How Twitter Can Support the HIV/AIDS Response to Achieve the 2030 Eradication Goal: An In Depth Thematic Analysis of World AIDS Day Tweets

Michelle Odlum, EdD, MPH1*, Sunmoo Yoon, RN, PhD1*, Peter Broadwell, PhD2*, Russell Brewer, DrPH1*, Da Kuang, PhD4*

1. Columbia University School of Nursing, New York, NY 10026
2. University of California Los Angeles: Charles E. Young Research Library, Los Angeles, CA 90095
3. Louisiana Public Health Institute, New Orleans, LA 70112
4. Amazon.com, Seattle, WA 98109

§ Corresponding author: Michelle Odlum

Columbia University School of Nursing
560 W 168th Street
New York, NY 10026
mlo12@columbia.edu
212-342-3837 (w)
212-305-6937 (f)
347-806-2090 (c)

* These authors have contributed equally to the work.
Abstract

Background: HIV/AIDS is a tremendous public health crisis, with a call for its eradication by 2030. A human rights response through civil society engagement is critical to supporting and sustaining HIV eradication efforts. Yet, ongoing civil engagement is challenging.

Objective: The current study sought to demonstrate the use of Twitter data to assess public sentiment to support civil society engagement.

Methods: Tweets were collected during World AIDS Days 2014 and 2015. A total of 39,940 unique Tweets (>10 billion users) in 2014 and 78,215 unique Tweets (>33 billion users) in 2015 were analyzed. Response frequencies were aggregated using natural language processing. Hierarchical rank-2 NMF algorithm generated a hierarchy of tweets into binary trees. Tweet hierarchy clusters were thematically organized by UNAIDS core action principles and categorized under HIV/AIDS Prevention, Treatment/Care, or Support.

Results: Topics tweeted ≥35 times were visualized. Results show a 2015 decrease in frequency of tweets associated with fighting to end HIV/AIDS, the recognition of women and achieving an AIDS-free generation. Also an increase in tweets associated with an integrative approach to the HIV/AIDS response. Hierarchical thematic differences in 2015 included: no prevention discussion and the recognition of the pandemic’s impact and discrimination. Additionally, a decrease was observed in motivation to fast track the pandemic’s end and combat HIV/AIDS.

Conclusions: The human right-based response to HIV/AIDS eradication is critical. Findings demonstrate the usefulness of Twitter as a low cost method for assessing public sentiment to enhance knowledge, increased hope, and revitalized expectations for HIV/AIDS eradication.

Keywords: Community, Human rights, Key and vulnerable populations, Structural drivers
Introduction

The United Nations - Sustainable Development Goals (SDGs) seek a holistic and balanced approach to social, economic, and environmental aspects of development[1, 2]. The SDGs emphasize the need to advance and complete the objectives set forth by the Millennium Development Goals (MDGs), including a call to end the HIV/AIDS pandemic by 2030 [1], [3, 4]. An effective response to the pandemic’s end demands critical, dedicated, and sustained action. However, the SDGs [2] have a board health goal, not recognizing HIV/AIDS as a distinct focus area.[1, 2],[3] This is a result of the vertical prioritization of HIV/AIDS during the 2000-2015 MDG period,[4] that reduced the allocation of resources from other important health issues [1, 2, 5]. With limited visibility in the post-2015 agenda and lack of additional resources to scale up efforts,[1, 6] achieving the 2030 eradication goal is of great concern.

The MGDs resulted in the expansion of 15 million people living with HIV/AIDS (PLWH) on life saving antiretroviral drugs (ARVs) [1, 3, 4],[7]. Although the need to scale up and sustain these biomedical solutions is recognized, human-rights issues (e.g., stigma, marginalization, discrimination) still serve as pervasive barriers to successful adoption [3, 8]. Yet, if HIV/AIDS ceases to be perceived as an ongoing global health emergency, the necessary services and resources to sustain and expand eradication efforts will rapidly diminish. Grassroots activism[8-10], and civil society mobilization [10, 11] are critical driving forces for the advancement of human rights and play a major role in the global response to HIV/AIDS. [12] Early in the pandemic, civil society organizations understood the limitations of a solely biomedical focus and need to progress a human rights approach, resulting in the global scale up of access to ARVs [3, 11]. Further strengthening of the human rights response can support and sustain eradication efforts.

According to the Joint United Nations Programme on HIV/AIDS (UNAIDS), strong civil society engagement is critical to the eradication of HIV/AIDS [11],[12]. Ongoing human rights related efforts; including giving voices to PLWH and empowering marginalized populations, are essential for the successful mobilization of treatment and prevention resources [10]. In fact, the SDGs demand grassroots activism,[4, 8, 9, 11] calling for a greater investment and support to civil society to achieve the eradication goal.[1, 2, 13] Effective civil society activism[8, 11] must
engage advocacy networks, private and public institutions, and global policymakers to advance the 2030 campaign [14]. However, the life cycle of activism [8, 9, 12] and its ability to sustain influence and engage citizens is one of the greatest challenges and opportunities to the eradication of HIV/AIDS. Furthermore, the success of civil society activism[8, 9] lies in exploring public sentiment to guide the effective exchange of information for improved knowledge, increased hope, and revitalized expectations.[11],[12]

The utilization of social networking sites (SNS) for civil society activism [8, 9] is a promising approach to help sustain and drive HIV/AIDS-related social movements. SNS, like Twitter, have played a vital role in the organization of global movements.[15] Twitter is a very powerful and popular microblogging communication tool.[16] Users share information through 140-character messages called tweets. Information is disseminated through direct messages or the forwarding (retweeting) of messages for broad propagation.[15, 17]

Twitter is a digital and far-reaching tool to study human behavior; capturing aspects of people, which can be missed through traditional data collection strategies in health research. Twitter is effectively used to predict disease outbreaks, including the flu and HIV, and has informed a variety of public health efforts. Public sentiment, expressed in tweets, provide a wealth of information utilized by public health professionals, politicians, governmental entities, activists and computer scientist, to engage in purposeful discussions and to play active roles around a variety of topics [11],[16],[17]. Moreover, Twitter has the capability to identify health trends and to support interventions and health campaigns for improved motivation, and behavior.

Twitter is also effectively used to assess and address health information needs during disease outbreaks, such as Ebola. [15, 17-19] With over 645 million registered users and the distribution of >58 million tweets daily, Twitter is a reliable source for tracking public sentiment to guide discussions for effective health awareness campaigns.[14, 17, 18]. Public motivation is essential to sustain the global HIV/AIDS response, and to achieve our global eradication goals. The current study sought to demonstrate the use of Twitter to explore HIV/AIDS public sentiment over a period of time, [15] to help guide social movements in support of HIV/AIDS eradication efforts which are guided by UNAIDS core action principles of a comprehensive HIV/AIDS
response [20] and commitments necessary to reach the 2030 goal of HIV Prevention, Treatment, Care, Support. Our study provides a unique and in depth analysis of Worlds AIDS Day tweets in 2014 and 2015. World AIDS Day, held on December first every year provides an ideal opportunity to assess public sentiment over time as people unite worldwide to support PLWH, honor those lost, and to combat HIV/AIDS [21].

Methods

Tweet Corpus
During the World AIDS Days of December 1, 2014 and December 1, 2015, tweets mentioning HIV/AIDS were collected from Twitter (https://twitter.com/) via a Google Chrome-based version of NCapture™, a web crawler that captures internet-based text.[22] The streaming Application Programming Interface allowed for the capture of a limited sample of all tweets (e.g., 18,000 tweets per 15 minutes) [22]. To overcome the challenges of time and amount limits, tweets were collected in 15-minute intervals for a representative sample [22]. Keywords were used for searching tweets mentioning HIV/AIDS (e.g., #HIVtreatment, #HIVservices, #HIVprogramming and #HIVprevention). Additional data elements collected were time stamps, content, geographical locations from IP addresses, user names, message type (unique or retweet), and followers (number of disseminated) [22]. English language tweets were included in the analysis with 39,940 unique tweets (10,027,038,772 users) in 2014 and 78,215 unique tweets (33,370,938,359 users) in 2015.

Natural Language Processing
Natural language processing was conducted to identify and depect topics of collected tweets about HIV/AIDS. Tweets were cleaned and transformed to a vector form and N-gram. [22] An N-gram is a subsequence of N items in a given sequence from characters to words. The N-gram method was used to compute a tweet term-frequency dictionary. [22] Notepad++ and Weka 3.7 reduced the dimensionality for the algorithmic processing of the data. Snowball stemmers were used to apply the stemming algorithm Porter’s algorithm, an affix-removal approach, which was applied through Weka [22]. To further remove dimensionality, stemming was conducted to identify a word root and for removing suffixes and prefixes. Tweet topics were detected and
summarized through descriptive statistics (e.g., frequency count), classification, visualization, and clustering.

Hierarchical rank-2 NMF algorithm and Rank-2 NMF
Clustering, the process of grouping a set of words into classes of similar topics was conducted [23]. HierNMF2 was used to determine the semantic organization of tweeted words [23], [24]. Tweet topic clusters were generated by HierNMF2 and visualized into tree nodes for 2014 and 2015 data to illustrate the topic structure. The Hierarchical rank-2 NMF algorithm used for clustering generated a hierarchy of tweet themes into binary trees [23]. A node-splitting rule was also employed to determine tree nodes to split from the original binary nodes. This methodology allows for the determination of tree structures. Data for each time point (2014 and 2015) was split into two clusters, creating a binary tree [23]. Rank-2 NMF was applied to generate the hierarchical tree structure. Each cluster created non-leaf nodes. A score was computed for each non-leaf node using Rank-2 NMF [23, 24]. Non-leaf nodes with the highest scores were then split into leaf nodes of two or more well-separated topics.

Thematic Analysis
The tweet hierarchy clusters were characterized based on content by subject matter experts. A coding framework was used to interpret and explore the data and to identify perceptions of the HIV/AIDS pandemic. Cluster analysis occurred iteratively, with the refining of major themes through review and discussion to shape the final coding structure. Themes were organized under overarching themes categorized by UNAIDS core action principles guiding a comprehensive response to HIV/AIDS (Table 1). Themes were then grouped according to the commitments necessary to reach the SDG’s 2030 goal [4]: Prevention, Treatment/Care, and Support.

Results
Response Frequencies
The geographic spread of HIV/AIDS-related tweets on World AIDS Days 2014 and 2015 spanned the globe (Figure 1). Top disseminators were United Nations agencies followed by
celebrities including singers, models, actors, and US governmental organizations and political figures. Topics tweeted 35 times or more were aggregated. Response frequencies were generated and compared for increased and decreased frequency between 2014 and 2015. Results show increased frequency in tweets associated with an integrative approach to HIV prevention, treatment, and care (e.g., community). An increase was also observed in the frequency of tweets associated with the recognition of barriers to HIV/AIDS eradication (e.g., stigma). A moderate decline was observed in tweets associated with ending the epidemic (e.g., fast track) and the provision/utilization of services (e.g., access). A significant decrease in tweet frequency associated with combating the epidemic (e.g., fight) were also observed (Figure 2).

Hierarchical clusters of tweets
PLWH (leaf node: Prevent HIV/AIDS) vs. Spread HIV/AIDS Awareness (leaf nodes: Commitment to End the Pandemic, Save Lives: Infected and At Risk and Halt Infections (Figure 3).

Figure 3. World AIDS 2014 Tweet hierarchy

Thematic Analysis


### Table 1. Thematic Analysis of 2014 and 2015 World AIDS tweet hierarchy

<table>
<thead>
<tr>
<th>Hierarchy</th>
<th>Themes</th>
<th>2014</th>
<th>2015</th>
<th>Representative Example Tweets</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PREVENTION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1° 2°</td>
<td><strong>Comprehensive Global Response to HIV/AIDS:</strong> UNAIDS Core Actions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Prevention, Treatment/Care &amp; Support</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4° 5°</td>
<td><strong>Decrease Vulnerability of Acquiring HIV/AIDS</strong></td>
<td>Informational Resources from Governmental Organizations</td>
<td>infographics, AIDS.gov</td>
<td>white house, AIDS.gov</td>
</tr>
<tr>
<td>5° 4°</td>
<td>Halt Infections</td>
<td>accessible, affordable, vaccine partnerships</td>
<td>epidemic, averted</td>
<td></td>
</tr>
<tr>
<td>4° -</td>
<td>AIDS Free Generation</td>
<td>Close the gap</td>
<td>AIDS free gen, children</td>
<td></td>
</tr>
<tr>
<td>4° -</td>
<td>Prevent HIV/AIDS</td>
<td>AIDS prevention, HIV facts</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>TREATMENT/CARE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5° 3°</td>
<td><strong>Access to Services</strong></td>
<td>Treatment for HIV/AIDS</td>
<td>Access, children, pregnant women services, closing the gap</td>
<td>treat, people</td>
</tr>
<tr>
<td>4° 1°</td>
<td>Save Lives: Infected and At Risk</td>
<td>cure</td>
<td>treatment for all</td>
<td></td>
</tr>
<tr>
<td>- 3°</td>
<td>Impact of the HIV/AIDS Pandemic</td>
<td>-</td>
<td>millions, lives-end, children, can avert, infections-save, world</td>
<td></td>
</tr>
<tr>
<td>1° 1°</td>
<td><strong>Expand Programmes</strong></td>
<td>Efforts for Targeted HIV/AIDS Eradication</td>
<td>facts, wipe homophobia, today</td>
<td>save, join, treatment for all</td>
</tr>
<tr>
<td>3° 6°</td>
<td>Fast Track the End of the Pandemic</td>
<td>today, end, epidemic</td>
<td>UNAIDS, fast track, treatment, response, assessing</td>
<td></td>
</tr>
<tr>
<td><strong>SUPPORT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4° 2°</td>
<td><strong>Safeguard Human Rights</strong></td>
<td>Spread HIV/AIDS Awareness</td>
<td>cure, love, sending, support, fight, AMFAR</td>
<td>awareness, spread, end AIDS</td>
</tr>
<tr>
<td>5° 3°</td>
<td>Recognition of the Pandemic</td>
<td>unaware, United States, HIV</td>
<td>aware, disease</td>
<td></td>
</tr>
<tr>
<td>4° 1°</td>
<td>Honor People Living with HIV/AIDS</td>
<td>honour, memory, continuing, lost, affected</td>
<td>living, people, positive, statement</td>
<td></td>
</tr>
<tr>
<td>4° 2°</td>
<td>Expression of HIV/AIDS Solidarity and Consciousness</td>
<td>wear, ribbon, close the gap</td>
<td>tee shirt ribbon</td>
<td></td>
</tr>
<tr>
<td>2° 5°</td>
<td>Combat HIV/AIDS</td>
<td>fight, today, support, cure, love</td>
<td>fight, helping</td>
<td></td>
</tr>
<tr>
<td>3° 2°</td>
<td>Support People Living with HIV/AIDS</td>
<td>sending, support, fight, cure</td>
<td>support, living, people</td>
<td></td>
</tr>
<tr>
<td>4° 2°</td>
<td><strong>Partnerships and Alliances</strong></td>
<td>Commitment to End the Pandemic</td>
<td>renew, vow, longer</td>
<td>statements, discrimination, make</td>
</tr>
<tr>
<td>2° 5°</td>
<td>Celebrities/Industries</td>
<td>MAC cosmetics viva glam, Kasper, rappers, Victoria</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion

HIV/AIDS is one of the greatest challenges to sustainable social, economic, and civil society development and is affects all sectors of our society [1, 6, 11, 25]. Strong civil engagement to drive social change remains critical to eradicating HIV/AIDS [3, 11]. The current study demonstrated how the analysis of social media data, specifically Twitter could inform purposeful discussions for effective civil society engagement.

Our thematic analysis of the 2014 and 2015 World AIDS Day Twitter hierarchies identified major shifts in public sentiment around a variety of human rights and biomedical related topics. The majority of themes fell primarily under Support, followed by Treatment/Care, then Prevention. In fact, the theme of Prevent HIV/AIDS was present only in 2014. Prevention words were not present in the 2015 hierarchy. An absence of this theme may be a direct result of the treatment as prevention belief. This may indicate the need for purposeful discussion, as prevention needs are not only biomedical (e.g., ARVs). Human rights-based prevention approaches will ensure a more permanent and sustainable solution to achieve HIV eradication. The future cost of daily ARVs in developing countries is not sustainable [7]. Therefore, the revitalization of human-rights-based initiatives is a priority. Early in the pandemic, human rights based approaches focused on health equity worked with populations living with and affected by HIV to expand ARVs in the pipeline and successfully halted the spread of HIV [7, 26] This indicates the need for ongoing civil society engagement in purposeful discussions to revitalize prevention campaigns [14].

Under Treatment/Care, the theme of Impact of the Pandemic was seen only in 2015. The emergence of this theme may indicate a heightened awareness or recognition of the pandemic’s effects on those infected and affected. The theme of Efforts for Targeted HIV/AIDS Eradication remained as a primary level theme in both years. In 2014, tweets discussed the elimination of homophobia and knowledge about HIV/AIDS. Tweets from 2015 discussed treatment, saving lives and joining in efforts. Although discrimination was present in 2015, the recognition of
homophobia was seen only in 2014. This may indicate a need for purposeful discussion, particularly around the importance of supporting key populations [27]. The theme of Fast Track End of the Pandemic decreased three levels on the hierarchy in 2015. In 2014, tweets discussed the end of the pandemic; an indication of action and words of excitement. In 2015, tweets called for an assessment of the HIV/AIDS response. This shift may be a direct reflection of falling short of the 2015 goals and the need for improved effort [3].

In 2015, all themes increased on the hierarchy under Support, except Combat HIV/AIDS, decreasing three levels. Such a decrease is in need of attention, as this theme is comprised of action words necessary to end the pandemic, including fight and support. This shift is a potential indicator of lost enthusiasm, similarly discussed under Treatment/Care. It is also important to note the absence of such words across years based on our response frequency analysis. Decreases were observed in tweets associated with efforts to eradication, (e.g., AIDS-free generation). This further reinforces the need to re-engage the public in purposeful discussions to reinvigorate grassroots efforts.[10] The theme of Reduced Stigma remained a second level tweet in 2015, with 2014 data acknowledging the LGBTQ community and 2015 data discussing the need to end stigma [8, 11, 27, 28]. Maintaining this rank is critical as key populations are disproportionately affected by HIV/AIDS. Key populations are criminalized, marginalized, and plagued with stigma, [29] contributing to HIV risk [8, 28]. Essential services for prevention are often times unavailable to these groups.[27] Furthermore, HIV-associated stigma has contributed to poor access to HIV treatment and care [27]. The theme of Discrimination of PLWH emerged in 2015. This may potentially reflect the November 2014 UNAIDS Fast-Track Strategy report, with zero discrimination recognized as a main target to end HIV/AIDS [3]. However, in the 2015 data on recognition of stigma and discrimination, tweets were not focused on [3, 8] key population or high-risk groups [3, 8, 27]. Results indicate the need for ongoing discussion around such barriers, as civil society plays a major role in supporting marginalized populations [11, 12, 28].

Hierarchical differences also revealed a 2015 absence of tweets mentioning women. This is of great concern, as not addressing issues of gender minorities will thwart eradication efforts. The focus on key populations does very little to change gender inequality [30]. Women are only considered key when they are pregnant, nursing or members of other high-risk groups (e.g., sex
workers) [3]. This approach, neglecting women living with HIV and other vulnerable women, fails to transform norms, beliefs, and perceptions around women’s rights to health and well-being [30]. In fact, 2014 tweets mentioned women only in the context of motherhood. With women comprising 60% of PLWH globally,[30] the biological and social factors making them most vulnerable to HIV infection must be addressed. Persistent discussions around women’s rights remain critical for eradication.

**Limitations.** The generalizability of this study is limited due to the English only analysis and the use of one SNS (i.e., Twitter) [15]. We did not include composed tweets, only those disseminated through tweeting or retweeting. For a more in depth analysis of social media data, future studies should explore other SNSs and analyses in other languages. Twitter’s global and pervasive spread of information can support civil society engagement [15, 19]. Our ambitious targets are critical to ending the pandemic and are possible with the support of technology and social media outlets like Twitter [15, 16]. Civil society’s human rights-based approaches and responses can be limited by material resources [11, 12]. The low cost and ubiquitous spread of information through SNS can diminish such barriers. The vision of zero new HIV infections, zero discrimination, and zero AIDS-related deaths must be transformed into tangible milestones and endpoints,[1, 6] and social media can help support these efforts [3], [16],[18],[19].

**Conclusions.** Our study’s demonstration of Twitter utilization to explore HIV/AIDS public sentiment over time can guide targeted social movement campaigns aimed at addressing grassroots level barriers and heighten public motivation necessary to drive eradication. We also demonstrated the feasibility of the use of cost-effective social networking technologies to identify health-related communication and, the utilization of such platforms to support improved outcomes. In fact, with the ever increasing amount of social media data, and the unique and refined analytic approaches such as ours, HIV/AIDS researchers and global health professionals will soon be able to build upon and enhance their methods, to accurately monitor and support a variety of HIV-related issues, and outcomes.

**Competing interests**

The authors have no conflicts to disclose.
Acknowledgements

NSF and the Institute for Pure and Applied Mathematics (IPAM) UCLA for data analytic support.

Authors’ contributions

All authors assisted in the data interpretation, manuscript drafting, revising of important intellectual content, and final manuscript approval. We have no conflict of interest to disclose.

We had no funding source and ethics committee approval was not required as tweets are de-identified with no identifiable information obtained.

References


