that was retweeted many times) did critique an advertising campaign for normalizing child labor.

*A lireQuand une publicit normalise la maltraitance des enfants et lesclavage moderne#Haiti*

## Findings: What Are People Talking About on Twitter?

The AIR research team identified four distinct themes in the data collected from Twitter. These themes were discovered by developing a topic model through Latent Dirichlet Analysis (LDA). LDA is a statistical model that identifies underlying topics in text based on the distribution of words within different documents, or in this case, tweets. The model assumes that the tweets have underlying word patterns that represent topics, and that the distribution of words in these tweets is driven by these underlying topics. For instance, we would expect tweets primarily about sexual violence to have different words than tweets focused on restavek.<sup>18</sup> Thus, the model divides the tweets into different topics based on the distribution of words within the tweets and labels these tweets based on the most common words in that topic.

Figure 5. Word Cloud Generated From All Relevant Tweets



<sup>18</sup> Blei, D. M., Ng, A. Y., Jordan, M. I., & Lafferty, J. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.

The following four topics emerged from the data. Please note that the labels are intended to characterize the topic and are not generated through automated processes.

1. **Restavek.** Tweets in the “restavek” theme centered on discourse specific to violence against children and the problem of restavek (which was referred to a “restavek” across all three languages). Many of these tweets mention restavek specifically, whereas others refer to “modern slavery” or “child slavery,” or focus on violence against children more broadly.

Distinguishing words include **restavek, child slavery, and violence.**

Example tweets:

*Experiences of Domestic Child Workers in Haiti aka Restavek | Ayiti Now Corp |*

*La violence contre les enfants est une grave violation des droits de la personne et un problme majeur ...*

*An n rasanble pou n konbat sistem restavek la*

2. **Child protection.** Tweets in the “child protection” theme focused on actions being taken to help end violence against children. The tweets either encouraged readers to help—by asking for donations, likes, or asking readers to spread anti-child abuse messages—or mentioned an effort being taken to end violence, often celebrating a successful event or campaign.

Distinguishing words include **child, act, help, children, violence, and young.**

Example tweets:

*10ans Rduction Violence Communautaire #Hati = 4 millions bnficiaires ds quartiers vulnrables 450projets 70millions*

*85 youth leaders from 3 hoods meet to share analysis and action plans on violence in their area.*

*How Ashton Kutcher is building tech to fight child sex traffickers*

3. **Violence against girls (women).** Tweets in the “violence against girls (women)” theme focused specifically on violence against women, violence on women and children and sexual violence. Unlike the other tweets focused on violence against children generally, most of these tweets mentioned sexual abuse specifically or referenced gender-based violence.

Distinguishing words include **sexual, violence, women, abuse, rape, violence, and youth.**

Example tweets:

*#Haiti Violence on women 127 cases recorded in the South*

*#Haiti Le Canada considre donner davantage de fond l'ONU pour les victimes d'abus sexuels*

4. **Child abuse.** Tweets in theme Child Abuse speak generally about violence against children as a problem in general. Tweets in this theme generally do not specify an event or to take a specific action (in contrast to the child protection theme) but rather describe the scope of the problem more broadly.

Distinguishing words include **child, children, youth, and violence.**

Example tweets:

*Parents first responsibility is to protect children against all forms of abuse, violence and #eksplwatasyon. #Eliminerlaviolence #timoun*

*Plus de 2 800 enfants dont plusieurs dizaines de fillettes vivent dans la rue selon l'Institut du Bien-Etre*

*Trump vows to fight 'epidemic' of human trafficking*

## Findings: Who Is Talking?

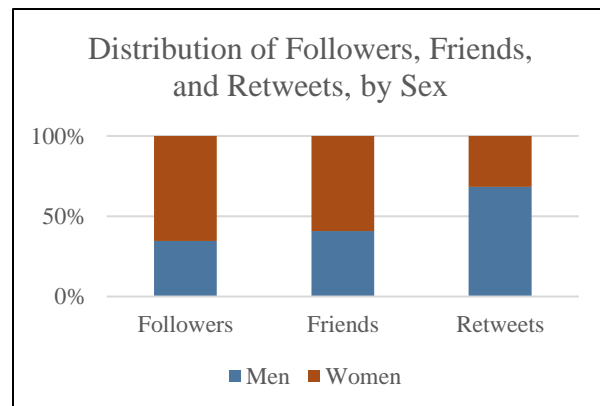
### **Organizations have more activity but individuals have the most influence**

Twitter users in Haiti are a mix of organizations (30.4%) and individuals (63.5%). Organizations tend to be more active (that is, tended to have more tweets about child violence)—organizations have an average of 2.3 relevant tweets per user with an average of 7,319 and 3,147 followers and friends respectively. This is almost 1 tweet per user greater than the average tweets per user for individuals (1.5) and more than double the number of followers and friends (3,561 and 1,403). However, even though organizations are more prolific on Twitter—they post more tweets, they have more followers and they have more friends –, individual users have significantly more influence. In simple terms, influence is based on the rate of re-tweets of a posts<sup>19</sup>. A user is said to have “high influence” if their posts tend to be re-tweeted by other users. The more posts with re-tweets, the more influence. Individuals have an average number of 7.0 re-tweets as compared to the organizations’ average of 1.6.

### **Females have more activity but males have more influence**

Most individual Twitter users are male. Of the 73 users in this sample who are individuals, 73% are male and 17% are female. This distribution is similar to—although more extreme than—the distribution of social media users in Haiti, where overall about 60% of social media users are men.<sup>20</sup> Women have a higher average number of followers and friends (5,817 followers and 1,827 friends for women vs. 3,067 followers and 1,255 friends for men); however, male users tend to have more influential posts. Male users have an average of 7.1 retweets, whereas females have a rate of 3.3.

**Figure 6. Distribution of Followers, Friends, and Retweets by Sex**



<sup>19</sup> Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. P. (2010). *Measuring user influence in Twitter: The million follower fallacy*. Paper presented at the Proceedings of the Fourth International Conference on Weblogs and Social Media, ICWSM 2010, Washington, D.C.

<sup>20</sup> Kemp, S. (2016, January 26). *Digital in 2016* [Slide show]. Retrieved from <https://www.slideshare.net/wearesocialsg/digital-in-2016>

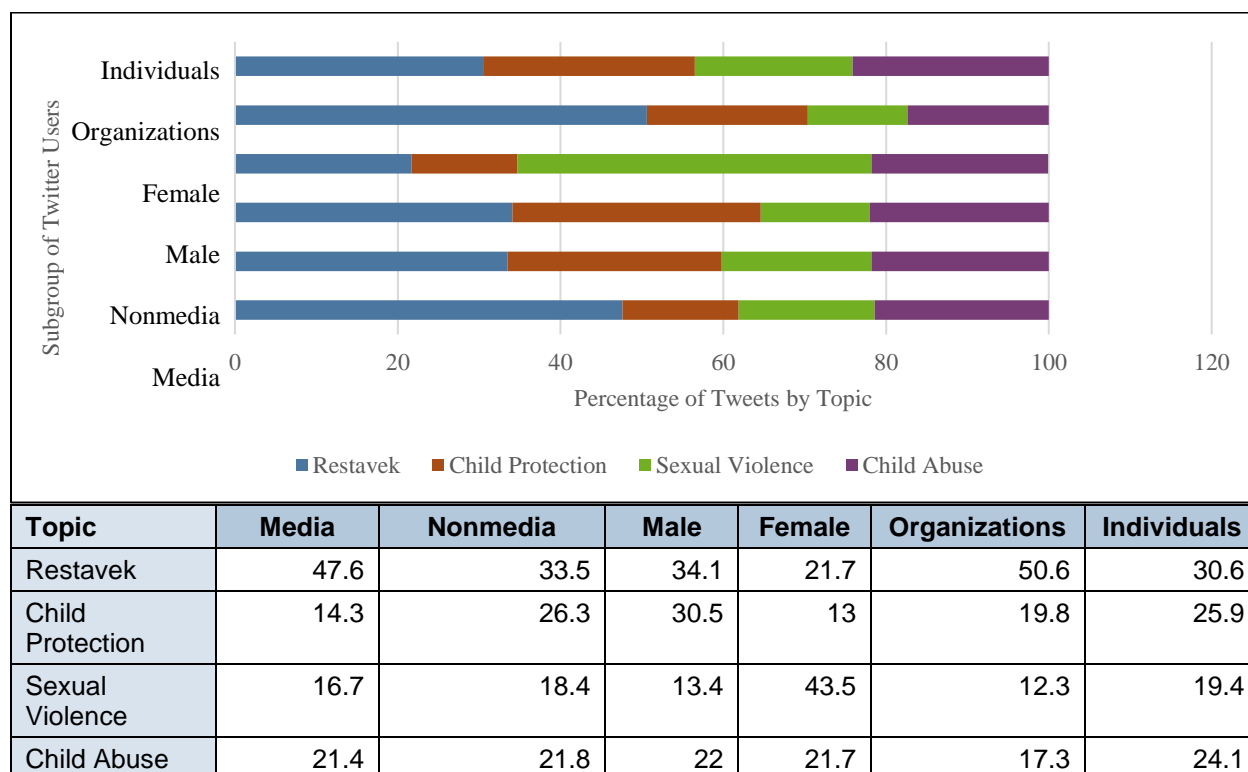
## **Media organizations have significantly more activity but nonmedia have more influence**

Last, most organization users posting to Twitter in Haiti are nonmedia organizations. Of all users, 81.7% are nonmedia and 18.3% are media. Media organizations have a higher number of followers and friends (16,081 and 3,297, respectively) than nonmedia organizations (1,913 and 1,557 followers and friends, respectively). Both nonmedia and media organizations post at about the same rate, with only a 0.1% difference at about an average of two tweets for each organization. However, the nonmedia organizations have significantly more influence (at least in the area of child violence) than media organizations do. Nonmedia organizations have an average of 5.5 retweets for tweets about violence against children compared with an average of 0.8 retweets for media organizations. Nonmedia organizations are not necessarily more influential in general, but they are more influential in the child violence space specifically.

## **Findings: What Are These People and Organizations Saying?**

Based on the top discourse themes developed in topic modeling, (Child Abuse, Sexual Violence, Child Protection, and Restavek), displayed in Figure 8, the distribution comparison between how often individuals and organizations, females and males, and media and nonmedia depicts differences in how individual subgroups are discussing the topic of child violence and protection. In particular, female individuals are more likely to have discourse on “sexual violence” than “child protection” or “restavek” compared with their male counterparts. All subgroups tend to speak about “child abuse” at the same rate and media organizations tend to focus their discourse more on “restavek” topics.

**Figure 7. Percentage of Tweets by Topic for Subgroups of Twitter Users**



## Findings: If Not Child Violence and Protection, Then What Are Organizations Saying on Facebook?

The AIR research team did not find enough posts relevant to child abuse on Haitian and International NGO Facebook pages to perform analysis similar to that used for the Twitter data. Most Facebook posts revolved around other issues facing Haiti (disaster relief and health issues) and solicitations for donations.

Many focused on recovery from Hurricane Matthew.

*The dire situation is persisting in Jérémie and the surrounding villages in the wake of #HurricaneMatthew, and the Health Foundation needs donations now more than ever.*

Some posts gave thanks to donors.

*Un grand merci à Alice Bonhomme-Biais et aux généreux sponsors qui ont permis d'équiper l'école Le Réconfort de Torbeck (Sud d'Haïti) en tableau numérique et en panneaux solaires. Cette école avait beaucoup souffert de l'ouragan Matthew.*

*Dans quelques semaines, Alice refait un triathlon pour équiper/renforcer l'installation solaire d'une école à Fonds Jean Noël.*

*Ce sera la troisième école équipée grâce au triathlon d'Alice aux États-Unis (la première étant une école aux Abricots, dans la Grande Anse).*

Other posts sought donations.

*We could use some help!*

*Due to all of the generous donations people have poured in, we need some help getting them to !*

*We have filled a container packed full of supplies we desperately need. This container holds things like food, baby supplies, medicine and so much more. The cost to ship this giant container will be about \$2,500. That's a lot of money! So we are prayerfully asking for your support. Please consider donating to help us get supplies to the orphanage.*

*Message us directly or email reachoutto @yahoo.com if you would like to contribute! Your support means the world to us!!*

Still other posts were just short captions of photographs.

*Nous sommes rentrés d'Haïti. En photos, une partie de la mission. Voyez comment nous avançons*

Overall, these posts focused on the operations of the nonprofits soliciting help, thanked those who have helped, and discussed the success of their projects or efforts. Figure 8, which shows the main words being used in NGO's Facebook posts, demonstrates that children, development, and Haiti were central concerns in these posts.



## Appendix A. List of Radio Stations

This table includes the 33 radio stations that were contacted and supplied programming information for this project. The full list can be found: [https://docs.google.com/spreadsheets/d/127ee8W97tNlIm\\_VR0RZIYBbVlSh1XH2r4INOSfOkGK4/edit?usp=sharing](https://docs.google.com/spreadsheets/d/127ee8W97tNlIm_VR0RZIYBbVlSh1XH2r4INOSfOkGK4/edit?usp=sharing)

Station	Frequency	Ville/Dept	Department	Number of Contacts (Telephone & In-Person)	Population of Audience	Relevant Programming
RADIO LJS	100.3	SAINT MARC	Artibonite	1 contact	Plus de 25,000	Yes
RADIO POWER	104.5	OUANAMINTHE	Nord Est	1 contact	Unknown	Yes
RADIO RAYON FM	98.3	MARCHAND DESSALINES	Artibonite	1 contact	Unknown	Yes
RADIO TÉLÉ DESSALINES	96.7	MARCHAND DESSALINES	Artibonite	1 contact	Plus de 25,000	Yes
RADIO MASSACRE	89.7	OUANAMINTHE	Nord Est	1 contact	Unknown	Yes
RADIO BELVAL	97.1	Leogane	Ouest	2 contacts	Unknown	Yes
RADIO LA VOIX DE L'EVANGILE	95.3	Port-au-Prince	Ouest	2 contacts	Unknown	Yes
RADIO LUMIERE	92.1 / 88.1	Port-au-Prince / GONAIVES	Ouest	2 contacts	Unknown	Yes
RADIO EXPLOSION	96.5	GONAIVES	Artibonite	3 contacts	10,000 - 15,000	Yes
RADIO GAMMA	99.7	FORT-LIBERTE	Nord Est	3 contacts	5,000 - 10,000	Yes
RADIO MELODIE	103.3	Port-au-Prince	Ouest	3 contacts	15,000 - 20,000	Yes
RADIO KYSKEYA	88.5	Port-au-Prince	Ouest	4 contacts	Unknown	Yes
RADIO VISION 2000	99.3 / 98.1	Port-au-Prince ; GONAIVES	Ouest	4 contacts	Unknown	Yes
RADIO SOLEIL	105.7	Port-au-Prince	Ouest	5 contacts	Plus de 25,000	Yes
RADIO TÉLÉ MÉGA MAX	102.3	GONAIVES	Artibonite	5 contacts	10,000 - 15,000	Yes
RADIO SOLIDARITE	107.3	Port-au-Prince	Ouest	7 contacts	Plus de 25,000	Yes

Station _	Frequence	Ville/Dept	Department	Number of Contacts (Telephone & In-Person)	Population of Audience	Relevant Programming
RADIO PACIFIC	101.5	Port-au-Prince	Ouest	NGO Questionnaire	Unknown	Yes
RADIO TIMOUN	90.9 / 89.5	Port-au-Prince / GONAIVES	Ouest	NGO Questionnaire	Unknown	Yes
RADIO NATIONALE	102.1 / 105.3	Port-au-Prince	Ouest	12 contacts	10,000 - 15,000	No
CECOYA	90.7	Leogane	Ouest	3 contacts	Unknown	No
RADIO PLANET KREYOL	97.1	Port-au-Prince	Ouest	3 contacts	Unknown	No
RADIO M5	93.5		Ouest	4 contacts	10,000 - 15,000	No
RADIO CARAIBES FM	94.5	Port-au-Prince	Ouest	5 contacts	Plus de 25,000	No
RADIO GALAXY	104.5	Port-au-Prince	Ouest	5 contacts	10,000 - 15,000	No
SKY FM	96.5	Port-au-Prince	Ouest	5 contacts	Plus de 25,000	No
RADIO ONE	90.1	Port-au-Prince	Ouest	6 contacts	10,000 - 15,000	No
RADIO SIGNAL FM	90.5	Port-au-Prince	Ouest	6 contacts	Plus de 25,000	No
RADIO MAGIK 9	100.9	Port-au-Prince	Ouest	7 contacts	Plus de 25,000	No
RADIO MEGASTAR	97.3	Port-au-Prince	Ouest	7 contacts	Plus de 25,000	No
RCH 2000	96.1	Port-au-Prince	Ouest	8 contacts	15,000 - 20,000	No
RADIO INDEPENDANCE	101.5	GONAIVES	Artibonite	9 contacts	5,000 - 10,000	No
TOUCHE FM	91.9	Leogane	Ouest	3 contacts	Plus de 25,000	No



## Appendix B. Social Media Data Collection

The AIR team researched the possibility of using Facebook, Instagram, Twitter, WordPress, Blogger, Disqus, and WhatsApp to collect data from Haitian users to describe the influences and followers of people conducting or speaking about child violence and protection in Haiti. Next, the information gathered about each source of social media data is described.

### Facebook

What is Facebook: Facebook is a social media application focused on connections among “friends” (mutually connected users) who share posts, photos, and videos. Friends can privately chat with one another through Facebook messenger. Users can interact with their friend’s pages—or with public pages, which generally belong to organizations—by liking, sharing, and commenting on content as well as by tagging users in photographs.

Privacy and security: Individuals can determine whether their posts and photographs are visible to the public or only to their friends on the application. Many organizations have public pages that are viewable by anyone—not just those who have connected to that organization. According to the Facebook terms of service, data from these public pages (including posts, likes, and shares) can be freely collected through the Facebook API, but access to private pages is strictly prohibited.

What this means for data collection: Because of the prohibition on collecting data from private individuals’ pages, Facebook data were only collected for nonprofits (both Haitian nonprofits and international nonprofits working in Haiti).

### Instagram

What is Instagram: Instagram is a social networking application used to share photos and videos from a smartphone. Similar to Facebook and Twitter, a user creates an Instagram account and upon login has a news feed and a profile. Users can follow other Instagram accounts and interact by commenting, liking, tagging, and private messaging.

Privacy and security: All users of the latest version of Instagram can opt to set their photos and videos to private so they will only be visible to approved followers. The privacy settings work the same for hashtags—the post with a searched hashtag will not appear publicly if it is from a private account. However, if private Instagram accounts share a photo on another social media platform that is not set to private (for example, Facebook), the photo will be publicly accessible by the privacy settings of the other platform.

What this means for data collection: With the current API restrictions, the AIR team concluded that although data can be collected from Instagram, it was too labor intensive for a limited few data points relevant to this project.

## Twitter

What is Twitter: Twitter is a social networking platform that enables users to share short messages (around 140 characters)—called tweets—with others. All tweets are publically viewable, and users can follow or friend other users in order to see their tweets regularly in their Twitter feed. Users can also retweet (share with their own followers) a tweet from another users. Twitter is used by both organizations and individuals.

Privacy and security: All Twitter posts are public, as are the likes, shares, and retweets associated with these tweets. Twitter posts are accessible via API based on user-authentication tokens.

What this means for data collection: The AIR research team collected tweets from both organizations and individuals that mentioned violence against children through the free public API.

## Wordpress

What is Wordpress: Wordpress is a content management or blog-publishing platform which is mainly used to share longer content in many different media types (text, images, video, etc.)—often called ‘blogs’. Users can customize the theme of their own WordPress website by choosing their preferred templates and adding a set of available functionalities. Users can retrieve data on their own websites including posts, pages, comments, categories, tags, custom fields, number of visitors from the WordPress database. Some statistics about the behavior of the visitor might also be collected.

Privacy and security: WordPress.org collects nonidentifying browser and server-supplied information on users who register to host their own WordPress websites, such as browser type, language preference, citing sites, date, and time of each visitor request. IP addresses—potentially identifying information—are also recorded but not disclosed or used to identify users. WordPress.org can publish nonidentifying aggregated data like trends in the usage of its website.

What this means for data collection: With the current API restrictions, the AIR team concluded that although data can be collected from WordPress, it was too labor intensive for a limited few data points relevant to this project.

## Blogger

What is Blogger: Blogger (also known as Blogspot) is a blog-publishing platform owned by Google. It is similar to WordPress except for two main differences: Blogger owns users’ websites and takes responsibility for maintenance and control. Users also have access to their website statistics including number of visitors and behavior of visitors.

Privacy and security: Since Blogger is often connected to a Google account, it collects users’ personal information and data on the usage of services. Blogger may share non-personally identifiable information, such as trends, publicly and with its partners (publishers, advertisers, etc.).

What this means for data collection: With the current API restrictions, the AIR team concluded that although data can be collected from Blogger, it was too labor intensive for a limited few data points relevant to this project.

## **Disqus**

What is Disqus: Disqus is a networked community platform used by websites to allow users to comment on webpages and blogs. It automatically adapts to site designs and colors, and it is supported on mobile and desktop devices.

Privacy and security: All public posts are available for automated data collection using the Disqus API. Data available via the API include topics (identification and name of topic), websites (website comments, threads, and settings), categories (grouped threads), threads (core discussions on Disqus), posts (individual comments within threads), and user information (registration information). Unlimited access to the data is limited to users with a Pro-level subscription, and limited access is tiered based on the types of data queried.

What this means for data collection: It was difficult to assess the quality and relevancy of the potential data to be extracted. As such, the AIR research team concluded that the cost of purchasing API access was cost prohibited in the uncertainty of collecting relevant data.

## **WhatsApp**

What is WhatsApp: WhatsApp is an encrypted endpoint-to-endpoint messaging and calling (voice and video) service that is available on phones worldwide. It is second only to Facebook in messaging usage. With WhatsApp on the web and desktop, users can seamlessly sync all chats to the computer so one can speak with contacts on whichever device is convenient for the user.

Privacy and security: All users of the latest WhatsApp versions have end-to-end encryption that ensures that only those communicating can read what is sent and nobody in-between, not even WhatsApp. Messages are secured with a lock and only the recipient and the sender have the special keys to unlock and read the messages.

What this means for data collection: There is currently no legal way to collect messages via WhatsApp for research purposes.

## Appendix C. Twitter Keyword List and Resulting Data

### English Keywords

English	Total Tweets Collected	Total Relevant Tweets	Example Relevant Tweet
sexual abuse	25	14	RT @Jacquiecharles United States Ambassador Nikki Haley the first to mention @UN sexual abuse scandal in #Haiti
rape	57	10	Indian Supreme Court upholds death sentences in Delhi gang rape
child/children	53	9	UN expert urges Dominican Republic to place child protection at core of tourism strategy
orphan	7	7	If we want to make a real difference for orphans and other children at risk of becoming orphans, we need to...
trafficking	5	6	How Ashton Kutcher is building tech to fight child sex traffickers
abuse	109	2	How to Find Healing From Sexual Abuse
refugees	29	2	#Refugee women and children fleeing conflict and poverty in Africa face beatings rape & starvation in Libya
bit	349	0	
fight	119	0	
crush	67	0	
child abandonment	39	0	
bother	26	0	
gender based violence	25	0	
stigmatization	24	0	
neglect	18	0	
children on the street	13	0	
whip	11	0	

Nonrelevant keywords that resulted in fewer than 10 tweets collected: errant (nine tweets), trauma (nine tweets), sexual violence (eight tweets), tame (eight tweets), bullying (seven tweets), youth (seven tweets), homeless (six tweets), beggar (five tweets), intimidate (four tweets), beat (one tweet), child labor (one tweet), child protection (one tweet), corporal punishment (one tweet), distress (one tweet), human trafficking (one tweet), street youth (one tweet), and punish (one tweet).

Keywords that resulted in zero tweets collected: discipline, displaced persons, exploitation, gray, maltreatment, offender, paralyze, physical abuse, remonstrance, serving as domestic, stateless persons, street children, tanner, tramp, vagabond, violence, and whip in school.

## French Keywords

French	Total Tweets Collected	Total Relevant Tweets	Example Relevant Tweet
violence	167	67	Peer educators ready to better tackle sexual gender based violence in ...
enfants	50	35	Session de sensibilisation sur la protection de l'enfance a l'ecole Methodiste de Duplan
jeunes	64	24	De jeunes #haitiens qui auraient du tre En train d'tudier pour prparer la relve sont esclaves ailleurs
abus	12	4	#Haiti Le Canada considere donner davantage de fond l'ONU pour les victimes d'abus sexuels
battre	42	2	Pour moi un vrai homme ne doit pas battre sa femme mais lui faire battre son cur - Beyonce
enfants dans la rue	19	2	Plus de 2 800 enfants dont plusieurs dizaines de fillettes vivent dans la rue selon l'Institut du Bien-Etre
harcelement	3	1	Ohia USA Victime de harclement un petit garon de 8 ans se suicide
discipline	85	0	
traumatiser	79	0	
exploitation	33	0	
vagabond	23	0	
orphelin	19	0	
discipliner	15	0	
delinquant	13	0	

Nonrelevant keywords that resulted in fewer than 10 tweets collected: maltraitance (nine tweets), ditresse (seven tweets), refugies (six tweets), Personne déplacée (five tweets), bataille ou combat (three tweets), ecraser (three tweets), intimider (three tweets), jeune de rue (two tweets), negligence (two tweets), protection de l'enfant (two tweets), punir (two tweets), tanner (two tweets), violer (two tweets), agression physique (one tweet), clochard (one tweet), dresser (one tweet), ennuyer (one tweet), and fouet (one tweet).

Keywords that resulted in zero tweets collected: apatride, bastonner, battre ou fouetter à l'école, chicoter, délinquant, détresse, domesticité, donner une leçon, enfant abandonné, enfants de rue, faire la peau, figer, fouetter, grappilleur, harcèlement, mendicant, mesure coercitive, punition corporelle, réfugiés, remontrance, réprimande, réprimander, sans domicile fixe, stigmatization, tétaniser, trafic humain, traumatisme, travail des mineurs, violence basée sur le genre, and violence sexuelle.

## Creole Keywords

Creole	Total Tweets Collected	Total Relevant Tweets	Example Relevant Tweet
restavek	56	25	Through the month of March a generous supporter of Restavek Freedom is giving \$1 for every Facebook page like
vyolans	32	3	Li ta two fasil pou nou ta ap kritike elev ki ap fe vyolans san nou ap blayi sou delenkans leta a ki apt janm...
bat	82	2	Fi ki aksepte yon nag bat li se fat paran li. #TeamTim
abi	8	2	Manman ak papa se premye #responsab #pwoteksyon pitit yo kont tout fam abi, #eksplwatasyon ak vyolans. #Eliminerlaviolence #timoun
batay	65	1	#fetpititfim #Mesi ou banm + raj pou batay pou Dwa + pase 50 % tifi ki sibi tout fom diskriminasyon epi touyeâ...
goumen	123	0	
make	107	0	
korije	61	0	
kokorat	48	0	
kale	38	0	
san peyi	35	0	
san manman	26	0	
sanzave	26	0	
moun yo leve	11	0	

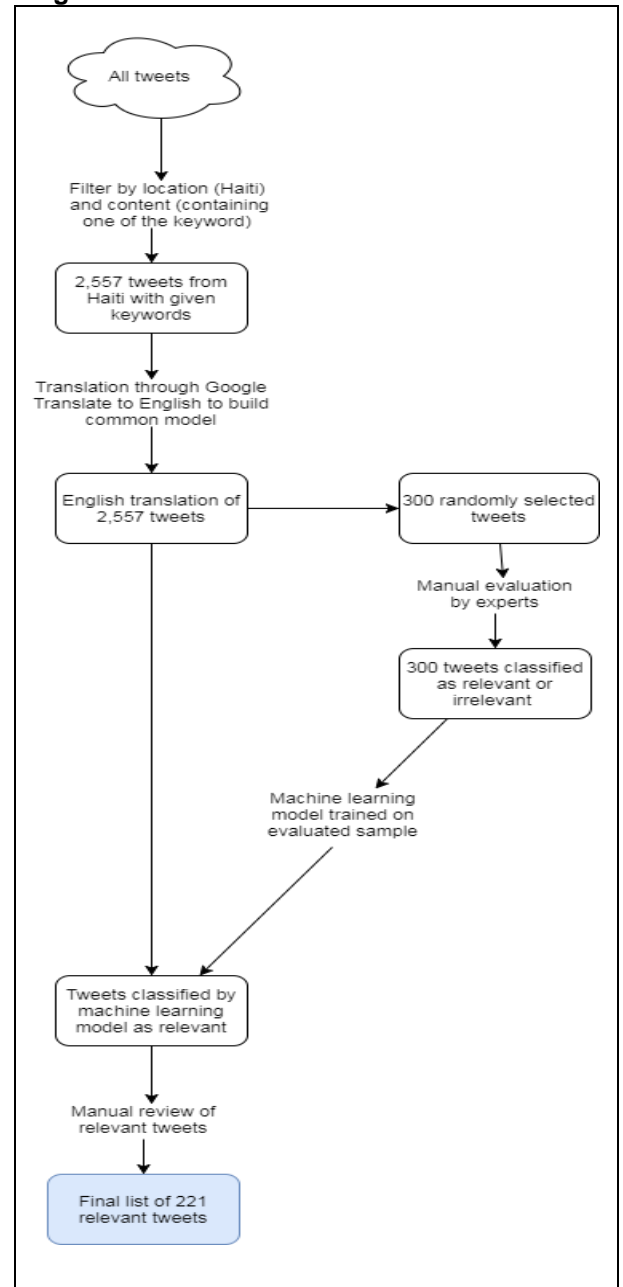
Nonrelevant keywords that resulted in fewer than 10 tweets collected: rigwaz (nine tweets), sentaniz (nine tweets), travay timoun (nine tweets), vyole (nine tweets), bay leson (eight tweets), bay presyon (eight tweets), anmegde (seven tweets), fwet (five tweets), neglijans (five tweets), san fanmi (five tweets), san papa (five tweets), maspinen (four tweets), san sal (four tweets), atoufe (three tweets), toupizi (three tweets), matinet (two tweets), pini (two tweets), timoun (two tweets), explwate (one tweet), freeze (one tweet), grapiay (one tweet), malmennen (one tweet), move tretman (one tweet), san zabri (one tweet), sou chok (one tweet), ti joel (one tweet), trafik moun (one tweet).

Keywords that resulted in zero tweets collected: ak matinet, an detres, sanzespwa, apatrid, bat nan lekòl, deplasman fose, discipline, donte, farouch, fe po leve anba baton, fwete, jen nan la ri, karikati timoun, kraponnen, malandrin, maltretans, mandian, maspinay, protection timoun, refijye, tetanize, timoun abandone, timoun la ri, timoun nan la ri, violyans seksyel, vyolans sou fanm, and woule anba baton

## Appendix D: Twitter Relevance Model

Although the AIR team was able to collect 2,557 tweets from Twitter, not all of those tweets were relevant to child violence and protection topics. Figure 5 shows the process of identifying relevant tweets. Tweets were selected that contained one of the given keywords and came from Haiti. Then, to classify the collected tweets as relevant or irrelevant, the team developed a supervised machine learning<sup>21</sup> model that used a set of hand-classified relevant and irrelevant tweets to teach the classifying algorithm which of the remaining unclassified tweets were relevant. For the purposes of relevance modeling, all French and Creole tweets were translated to English through Google Translate in order to build a common model. Three research team members assessed a randomly selected sample of 300 tweets (weighted toward more relevant tweets because of the relatively smaller number of relevant tweets) and classified them based on relevancy. Disputes between the three experts were reviewed and decided based on majority vote. The final set of tweets were modified to remove stop words (words like a, the, is, etc.) and then converted to a list of words (referred to as a bag-of-words approach) to serve as input to a Random Forest model (a particular type of machine learning model).<sup>22</sup> The model then predicted whether a tweet was relevant or irrelevant based on the words in the tweet. After classification, all tweets were reviewed manually and the classification

Figure 9. Process of Tweet Evaluation



<sup>21</sup> Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2012). *Foundations of machine learning*. Cambridge, MA: MIT Press.

<sup>22</sup> Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3), 146–162.

was updated if an error was detected. The final model had a cross-validation accuracy<sup>23</sup> of 93%, precision<sup>24</sup> was 96%, and recall<sup>25</sup> was 89%.

---

<sup>23</sup> Accuracy is defined as the number of tweets correctly classified divided by the total number of tweets.

<sup>24</sup> Precision is defined as the number of tweets classified as relevant that are actually relevant divided by the total number of tweets classified as relevant. This measure determines how precise the model is in only classifying tweets that are actually relevant as relevant. Our high precision means that the model was rarely classifying irrelevant tweets as relevant. See also Powers, D. M. W. (2011). Evaluation: From precision, recall and *f*-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.

<sup>25</sup> Recall is defined as the number of tweets classified as relevant divided by the total number of tweets that are actually relevant. There is generally a trade-off between precision and relevance. This trade-off is discussed in the previous footnote.



## ABOUT AMERICAN INSTITUTES FOR RESEARCH

Established in 1946, with headquarters in Washington, D.C., American Institutes for Research (AIR) is an independent, nonpartisan, not-for-profit organization that conducts behavioral and social science research and delivers technical assistance both domestically and internationally. As one of the largest behavioral and social science research organizations in the world, AIR is committed to empowering communities and institutions with innovative solutions to the most critical challenges in education, health, workforce, and international development.



AMERICAN INSTITUTES FOR RESEARCH®

1000 Thomas Jefferson Street NW  
Washington, DC 20007-3835  
202.403.5000

[www.air.org](http://www.air.org)

*Making Research Relevant*

## LOCATIONS

### Domestic

Washington, D.C.  
Atlanta, GA  
Austin, TX  
Baltimore, MD  
Cayce, SC  
Chapel Hill, NC  
Chicago, IL  
Columbus, OH  
Frederick, MD  
Honolulu, HI  
Indianapolis, IN  
Metairie, LA  
Naperville, IL  
New York, NY  
Rockville, MD  
Sacramento, CA  
San Mateo, CA  
Waltham, MA

### International

Ethiopia  
Haiti  
Honduras  
Kyrgyzstan  
Tajikistan  
Zambia