

Anti-Immigrant Computational Propaganda in the United States

Mapping the flow of image-based
alternative narratives in 2018



ABSTRACT

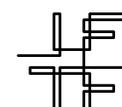
This research is an inquiry into the role that images play in shaping alternative narratives leveraged for targeted harassment and political dysfunction surrounding online activism and immigration justice.

Tweets related to the crisis at the US/Mexico border were collected using the Twitter API over the span of a week during the summer of 2018. The dataset included tweets from people and organizations in support of and opposition to Immigration and Customs Enforcement (ICE) policies and practices. We utilized quoted and retweeted unique media IDs for each piece of visual content to generate a dynamic network of content, connecting images (photos, videos, and GIFs) shared by the same individual users, then applied content analysis to understand how the content and context of these images evolve over time on Twitter. Iterative mixed methods analysis of these images identified high-level trends, and low-level case studies of content appropriation and re-contextualization patterns on social media. In addition, the network analysis is enriched by interviews with WITNESS immigrant organization partners, which detail how the fight for social justice has both benefited from social media affordances and been disrupted by digital disinformation. This study expands the current understanding of how rumors and alternative narratives propagate online; specifically, it contributes new knowledge with regards to the negative impact of digital content manipulation on the narratives surrounding immigrants and refugees within the United States. As more sophisticated, personalized photo and video manipulation techniques emerge, this research addresses the necessity to understand disinformation surrounding visual content and the ramifications on social justice movements.

Erin McAweeney
Palika Makam
WITNESS

ABOUT THESE PAPERS

This case study series explores the social implications of computational propaganda. Each report outlines how tools and tactics, including bots, disinformation and political harassment, were used over social media in attempts to silence social and issue-focused groups prior to the 2018 US midterm elections.



INSTITUTE FOR THE FUTURE



introduction

In the evolving mire of online information practices, the apparently transparent nature of visual content has surfaced as an exceptionally useful and reliable tool in protecting and defending human rights. Activists and the organizations supporting them have demonstrated the value of the image—as an authentic documentation of injustices and a catalyst for social change, facilitated by the horizontal and informal interactions afforded by social media platforms and the increasing accessibility of mobile phones. Visual content has a particularly strong influence on the formation of *narrative* and *frame* within movements, both on and offline, given their use to construct social and cultural representations and identities for groups (Entman, 1993). Pervasive social media use has enabled activists to circumvent hierarchical power structures and share the stories of marginalized groups, creating a new era of online human rights activism driven by visual content.

However, the same features on social media platforms that build important channels for human rights activists are being adapted for insidious influence to create political disarray. Visual mis- and disinformation, both in context and content, have become exceedingly prevalent online, including progressively sophisticated use of synthetic media, “deep fakes,” which pose social, economic, and political threats (World Economic Forum, 2017). If the authenticity of visual content is called into question, communities that rely on accurate documentation of human rights injustices are stripped of tools that have tangible impacts on their livelihood. While the participatory nature of social media has been shown to help individuals foster political will and organize with a collective purpose, bad actors have increasingly exploited the horizontal nature of networks to propagate new narratives and disinformation surrounding politically-charged visual content (Bakdash et al., 2018). Despite the increasing availability of effective and accessible means for visual content manipulation, journalistic and platform responses to “fake news” issues have predominantly focused on text-based mis/dis/mal-information and text-based responses.

This project is a mixed methods network analysis of visual content created and shared by activists online in response to recent US immigration policies and Immigration and Customs Enforcement (ICE) activities. This study is intended to map the evolving contextual narratives surrounding visual content online as users both in support of and opposed to immigration negotiate the interpretations and contents of images and videos. Our analysis recursively integrates quantitative network analysis with qualitative visual and text analysis, drawing on previous work in online rumor propagation and social movement framing, to illustrate the evolving online lives of images and videos.

This research intends to answer the following questions:

- Q1. What common trends in conversational patterns on social media can we identify that develop around images and videos in a social movement context?**
- Q2. What are the more granular patterns in how individual users interact with these images and videos as they deliberate on the content and context of visual content, otherwise known as framing?**
- Q3. How does misinformation develop from the aforementioned framing negotiations of visual content?**

The Methods and Results sections of this paper are organized by these three questions. By following these lines of inquiry our analysis finds:

- While some rumors and misinformation can be siloed in insular online communities, the recontextualization of images requires reappropriation of content between two online communities that otherwise interact very little. Discursive patterns in the framing of images on Twitter, pertaining to a specific social movement, are collectively developed, and the narratives that emerge may become, and in some instances co-opted into, counter-narratives.

background

- Influence and amplification of counter-narratives on Twitter are used to recontextualize an image and, in some instances, gain popularity over the original context. In total, there were five individual instances where a counter-narrative rose to prominence over the original tweet in our dataset. All instances of prominent counter-narratives were propagated through members with significantly more Twitter followers and reached a larger audience than the original image.
- Tactics similar to historical practices of anti-immigrant propaganda were observed in proliferating counter-narratives, including stereotypes, harassment, and calls for violence.
- The patterns and tactics identified in quantitative analysis translate into tangible impacts for human rights organizations that rely on social media to more broadly reach their communities. The transition to image-based immigration rights campaigns invites new tactics for the targeting of immigrants and activists. Interviewers shared stories of digital “know your rights” flyers from their organizations being photoshopped to spread misinformation, images misportraying activists as ICE officers online, and the reframing of images to discredit immigrant rights organizations.

Visual news

In an increasingly image-based society, online visual content has become a powerful tool for political communication, including computational propaganda.¹ Visuals are elemental in shaping public opinion (Grabe, 2009), and an increasing amount of the US population, especially young people, rely on text-constrained, image-heavy social media to consume news. The Reuters Institute Digital News Report found in 2017 that 33% of adults aged 18–24 rely on digital and social media for their main source of news—the most preferred compared to all other sources of news consumption (Newman et al., 2017). Yet, the rise of the image as a vehicle for news dissemination is not a new phenomenon, nor surprising given our biological inclinations towards visual information consumption. Visual experience is the most dominant mode of learning (Grabe, 2009) and has prevailed as a primary channel of political communication since the inception of the television and network news coverage. Sight is the most sophisticated of all senses; our visual perception processes incoming data more quickly and efficiently than any other sense (Grabe, 2009). The human brain processes images so efficiently that basic recognition and emotional response can occur before we are consciously aware, which entails reaction and judgement based on emotions before mediated thought (Grabe, 2009). The resounding emotional impact, given the human capacity for image processing, has resulted in moving pieces of visual content ranging from Pulitzer prize-winning photojournalism to anti-immigration propaganda throughout US history (Hernandez, 2016).

1. Computational propaganda is a term and phenomenon that encompasses recent digital misinformation and manipulation efforts. It is best defined as the use of algorithms, automation, and human curation to purposefully distribute misleading information over social media networks (Woolley & Howard, 2016).

Networked activism on social media platforms

A well-documented process essential to social movement development, known as framing, has carried over from physical to online public spheres through social media channels (Meraz & Papacharissi, 2013). A popular definition of framing comes from Entman, who specified that “to frame is to select some aspects of a perceived reality and make them more salient in a communicating text” (Entman, 1993, p. 52). Framing has been repeatedly applied to social network analyses³ (SNA) of online social movement organizing, using affordances specific to social media platforms—such as mentions, retweets, and hashtags—as a methodological technique to make sense of traditional political organizational processes like information gatekeeping online (Stewart et al., 2017; Meraz & Papacharissi, 2013).

The Arab Spring uprisings in 2011 were heavily studied political protests by social network researchers as they were the first to be largely facilitated by social media. Live tweeting and live video feeds played an especially important role for activists, enabling them to share their experiences with a larger global audience, regardless of traditional media coverage. This produced eye-witness visual content that was disseminated online through retweeting, and narrative frames for the movement were essentially crowdsourced into prominence on a global stage (Starbird & Palen, 2012). Meraz & Papacharissi (2016) used previous seminal work on social movement framing as a lens to explain evolving patterns of communication on Twitter surrounding the Egyptian uprisings. Findings underscored the significant role of “ordinary” users who “both rose to prominence and elevated others to elite status through networked gatekeeping actions” (Meraz & Papacharissi, 2016).

Similarly, Kow et al. (2016) studied a more technologically complex situation during the Umbrella Movement, a series of student protests that took place in Hong Kong in 2014. Across platforms, Kow et al. (2016) identify four primary uses of social media

for social movements: (1) information seeking and dissemination, (2) frame articulation, (3) expression of solidarity, and 4) development of counter-narratives. This final purpose—development of counter-narratives, involves the expression of alternative perspectives, often in conflict with official reports (Al-Ani et al., 2012; Monroy-Hernández et al., 2013). Counter-narratives “rebut, undermine, or neutralize” a group’s framing (Benford & Snow, 2000), a notion that is especially important for our analysis as we examine the network structure and motivations behind manipulated media. Networks around alternative media sites, which connect domains with the same users, have also been examined to triangulate the actors creating and propagating alternative narratives (Starbird, 2017) and have been used to track the patterns of online rumor propagation, as a step towards building misinformation detection tools (Maddock et al., 2015).

This study extends previous research on online activism and rumor propagation by focusing on visual elements that are used to represent narratives and their subsequent counter-narratives. The visual content shared by social rights activists inextricably ties them to the current political social contexts and movements that they represent, making visual content both useful for framing and, simultaneously, constantly subject to reframing and recontextualization.

3. Social network analysis has roots in sociology and has more recently come into popularity due to advances in computational power and availability of social network big data. At its core, SNA aims to study the patterns in relationships that connect social actors through the use of network graphs and graph theory.

methods

Data collection via the Twitter API

Data was collected from the Twitter streaming application programming interface (API) from June 28 to July 5, 2018. This time period was chosen due to the “End Family Separation” marches planned to occur across the United States on Saturday, June 30. The zero tolerance and family separation immigration policies⁴ were already being largely debated after images and audio of detained young children from inside one of the children’s shelters had circulated earlier that day. Significantly, there was data loss caused by Internet disconnection early Sunday, July 1, the day following the marches, until the next day, July 2.

The authors chose popular hashtag keywords to follow based on their experience working closely with online immigration rights communities. Three neutral hashtags were used as keywords: “#immigration,” “#border,” and “#HappeningNow.” Five pro-immigration hashtags were circulating: “#EndFamilyDetention,” “#WhereAreTheChildren,” “#AbolishICE,” “#FamiliesBelongTogether,” and “#OccupyICE.” Additionally, five anti-immigration hashtags were gathered: “#IllegalImmigrants,” “#SendThemBack,” “#BuildTheWall,” “#ImmigrationReform,” and “#SecureTheBorder.” Tweets were processed and stored in an SQL database.

DESCRIPTIVE STATISTICS

In total there are 943,300 tweets and 53 pieces of metadata for each tweet, including tweet ID numbers, user ID numbers, ID numbers for media that is native, retweeted, or quoted, tweet ID of original content if it was retweeted or quoted, hashtags, URLs, etc. Of the nearly 1 million tweets, 759,397 were retweets (68%), 111,800 were original (10%), and 243,153 were quotes of another tweet (22%)—some tweets were retweets of a quote and overlapped in two of these categories, explaining the combined total of these metrics (1,114,350) being greater than total tweets in the dataset.

Two outliers exist in the dataset—one user tweeted 2,305 times, all of which were retweets, and another tweeted 1,708 times with a username “OpenLetterBot” that was created to retweet posts of letters to congressmen and governors. Furthermore, 925,787 tweets were from unverified sources, and the remaining 17,513 (2%) were from verified sources. Half of the verified tweets were retweets (52%), 9% were quotes, and 39% were original. Comparatively, for unverified, 81% were retweets, 8% quotes, and 11% original. Overall, 22,898 tweets had at least one native piece of visual content (photo, movie, GIF), and the number of tweets with more than one piece of media decreased respectively until the maximum fourth image allowed per tweet. The total visual content count was comprised of 21,036 photos, 1,600 videos, and 262 GIFS. The dataset also contained 80,030 tweets that were retweets of visual content and 57,715 quotes of visual content.

In order to capture collective narratives around visual content posted in the dataset, we conducted a process of iterative quantitative exploration of structural data and qualitative inquiry into products of the preceding quantitative analysis. This process allowed us to narrow in on specific prevalent visual content frames circulating during data collection and identify discursive and organizational patterns common for users when interacting and framing visual content.

Research Question 1: constructing the networks

To investigate our first research question and construct a larger picture of the visual content networks, network graphs were generated by querying the database for quoted and retweeted unique media IDs associated with the original pieces of visual content. The distinction between retweets and quotes was made to distinguish when a narrative about the shared content was being pushed, or when a new (potentially counter-) narrative was being created. Then, images (photos, videos, and GIFs) shared by the same users were connected and visualized utilizing Gephi, an open-source network

4. The Trump administration mandated a “zero-tolerance” policy by executive order on June 20, which calls for the prosecution of all individuals who illegally enter the United States. This policy has the effect of separating parents from their children when they enter the country together, because parents are referred for prosecution and the children are placed in the custody of a sponsor, such as a relative or foster home, or held in a shelter (Politifact).

analysis and visualization software. The structure of the network was built using the Force Atlas 2 algorithm (Jacomy et al., 2010), a force-directed algorithm that comes standard with the software. Each node represents the original tweet of visual content and is connected by edges that represent how many users retweeted or quoted both images—the thickness of the edges indicates the amount of shared users that engaged with both images, which attracted or repelled the nodes.

The Gephi modularity algorithm (Blondel et al., 2008) was used to detect communities of images within the network for further analysis. Generally speaking, this algorithm is defined as the difference between the number of edges within communities and the expected number of such edges, and is often implemented in SNA to identify communities in large and complex networks. Past studies have identified hashtags within clusters as important mechanisms in framing for social movements online (Stewart et al., 2017), and accordingly the most popular hashtags from these clusters were computed and used to give an overview of the most prevalent original narratives associated with each image cluster.

Research Question 2: evolving narratives

Qualitative analysis identified key original narrative stances, while content analysis, using an existing framework, surfaced intent around each original image or video in the graph. This analysis further corroborated what the hashtags from Question 1 thematically suggested about prevalent narratives in each cluster. All coding was done by the first author, who was immersed in this work. Content analysis of the original content drew from four identified uses of social media for social movements as a framework developed by Kow et. al (2016) to understand the intent behind each prevalent narrative. The codes were:

- **Information seeking and dissemination:** this included images of “protester rights” flyers and videos posted by the ACLU.

- **Frame articulation:** this included tweets that were adding new angles to existing immigration rights frames, such as the “no one is illegal on stolen land” frame.
- **Expression of solidarity:** this included tweets that were mainly composed of hashtags already established for the movement.
- **Development of counter-narratives:** this included visual content opposing an already existing frame, such as an image of a house with open front doors mocking the stance of open borders.

Then, subsequent discussion around a subset of the top five images from each of the four clusters were analyzed by identifying the quoted tweets with the most volume, or highest number of retweets (Arif et al., 2016). The high-volume quotes were determined by either being one of the top ten most retweeted quotes for each image or until a quote had at least two retweets (in the cases where there were less than ten quote tweets with high volumes). These popular narratives were compared to the original image narrative to better understand the evolution from narrative to counter-narrative.

Research Question 3: mapping discursive patterns

Drawing on Maddock et al.’s (2015) work on rumor propagation, qualitative and visual analysis were applied to understand the origin and evolution of prominent counter-narratives identified in Question 2. The retweets of these narratives and prominent counter-narratives were plotted and compared over time; this illustrated patterns in how new context for visual content rose to prominence. The reach, or follower size of each user (Arif et al., 2016), was taken into consideration, and other visual signatures between narratives and counter-narratives were visualized. While examining the dynamic structure of narratives and their respective counter-narratives, we took specific notice of the reach and volume previously identified as important markers in the spread of a counter-narrative (Arif et al., 2016).

quantitative results

Data collection via qualitative interviews

Anecdotal evidence was collected through interviews with five separate members of the online activism community that regularly rely on social media to defend immigration rights. Interviewees were all members of WITNESS's network of human rights organizations. Participants were based in organizations across the country, and because of this geographical diversity, all interviews were conducted remotely: two over the phone and three by video chat. Interviews were typically an hour and included a set of questions that covered strategies around social media platform use, experiences with online attacks (including media manipulation), and preventative measures used for future attacks. We conducted a grounded theory-based iterative and inductive analysis of the interview notes using coding and memoing (Strauss & Corbin, 1998) to sift out themes across interview data. The codes were reduced using axial and selective coding to the themes presented in the findings of this paper.

Research Question 1: exploring image network structures

The first graphs constructed were networks of visual content with high levels of engagement: the first of tweets containing visual content that were retweeted at least once (Figure 1, page 9) and the second of tweets containing visual content that were quoted at least 20 times in the dataset (Figure 2, page 10). As in past research, retweeting signifies support and contribution to a growing narrative, whereas the process of quoting visual content entails additional context that shapes and tries to further propagate support or articulate a counter-frame for the visual content (Meraz & Papacharissi, 2013).

Construction of the networks of shared users between this content, distinguishing between shared quoters or shared retweeters, visualizes patterns in creation and support of narratives around visual content. Communities within each network were computed and color-coded for each graph to distinguish visual content clusters.

The first shared users graph is shared retweets of images (Figure 1, page 9) and shows two distinct super clusters of visual content: a single cluster to the right (black) and another on the left (multi-colored). The top ten hashtags of the right cluster and left multi-cluster were examined and (Table 1, page 9) revealed a homogenous, clear left-leaning/right-leaning political divide in, and siloed consumption of, visual content. This siloed or "echo chamber" behavior is a common characteristic noted by past network analysis of the spread of contentious content online (Starbird, 2017; Stewart et al., 2017; Arif et al., 2015). Comparatively, the second shared users graph of shared quotes (Figure 2, page 10) illustrates more intertwined clusters of images, indicating more connected images and shared users. Hashtags for clusters were more varied (Table 2, page 10), especially the green cluster, which notably has the least number of images but the highest number of unique hashtags. This implicates more diverse users interacting with visual content outside of their cluster boundaries, contrary to the more isolated spread of visual content in the retweet graph (Figure 1, page 9).

Further inspection of hashtag use in the quoting graph showed that co-opting was also prevalent. Co-opting is commonly used to disrupt frame development on social media (Stewart et al., 2017). Notably, this phenomenon is not one-sided; in several instances the MAGA hashtag was appropriated by pro-immigration attendees of the “End Family Separation” marches. Two examples from this cluster are seen in Figure 3 (page 11).

Node size is determined by the degree—or how many edges connect it to another node—of the visual content. Accordingly, degree is a robust measurement of image popularity and telling of the main narratives in each cluster, since the more users that quoted that photo or video and another image in the graph, the larger the degree. Large nodes can be seen more prominently at

the center of a cluster or acting as a bridge between two clusters. The evolving narratives around these popular images will be discussed further in proceeding sections, but these graphs alone indicate notable images acting as bridges, or a piece of content that connect two otherwise disparate clusters of images. For the first graph (Figure 1), these shared retweeters are much less common, creating the siloed clusters, as made visible by the graph, in which the same narratives are perpetuated through a refraining process without interruption from counter-narratives (Meraz & Papacharissi, 2013). In the case of the shared quoting users (Figure 2, page 10), in which text can be added to visual content and shared with new opinion or context, the bridges represent a negotiation of framing around images. The shared quoter graph (Figure 2, page 10) is a visual

Figure 1. Network of visual content with shared retweeting users from immigration rights protests.

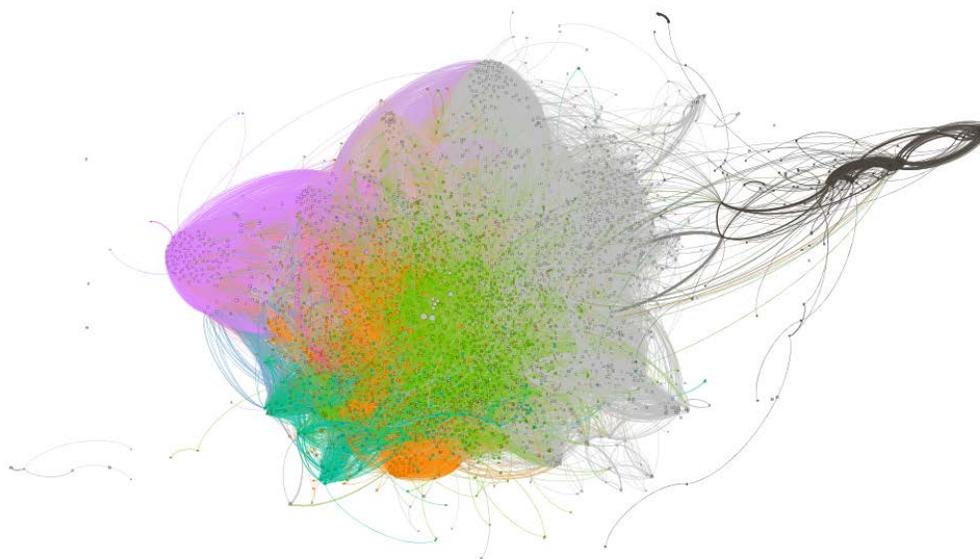


Table 1. Most used hashtags in clusters of retweeted visual content.

Cluster	Hashtags
	BuildTheWall (63), MAGA (26), FamiliesBelongTogether (16), KAG (9), AbolishICE (9), buildthewall (7), Americafirst (7), ICE (6), OccupyICE (5), Trump2020 (4)
Multi-colored	FamiliesBelongTogether (2459), AbolishICE (663), FamiliesBelongTogetherMarch (450), familiesbelongtogether (252), EndFamilyDetention (192), abolishICE (127), FreeOurFuture (73), OccupyICE (73), KeepFamiliesTogether (67), abolishice (62)

representation of a constant negotiation and framing around images where the retweet graph represents the images and their respective frames surfacing through repetition and shared ideology with little interruption.

Research Question 2: prominent narratives and counter-narratives

The metadata (tweet text, user, description, URLs, etc.) surrounding each of the *original* 120 images in the shared-quoter network was collected and analyzed, and subsequently the top five *quotes* (i.e., most popular

narratives) of highly-circulated images were analyzed. This process of comparing original narratives to the quoting that developed around these images allowed for a deeper understanding of how narratives evolved over time around each popular image. The common narratives and counter-narratives that surfaced in this analysis are detailed below. In each cluster, common frames around immigration were found, as well as the applied framework from Kow et al. (2016) to help identify distinct motivations in posting the popular images. Framing of each original piece of visual content aligned with popular hashtags were discovered in the network

Figure 2. Network of visual content with shared quoting users from immigration rights protests.

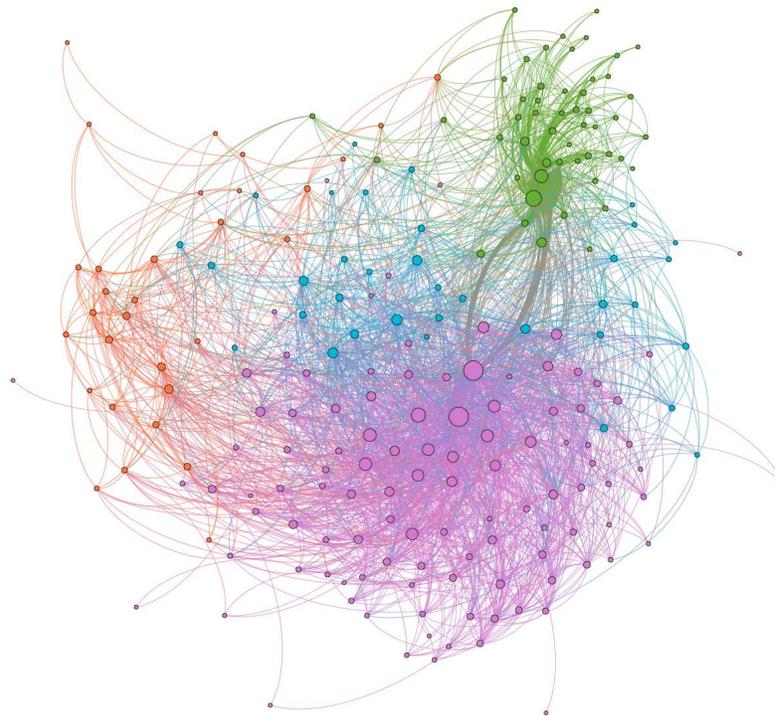


Table 2. Most used hashtags in clusters of retweeted visual content.

Cluster	Hashtags
Green	MAGA (8), BuildTheWall (7), AbolishICE (7), FamiliesBelongTogether (4), KAG (2), Illegal (2)
Purple	FamiliesBelongTogether (32), AbolishICE (12), EndFamilyDetention (4), WomenDisobey (2), womendisobey (2), FreeOurFuture (2), FamiliesBelongTogetherMarch (2)
Blue	FamiliesBelongTogether (24), FamiliesBelongTogetherMarch (5), EndFamilyDetention (3), AmericaWelcomes (2)
Orange	AbolishICE (13), FamiliesBelongTogether (4), abolishICE (3), OccpyICELA (2), OccupyICE (2)

graph section, both in content and homogeneity. Conversely, though, new narratives generated through popular quoting did not always align with the intent of the original content and often ran contrary to the initial framing.

CHARACTERIZATION OF THE CLUSTERS

A complex right-leaning fringe (Figure 2 green cluster).

Popular visual content in this cluster came from the account @AlwaysActions, which consists of mugshots from “illegal aliens” that have committed horrendous crimes. Accordingly, common narratives that surfaced through content analysis, both of the original text and the subsequent quotes, revolved around safety, truth, and the fear of open borders. The original frames of these images remained generally uncontested in preceding quotes of the image, 95% of high-volume preceding quotes were coded as solidarity with the original narrative. Only one quote from these images was a counter-frame: a video of police lined up for a march with the user pointing out that the police should instead be protecting the city from crime.

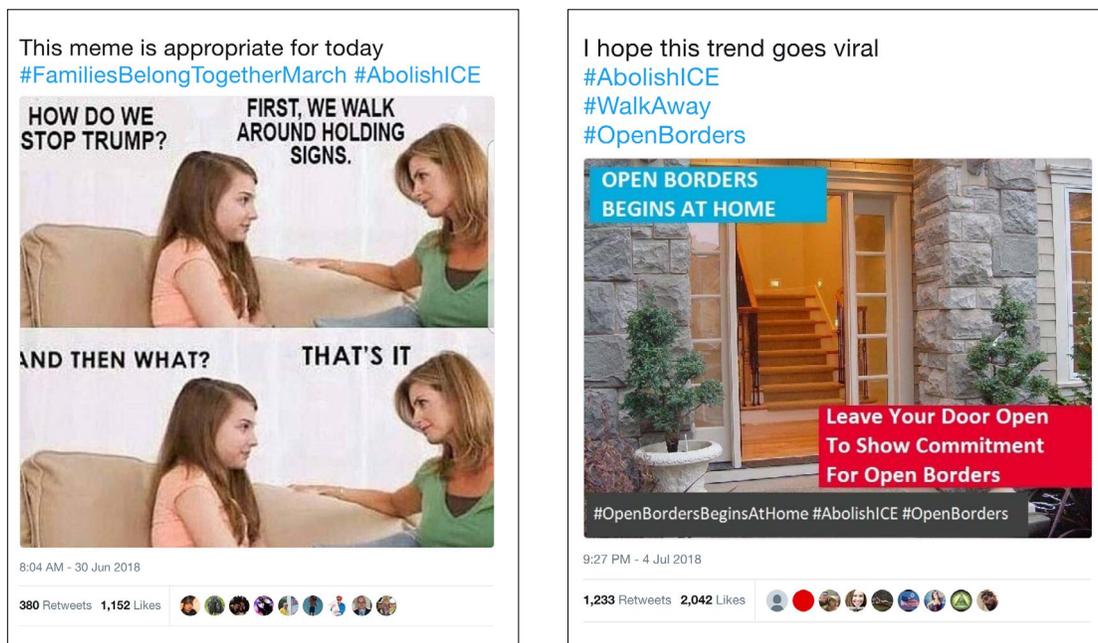
From these instances, the cluster of original frames appears generally homogenous in narrative and user

ideology, but of the top five most popular images, only two were aligned with the right-leaning narratives mentioned above. Other popular tweets of visual content had originated from left-leaning activists and were co-opted by this cluster, to the extent that they were grouped into the same (green) community. For example, the bridge between this cluster and the others is a photo of protesters from an account with the username NYCDSA, the New York Democratic Socialist Association. This was the most quoted image in the cluster and the entire dataset—alone this image was quoted 4,904 times. Of the top ten most popular quotes of this image, all were counter-narratives to the original post. Popular quotes of the image ranged from civil discourse on socialist positions to claims that the photo depicted a modern Maoist uprising. Notably, all high-volume quotes reframing this image were verified, whereas all other sets of high-volume quotes of images in the cluster were at least 75% unverified.

Closely tied protest clusters (Figure 2 blue and purple cluster)

These two clusters represent similar communities of images and narratives in an already well-articulated movement. Frames surfaced around humane treatment of immigrants, dissent against the current

Figure 3. Two examples of hashtags co-opting the End Family Separation movement.



administration, mobilizing marchers for protesting, and democracy. These clusters had premeditated frames for the End Family Separation marches, and thus solidarity was the most common type of engagement (75%), mostly reiteration of existing march hashtags, followed by information dissemination (15%), and frame articulation (10%). These clusters were largely homogenous, and high-volume quotes of the original visual content was engaged with to show solidarity or disseminate information for protesters. Only about 10% of the high-volume quotes of these images were counter-narratives, explaining the homogeneity. In other words, users that were quoting these photos and videos were showing agreement and support of what the visual content represented.

An organized Occupy movement (Figure 2 orange cluster)

Of all clusters, the orange used social media engagement to disseminate information (23%) for organizational purposes and was most characterized by Occupy and Abolish ICE narratives. As these were more information-based and generally not used to promote an agenda, these tweets had the least amount of subsequent quoting, and instead demonstrated a higher percentage of retweets. As with the blue and purple clusters, solidarity was most, but the second most common engagement was Occupy/Abolish members using photos and videos to communicate where and when sit-ins were happening and information on arrests of protesters that were taking place. The Occupy cluster, although also representing far-left ideology, seemed less vulnerable to online attacks because they used social media more than the other clusters to spread information rather than share a political stance.

QUALITIES OF COUNTER-NARRATIVES

The interdependence of image narrative and counter-narrative

For every narrative that was coded in the original set of visual content, there was a counter narrative, such as Democracy/Snowflakes protesting, Immigration/Open Borders, and Legal/Illegal immigrants. For example, in response to the NYCDSA's image of protesters holding "Abolish ICE" signs during the End Family Separation march, one user quoted the photo and responded with a warning of communism:

(June 30, 6:54pm): We are witnessing a Maoist uprising in realtime [link]

The "Impending Communist" description was a common theme in many counter-narratives and easily played into the "Snowflake" narrative that discredited protesters and young marchers expressing their socialist-leaning views. This communist reframing of a socialist protest image was retweeted 257 times in this dataset alone.

Stoking fear with images

The long history of image campaigns targeting social groups has carried over into the digital sphere. Counter-narratives, and the photos and videos that are used to contextualize them, play up stereotypes that have recurred throughout the history of anti-immigrant propaganda in the United States. As mentioned above, central to the green cluster was an account, @AlwaysActions, that claimed to share "breaking news and developing stories" in their description. This was a highly active account, with 106 tweets in the week span of collection alone, and appeared to target several social groups, including Muslims, African Americans, and immigrants, through tweets and retweeting of conspiracy theories and out-of-context images. In our dataset, this account shared four images, each depicting the mugshot of a Latino man, and detailed that he was seeking asylum in the United States from a horrendous crime that he had committed as an "illegal immigrant." For example, the following tweet was accompanied by a photo of an immigrant:

**(July 3, 1:54pm) #Mexican Killer Illegal alien
#Sneaks into the United States 7 times
#MAGA #BuildTheWall [link].**

While this sounds like a fringe incident, these images were quoted and retweeted 632 times just within the dataset and can be assumed to be disseminated further given the 47,594 followers the accounts had at the time of collection. The ensuing comments of these tweets were comprised of users genuinely scared for their safety and calling for a wall to be built at the Mexican border that could prevent immigrants like the one pictured from entering the United States. These posts also prompted more gruesome, violent comments that called for physical punishment of the immigrant and those like him.

Harassment

Racial slurs, sexism, and calls for violent actions against protesters and immigrants were also noted in the counter-narratives. Several attacks on protesters were geared toward sexual harassment of women and used photos to invite further commenting and misogyny. One user posted a video of lined-up police preparing for Occupy protesters in full armor, which baited conversation of violence against the peaceful protesting. Others used derogatory terms for Latino immigrants, one of which was retweeted over 100 times in the dataset alone. While it is widely known that these unfortunate and demeaning attitudes exist online, the normalized harassment of immigrants and activists plays an important role in deterring participation in online advocacy spaces, as activists recounted in interviews covered in the anecdotal section.

Research Question 3: development of misinformation around images

This section analyzes who spread the narratives uncovered in the preceding section and how successfully they spread. To understand the underlying processes needed to develop misinformation about visual content requires a more granular examination of discursive patterns for individual visual content, and, more specifically, how narratives and subsequent counter-narratives about immigrants are co-created

and developed over time. This section focuses on counter-narratives identified by previous analysis that rose to prominence over the original tweet and was quantitatively determined by comparing volumes of the original tweets and the quotes of the tweets. In the case of the green cluster, there were five individual instances where a counter-narrative rose to prominence over the original tweet while co-opting the visual content (Table 3). Notably, only one of these prominent counter-narratives was articulated around a video.

After examining the reach, or follower size, of each user that generated original content and counter-narratives, it surfaced that 100% of co-opters had a broader reach than the poster of the original tweet. This is in line with Arif et al.'s (2016) research that found reach and volume to be just as important to rumor spread as the actual content. Interestingly, while verified sources were twice as likely to share native visual content than unverified sources, high-volume counter-narratives of the same content were over twice as likely (70%) to be created by unverified sources than verified.

There was a mix of patterns identified for counter-narratives rising to prominence. In the instance of the New York DSA image—the major bridge that connected the right-leaning and left-leaning clusters—co-opting the counter narrative consistently overpowered the message

Table 3. Immigration rights narratives and the counter-narratives that rose to prominence over them.

Narrative	Counter-narrative	Narrative Audience	Counter-Narrative Audience	CN Prominence Increase	Media Type
Dissent	Communism	17,676	27,282	10%	Photo
American values	Working Class	2,370	23,258	40%	Photo
Freedom	Illegals	5,056	157,985	75%	Photo
Humane Treatment	Criminals	8,407	18,734	5%	Photo
Dissent	Authority	2,309	24,627	48%	Video



of the original photo (Figure 4) within hours. The instance in Figure 4 shows one of three counter-narratives of the original DSA photo that rose to prominence. This one specifically targeted Democratic primary victor Alexandria Ocasio-Cortez for being a dues-paying member of the group.

In several instances, there were prior attempts to offer a reframing that did not reach prominence, but largely adopted counter-narratives took hold once more prominent figures created the necessary counter-frame for the visual content. For example, an account with only 13 followers and no original posts quoted an image a young woman confronting the police at a protest and mentioned another user with a high audience reach (24,627 at the time) on Saturday at 16:00 (Figure 5, page 15). The prominent user proceeded to share the image

of the young woman, criticizing her protesting the police, which quickly overtook the popularity of the original images (Saturday at 17:00).

In all cases but one, a counter-narrative consistently drowned out a narrative. In the other instance, a dormant counter-narrative gained popularity after the original narrative had plateaued (Figure 6, page 15, Tuesday at 7:00) and eventually rose to prominence over the original framing. This could be explained by a second prominent figure retweeting the counter-narrative hours later. A user with close to 12,000 followers retweeted the counter-narrative around Tuesday at 13:00 followed by another with close to 72,000 followers, explaining the new spike in the graph and supporting this hypothesis (Figure 6, page 15).

Figure 4. Example of a growing counter-narrative rising to prominence over the original narrative.

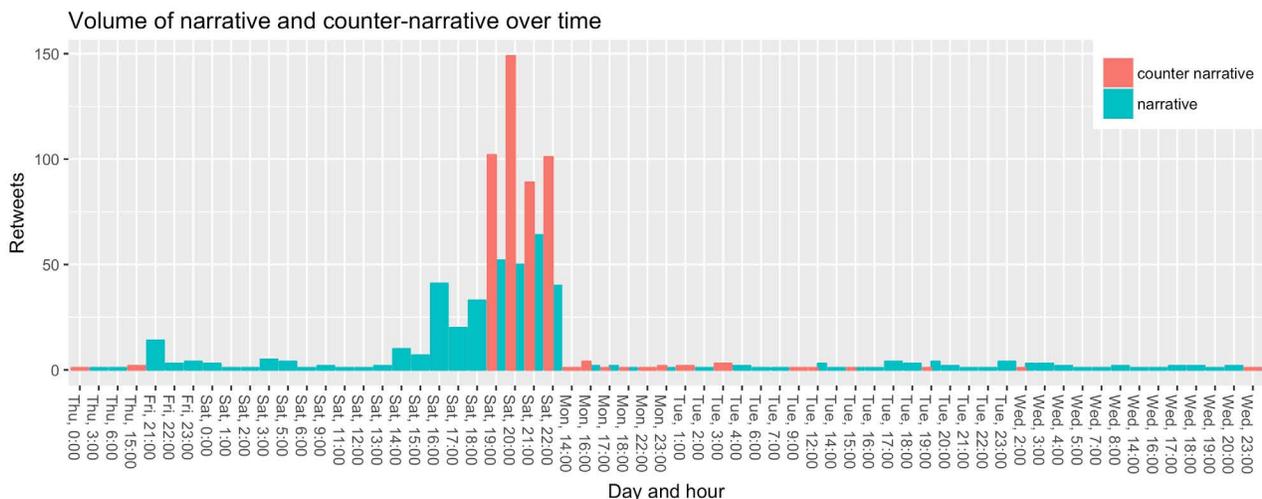


Figure 5. Example of an immediate drowning out of narrative by counter-narrative.

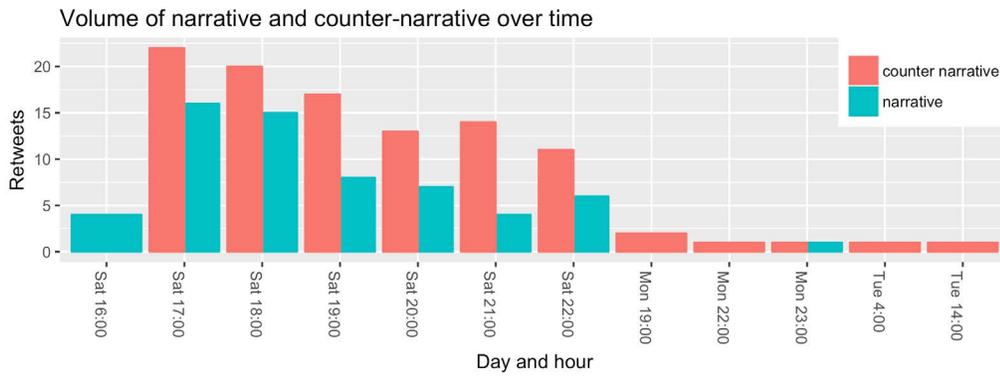
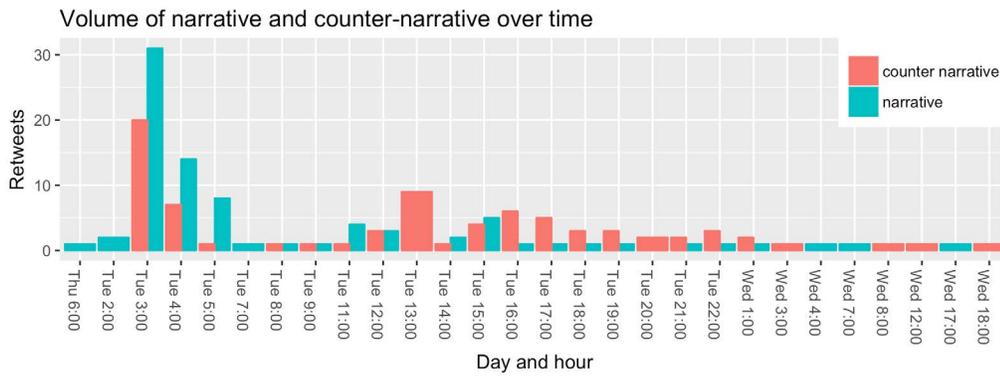


Figure 6. Example of a resurgent counter-narrative.



qualitative results

Interviews: experience of immigration rights activists online

The Twitter data poses limitations on the conclusions that can be made concerning the qualitative experiences of activists and immigrants using social media, so interviews were conducted to supplement the network analysis. These interviews added anecdotal data to enrich the stories told by the tweet data alone. Five diverse immigrant rights activists that were recognized for their personal and/or organizational online presence were interviewed about their advocacy experiences online. Invariably, each activist iterated the importance of social media for immigrant rights activism work. The ability to reach millions of people in and outside of immigrant audiences, engage in policy-changing conversations, and share crucial information on legal issues is invaluable for these communities. Social media benefits always outweigh the harm done by bad actors perpetrating harassment on platforms. That said, there are four key themes that surfaced amongst the interviewees that emphasize the use and threats of immigrant rights activism online: (1) affordances of the image; (2) digital to physical spaces; (3) absence of formal strategy, and (4) inclusive solutions.

AFFORDANCES OF THE IMAGE

Interviewers noted transitions in the last few years towards image-based campaigns, especially for raising awareness of lived immigrant experiences. On a daily basis, activists utilized stories told through visual content to raise bond for immigrants, organize call-ins to advocate for the termination of a deportation, and spread information through formal videos on DACA applications and legal advice. One activist's organization started engaging with their community through YouTube, since immigrants were already sharing their stories with each other through videos on the platform. Another interviewee was raising money to support bail for immigrants with unjust misdemeanors, which was difficult because they were seen as "criminals." Instead of using written words from people, they told their story through video and raised \$2,000,000—more than the expected \$70,000.

The transition to image-based campaigns, though, has invited new tactics for the targeting of immigrants and pro-immigration activists. Interviewers shared stories of how their organizations' digital "know your rights" flyers were photoshopped to spread contradicting information about ICE raids; how trolls created videos of pro-immigration activists, portraying them as ICE officers; and, how images were reframed to represent immigrant organizations as white supremacist supporters. Multiple interviewers said that perpetrators are tech-savvy and have brought their own cameras to live-stream immigrant event crashing, have used social media to track and disrupt activism events, and have created memes of minorities looting after a natural disaster.

DIGITAL TO PHYSICAL SPACES

The evidence of physical violence—incited by false narratives, malicious tactics, and mobilization on social media—carrying over into physical spaces is the most pressing concern for activists. Groups are being tracked on social media, and their information sessions and protests are being intruded on by bad actors who aim to harass immigrants. Activists are seeking solutions *around*, not *with*, platforms to protect their communities from physical repercussions of online behaviors. One interviewee said, "Monitoring hasn't really happened, especially stuff around the anti-immigrant stuff and blatant lies about immigration not being addressed. There's not much faith in corporations to provide this. The real question is: How can we provide that safety outside of these digital structures?"

Newly instated ICE policies and the larger political climate these were created in were, and continue to be, confusing for immigrant communities in the United States. This uncertainty is intentionally amplified online where images are strategically used to sow further confusion about the identity of ICE officers and deter immigrants from attending informational sessions. This physical carry-over from digital spaces is what some interviews cite as the most dangerous and characterizing feature of disinformation.

ABSENCE OF FORMAL STRATEGY

At the time of the interviews, there were no major attacks recounted by participants, but small-scale attacks were ubiquitous. In general, the strategy used most by the interviewed human rights organizations is to drown out misinformation surrounding their content, which was an effective strategy given the most common attacks were accounts with small followings. To this end, no pre-emptive formal strategies were in place for these human rights organizations if a more orchestrated attack occurred, although these types of attacks from prominent actors showed to be the most effective in spreading misinformation, according to our network analysis. While organizations iterated that they felt comfortable in their ad hoc strategies for managing attacks, they recounted that often the individuals from their communities are more frequently targeted after engaging with their pro-immigrant content, which fettered overall participation and self-advocacy. Overall, no organization voiced trust in platform accountability or intervention as an option for misinformation management.

INCLUSIVE SOLUTIONS

Although platform responses have begun to take shape, immigration activists believe these solutions aren't taking into account the most affected by malicious online practices. For example, one interviewee cited a recent Twitter account audit that not only terminated harmful accounts, but also a key organizer in the Black Lives Matter marches. Subsequently, the march was sent into disarray. This could have been avoided if major voices of social rights activist organizers had been present in the conversation. Unfortunately, although immigrant populations make use of social media, several interviewees noted that their audience are comprised mostly of white allies without the lived experiences of an immigrant, who are less subjected to harassment and intimidation while advocating for immigration rights online. If those most affected are reluctant to participate online to advocate for themselves, then allies that feel comfortable in these digital spaces will still run the conversations.

These interviews gave us a glimpse at the day-to-day hurdles of online immigration activists that the Twitter data alone cannot tell. Overall, activists voiced their daily processes in working around the systems in place since platforms aren't doing enough to guarantee the safety of their immigrant communities. The users have developed their own tactics to deter bad actors, such as creating background patterns on flyers to make photoshopping more difficult, that provide safety outside of the structures in place that aren't working. Disinformation, if not on a mass scale, isn't so obvious, especially to mediators of platforms that are controlling billions of users and their data at a time. But these interviews shed light on the tangible and daily realities overflowing from digital spaces that immigrant communities and their allies encounter.

conclusion

The assertion that the objective camera lens will invariably convey truth is an unreasonable expectation for the evolving image online. In the same way that content within a video or picture can mask or uncover realities, human fallibility, or worse, mal-intent, it can also shape the narrative around visual content as it circulates on social media. Those most vulnerable to misuse of visual content are already marginalized voices, which can become subjected to harmful counter-narratives playing on historical propaganda techniques like stereotypes, and/or drowned out by more prominent figures on social media extolling these harmful narratives. Past research has suggested that opposing ideologies are generally siloed throughout the process of frame and counter-frame development when deliberating over text alone. The analysis of discursive patterns around visual content has shown the framing process to be more complex and encouraging of multiple perspectives to participate in the framing of images, although qualitative work has shown these negotiations to vary in civility and effectiveness. Human rights activists have an equally complex relationship with social media—the ability to organize audiences that otherwise would be unreachable is indispensable, and rallying around images and symbols of movement creates a larger sense of purpose and identity that has become vital. Yet, there is the inevitable cost of bad actors using the same affordances of social media.

This study illuminates two important facets of computational propaganda: the life an online image takes on as it is transmuted into misinformation and the subsequent tangible effects visual misinformation has on immigration rights efforts. In various forms, the analyses in this study suggest the power of prominent figures in shaping how we perceive the images presented to us online. Given the rapid exposure and quick emotional response to visual content, the promise of a well-followed source adds an immediate layer of authentication to content and context of an image—whether or not that account is verified as a credible source. With this in mind, harmful framing and misinformation of visual content is easily perpetuated by prominent online figures that create powerful counter-narratives playing on stereotypes and harassment akin to anti-immigrant propaganda witnessed throughout history.

references

- Al-Ani, B., Mark, G., Chung, J., & Jones, J. (2012, February). The Egyptian blogosphere: a counter-narrative of the revolution. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* (pp. 17-26). ACM.
- Andén-Papadopoulos, K., & Pantti, M. (Eds.). (2011). *Amateur images and global news*. Intellect Books.
- Arif, A., Shanahan, K., Chou, F., Dosouto, Y., Starbird, K., & Spiro, E.S. (2016). *How Information Snowballs: Exploring the Role of Exposure in Online Rumor Propagation*. CSCW.
- Bakdash, J., Sample, C., Rankin, M., Kantarcioglu, M., Holmes, J., Zaroukian, E. & Szymanski, B.K. (2018). The Future of Deception: Machine-Generated and Manipulated Images, Video, and Audio? *Proc of the IEEE International Workshop on Social Sensing (SocialSens)*
- Benford, R. D., & Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual review of sociology*, 26(1), 611-639.
- Blondel, V. D., Guillaume, J. L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008.
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of communication*, 43(4), 51-58.
- Grabe, M. E., & Bucy, E. P. (2009). *Image bite politics: News and the visual framing of elections*. Oxford University Press.
- Gregory, S. (2018, July). WITNESS leads convening on a proactive solutions to mail-users of deepfakes and other AI-generated synthetic media. Retrieved from <https://witness.org/witness-leads-convening-on-proactive-solutions-to-mal-uses-of-deepfakes-and-other-ai-generated-synthetic-media/>
- Hernandez, T. (2016). *A brief History of Anti-Immigrant Propaganda*. Retrieved from <https://immigrantarchiveproject.org/brief-history-anti-immigrant-propaganda/>
- Jacomy, M., Venturini, T., Heymann S., Bastian, M. (June 10, 2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software.
- Kow, Y. M., Kou, Y., Semaan, B., & Cheng, W. (2016, May). Mediating the undercurrents: Using social media to sustain a social movement. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3883-3894). ACM.
- Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. New York: Data & Society Research Institute.
- Maddock, J., Starbird, K., Al-Hassani, H. J., Sandoval, D. E., Orand, M., & Mason, R. M. (2015, February). Characterizing online rumoring behavior using multi-dimensional signatures. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing* (pp. 228-241). ACM.
- Meraz, S., & Papacharissi, Z. (2013). Networked gatekeeping and networked framing on #Egypt. *The international journal of press/politics*, 18(2), 138-166.
- Mitchell, A., & Page, D. (2014, March 26). *State of the News Media*. Retrieved from <http://www.pewresearch.org/wp-content/uploads/sites/8/2017/05/state-of-the-news-media-report-2014-final.pdf>
- Monroy-Hernández, A., Kiciman, E., De Choudhury, M., & Counts, S. (2013, February). The new war correspondents: The rise of civic media curation in urban warfare. In *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 1443-1452). ACM.
- Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D.A.L., & Nielsen, R.K. (2017). *Reuters Institute Digital News Report 2017*. Retrieved from https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital%20News%20Report%202017%20web_0.pdf
- Paul, C. & Matthews, M. (2016). "The Russian 'Firehose of Falsehood' Propaganda Model: Why It Might Work and Options to Counter It." *International Security and Defense Policy Center*
- Seo, H. (2014). Visual propaganda in the age of social media: An empirical analysis of Twitter images during the 2012 Israeli-Hamas conflict. *Visual Communication Quarterly*, 21(3), 150-161.
- Smith, A., & Anderson, M. (2018, March 1). *Social Media Use in 2018*. Retrieved from <http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/>
- Starbird, K. (2017, May). Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter. In *ICWSM* (pp. 230-239).
- Starbird, K., & Palen, L. (2012, February). (How) will the revolution be retweeted?: information diffusion and the 2011 Egyptian uprising. In *Proceedings of the acm 2012 conference on computer supported cooperative work* (pp. 7-16). ACM.
- Stewart, L. G., Arif, A., Nied, A. C., Spiro, E. S., & Starbird, K. (2017). Drawing the Lines of Contention: Networked Frame Contests Within #BlackLivesMatter Discourse. *PACMHCI*, 1(CSCW), 96-1.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Procedures and techniques for developing grounded theory*.
- Wardle, C. (2017, February 16). *Fake news. It's complicated*. Retrieved from <https://firstdraftnews.org/fake-news-complicated/>
- World Economic Forum. (2017, January 11). *The Global Risks Report 12th Ed*. Retrieved from <https://www.weforum.org/reports/the-global-risks-report-2017>
- Woolley, S. C., & Howard, P. N. (2016). Automation, algorithms, and politics| Political communication, computational propaganda, and autonomous agents— Introduction. *International Journal of Communication*, 10, 9.

INSTITUTE FOR THE FUTURE

Institute for the Future is the world's leading futures thinking organization. For over 50 years, businesses, governments, and social impact organizations have depended upon IFTF global forecasts, custom research, and foresight training to navigate complex change and develop world-ready strategies. IFTF methodologies and toolsets yield coherent views of transformative possibilities across all sectors that together support a more sustainable future. Institute for the Future is a registered 501(c)(3) nonprofit organization based in Palo Alto, California.

DIGITAL INTELLIGENCE LAB

The **Digital Intelligence Lab** at Institute for the Future is a social scientific research entity conducting work on the most pressing issues at the intersection of technology and society. We examine how new technologies and media can be used to both benefit and challenge democratic communication.



INSTITUTE FOR THE FUTURE
201 Hamilton Avenue
Palo Alto, CA 94301
www.iff.org