ADL
Anti-Defamation League®

# Responding to Cyberhate

## Progress and Trends

March 2016

# Table of Contents

# INTRODUCTION

The Internet is the largest marketplace of ideas the world has known. It enables communications, education, entertainment and commerce on an incredible scale. The Internet has helped to empower the powerless, reunite the separated, connect the isolated and provide new lifelines for the disabled. By facilitating communication around the globe, the Internet has been a transformative tool for information-sharing, education, human interaction and social change. All of us treasure the freedom of expression that lies at its very core.

Unfortunately, while the Internet's capacity to improve the world is boundless, it is also used by some to transmit anti-Semitism and other forms of hate and prejudice, including anti-Muslim bigotry, racism, homophobia, misogyny, and xenophobia. The Anti-Defamation League (ADL) has been addressing the scourge of online anti-Semitism since pre-Internet days, when dial-up bulletin boards were a prominent communications tool. As the Internet emerged for personal use in the 1990's, ADL was there to monitor, report and propose approaches to fight online hate. Today, ADL is known as one of the preeminent NGO's addressing online anti-Semitism and all forms of online hate.

Cyberhate, defined as "the use of any electronic technology to spread bigoted, discriminatory, terrorist and extremist information," manifests itself on websites and blogs, as well as in chat rooms, social media, comment sections and gaming. In short, hate is present in many forms on the Internet, creating a hostile environment and reducing equal access to its benefits for those targeted by hatred and intimidation.

In an ideal world, people would not choose to communicate hate. But in the real world they do, all too often. And hate expressed online can lead to real-world violence, nearby or far away. Cyberhate poses additional challenges, because everyone can be a publisher on the Internet. Hateful content can spread around the globe literally in seconds, and it often goes unchallenged. So it is necessary to find effective ways to confront online hate, to educate about its dangers, to encourage individuals and communities to speak out when they see it, and to find and create tools and means to deter it and to mitigate its negative impact. In doing so, it is also important to keep in mind the need to find the right balance, addressing cyberhate while still respecting free expression and not inhibiting legitimate debate.

The unique challenge of hate speech online is unfortunately not the only challenge we face today. Extremists and terrorists have become much more sophisticated in their use of social media. This growing threat has been particularly evident with a rise in "self-radicalization," encouraged and abetted by terrorist groups. Terrorist exploitation of the Internet is an order of magnitude different from hate speech online, and new strategies may be necessary to respond to it.

# CHARTING PROGRESS

The following charts have been created to illustrate changes in the predominant Internet environment over just the past two years. In snapshot form, they show where we are today, revealing changes in both how cyberhate manifests itself on the Internet and how the industry has become more serious and more sophisticated in dealing with the problem. They document progress – often incremental, but in its totality significant, impressive and important – mostly the result of industry-sponsored initiatives.

Unfortunately, spewing anti-Semitism and hate online is much easier than finding effective ways to respond to it. We have come to understand that as long as hate exists in the real world, that hate will be reflected in the virtual world as well. What happens on the Internet is a reflection of society, and not the other way around. Consequently, as long as technology keeps evolving, and bias, racism and anti-Semitism persist, the haters will likely find ways to exploit the new services and new platforms to spew their corrosive message. We need to be just as creative, and just as determined, to counter them.

## Highlights of changes in various platforms when it comes to dealing with cyberhate

| PLATFORMS | 2013 | 2016 |
|---|---|---|
| Websites | Limited existence and enforcement of hosting company rules | Mixed picture. Many companies with appropriate terms of service are responsive. Companies with lax terms of service are less responsive. Potential impact of proposed new Federal Communications Commission (FCC) regulations considering U.S.-based hosts as "common carriers" is still to be determined. |
| Comments/Reviews | Sporadic enforcement of hate speech in review and comment sections of websites. Terms of Service not always clear or easily found. | Word-sifting software coming into increasingly frequent use. Websites far more responsive to complaints about issues in reviews and comments.<br>Anonymity, which is used to hide identity of haters, increasingly is being addressed by online services. |
| Monetization and E-Commerce | Many hate groups used PayPal, Amazon, GoFundMe and similar services | Most transactional and funding websites have Terms of Service prohibiting use by hate groups (as defined by the company) and responsiveness increasing |
| Social Media | Limited prohibitions on hate speech, defamation or abusive posts | Universal acknowledgement of hate speech as a problem. More but confusing array of standards and mechanisms in use. Some companies more responsive to complaints than others. |
| Blogs | Lack of meaningful Terms of Service | Google launched updated and unified Terms of Service (3/2014) affecting content rules for most user generated content services. |
| File Drops/Cloud Storage | Limited attention to use of file drops by terrorists and hate groups | Most file-drop sites only prohibit illegal content but do not monitor on privacy grounds. Abuse by hacker and terrorist groups still occurs. |
| Smart Devices/Apps | No ratings for apps, games or other smart device content | Google play-initiated app rating and review system 3/2015. iTunes acknowledged as having the strictest review procedure and permissions. |
| Games | Microsoft Xbox Unit virtually alone in enforcing gaming environment rules | Many online game platforms now using filtering software to monitor and limit inappropriate language used within games as well as users name/profile information. Real time abuse in game environment still problematic. |

## Basic industry practices related to cyberhate and how they have changed in the last three years

| PRACTICES | 2013 | 2016 |
|---|---|---|
| Hate Speech Policies | The Terms of Service for many platforms did not address hate speech directly or used vague terminology in policies | Multiple platforms, including Facebook, Google, Twitter, Amazon, Microsoft gaming, and Yahoo, now include specific prohibition of hate speech |
| User-Friendly Reporting | Complaint mechanisms or contact details were often buried or limited in functionality | Virtually every major service and platform uses post, profile and image flagging. Now standard practice to send receipt of complaint acknowledgements and provide links to further policy/process information. |
| Enforcement Mechanisms | In cases where hate speech was prohibited, penalties were mostly delineated | Google, Facebook, Twitter have instituted flagging for specific posts and partial content removal. Several social media platforms have implemented "stop and think before sending" messages and campaigns. |
| Transparency | Pervasive tendency for companies not to explain why content allowed to remain after a complaint; little explanation offered to users whose material was deleted | Most platforms offer explanations to users whose content has been deleted and provide an appeals process. Complainants on Facebook and YouTube are advised if content has been removed. Public disclosure of rationales for removals is limited. |
| Counter-speech | Counter-speech education by only limited number of companies, and un-coordinated between companies | Counter-speech projects are being studied and changes implemented by major platforms. |

## Challenges that the industry as a whole confronts when dealing with cyberhate

| INTERNAL INDUSTRY CHALLENGES | 2013 | 2016 |
|---|---|---|
| Industry Realities | No effort to broadly explain the challenges created by evolving technology, unintended consequences and the volume of content | Industry platforms are sharing more data on traffic, members' complaints and responses than ever before - but still falling short in adequately illuminating the enormous and ever-growing volume of content and the challenge of addressing issues that require human evaluation and intervention |
| Anonymity | Anonymous participation on many platforms tolerated despite policies to the contrary | Anonymity continues to pose challenges for enforcement of Terms of Service. New technologies are better at detecting users with multiple accounts being used to evade website policy. |
| Industry Coordination | No coordinated industry statements or projects obvious to the public | The Anti-Cyberhate Working Group has become a major venue for the industry to coordinate anti-cyberhate activity.<br>Major breakthroughs: publications of ADL's "Best Practices for Responding to Cyberhate" and well-received Cyber-Safety Action Guide.<br>There is more dialogue between companies on hate related issues than ever before. |
| Hate speech links and linked material | Platforms took no substantial responsibility for third party or linked content | Ongoing debate and discussion regarding platform as publisher and impact of link distribution |
| Corporate Voices | Few if any corporate voices spoke about online hate | Anti-hate speech voices in industry now led by Facebook, Microsoft, and Google with recent important statements by Twitter |

## External challenges that impact the industry's ability to address cyberhate

| EXTERNAL INDUSTRY CHALLENGES | 2013 | 2016 |
|---|---|---|
| Cross Border | Limited coordination of cross border issues | In the borderless environment of the Internet, almost all initiatives and resolution programs remain geographically based |
| Government Intervention | Uncoordinated or unenforceable regulations | Increasing disconnect between online ideals and achievable targets for action compared to laws under consideration and being enacted to curb online hate |
| Cyber-Terror/Hacking | Hacking (website defacement) mainly performed on an opportunistic basis without consistent political motivation or targeting | Sharp increase in politically motivated hacking targeting Jewish institutions and Western interests |

## Activities by non-industry stakeholders to address cyberhate

| STAKEHOLDERS | 2013 | 2016 |
|---|---|---|
| International Bodies | Numerous country-specific orgs- few international networks or associations | Unchanged |
| Academia | Limited external and stakeholder events by major institutions | Centers flourishing at major universities, including Stanford, Harvard, Brandeis, UCLA and Yale in the U.S. |
| Industry — Anti-Cyberhate Working Group | Anti-Cyberhate Working Group-First Steps | Anti-Cyberhate Working Group continues to promote coordination among stakeholders; such coordination is probably still the best hope for productive results |

# A NEW CHALLENGE: TERRORIST USE OF SOCIAL MEDIA

As Internet proficiency and the use of social media grow ever-more universal, so too do the efforts of terrorist groups to exploit new technology in order to make materials that justify and sanction violence more accessible and practical. Terrorist groups motivated by Islamic extremist ideologies are not only using Facebook, Twitter, YouTube and various other emerging platforms to spread their messages, but also actively to recruit adherents who live in the communities they seek to target.

While the fundamental ideological content of terrorist propaganda has remained consistent for two decades – replete with militant condemnations of perceived transgressions against Muslims worldwide, appeals for violence and anti-Semitism – terrorists groups are now able to reach, recruit and motivate extremists more quickly and effectively than ever before by adapting their messages to new technology.

In the past, plots were directed by foreign terrorist organizations or their affiliates, and recruitment and planning generally required some direct, face-to-face interaction with terrorist operatives. Indoctrination came directly from extremist peers, teachers or clerics. Individuals would then advance through the radicalization process through constant interaction with like-minded sympathizers or, as the 2007 New York Police Department (NYPD) report on radicalization described, with a "spiritual sanctioner" who gave credence to those beliefs.

## The Internet and Self-Radicalization

Today, individuals can find analogous social networks, inspiration and encouragement online, packaged neatly together with bomb-making instructions. This enables adherents to self-radicalize without face-to-face contact with an established terrorist group or cell. Furthermore, individual extremists, or lone wolves, are also increasingly self-radicalizing online with no physical interactions with established terrorist groups or cells – a development that can make it more difficult for law enforcement to detect plots in their earliest stages.

The majority of American citizens and residents linked to terrorist activity motivated by Islamic extremist ideologies since 2013 actively used the Internet to access propaganda or otherwise facilitate their extremist activity.

## ISIS Recruitment Online

Since 2014, the Islamic State of Iraq and Syria (ISIS) has been particularly aggressive in pursuing multiple sophisticated online recruiting and propaganda efforts. ISIS's far-reaching propaganda machine has not only attracted thousands of recruits, but has also helped Syria and Iraq emerge as the destinations of choice for a new generation of extremists.

This activity has likely contributed to the increasing number of individuals accused of joining or aiding the terrorist organization. Eighty U.S. residents were linked to Islamic extremist plots and other activity in 2015 nearly triple the total of each of the past two years (28 individuals in 2014 and 22 in 2013).

Globally, at least 20,000 fighters are believed to have traveled to join the conflict in Syria and Iraq, many of whom have joined ISIS. The largest numbers come from the Middle East and North Africa. But non-majority-Muslim countries have seen steady numbers of individuals leaving to fight as well. This includes 800-1,500 from Russia, 1,200

from France, 500-600 from Germany, 500-600 from the United Kingdom, and about 300 from China.

There have also been a surprisingly large number of minors. For example, focusing on the United States, five Americans under the age of 18 were linked to activity motivated by Islamic extremist idealogy in 2014, and four in 2015. This included three Denver, Colorado teenagers, aged 15, 16 and 17. At least one of the girls was encouraged to travel to Syria by an individual she was communicating with online, according to reports. The 15-year-old described her radicalization in a series of Tweets. "I started to notice the people I called 'friends' weren't my true friends. But the people who reminded me about my Deen (religious path) were my TRUE friends." Some of the 16-year-old's Tweets reveal the degree to which she identified with this extreme ideology: "those who identify as 'gay' and 'Muslim' at the same time deserve death," and "Muslims handing out apologizes (sic) because of 9/11 are a disgrace to the Ummah (global community of Muslims)."

Twitter emerged as ISIS's platform of choice in part because it is able to conceal the identities of its users more effectively than other forums and social networking sites. And while accounts are regularly shut down – Twitter indicated that it has shut down more than 125,000 profiles linked to ISIS content since mid-2015 - ISIS supporters continuously attempt to establish new ones.

As Twitter and other platforms attempt to mitigate efforts by ISIS to actively encourage violent extremism by removing content that violates their terms of service, terrorist groups and their supporters continue to seek out new platforms to broadcast their propaganda and connect with adherents. In late 2015, for example, ISIS supporters started migrating to Telegram, a chat and group application available for smartphones and desktop, as their primary medium for official propaganda. While Telegram has since removed all public ISIS affiliated groups, ISIS supporters continue to utilize its private services.

Some efforts by terrorist groups to move to other platforms or create new ones have been less successful. In July of 2014, for example, ISIS announced that its official Internet presence was moving from Twitter to alternate social media sites Friendica and Quitter. Following exposure by ADL, however, all ISIS presence was quickly deleted from Friendica and Quitter, and the group returned to Twitter.

ISIS's online presence is worldwide and presented in multiple languages, as is the propaganda it distributes via social media platforms. The terror group releases online magazines in Arabic, English, Turkish and French, and it has also released statements and videos in other languages, including Hebrew, Spanish, Russian, Kurdish and German.

Official social media accounts are augmented by supporters on social media, some of whom seem to have quasi-official status. These supporters both share official propaganda and contribute to the barrage of online voices supporting terrorist ideology. Some supporters add personal details about their experiences in the group – information that adds to the authenticity of their narratives by providing concrete experiences.

In order to unify its messaging, ISIS has also organized hashtag campaigns on Twitter, encouraging supporters to repeatedly Tweet various hashtags such as #CalamityWillBefallUS, which threatened attacks against the U.S.; #AllEyesOnISIS, which attempted to magnify the number of ISIS supporters on Twitter; and #FightForHim, which called for copycat attacks following the 2014 attacks on the French magazine *Charlie Hebdo*. The apparent goal is for these terms to trend on Twitter, vastly increasing the visibility of tweets.

Similarly, ISIS has used hashtag campaigns to insert its messages into other trending topics on Twitter that have nothing to do with violent extremism. Thus, it will encourage its supporters to tweet ISIS messages with popular hashtags such as #worldcup or #Ferguson so that people searching for those hashtags will inadvertently come across pro-ISIS posts. Hashtag campaigns have been conducted in a number of languages, including English, French, Arabic and Turkish.

ISIS supporters are often active on a variety of platforms beyond Twitter, including the social networking site Facebook, the picture-sharing site Instagram, the chat services Kik and WhatsApp, the video sharing site YouTube, and the question and answer service Ask.FM. These individuals also encourage direct contact with potential recruits via encrypted messaging services such as SureSpot.

On Ask.FM, where users can post questions anonymously, known members of extremist organizations are asked questions by potential recruits. For example, the user Mujahid Miski (believed to be Mohamed Abdullahi Hassan, an Al Shabaab member from Minnesota who has since been taken into Somali custody) answered questions including, "My brother wants to be a *mujahid* (fighter) but he's got glasses. Will that stop him from becoming one?" Many of his answers also include encouragement for readers to join terrorist groups, including ISIS. In one, for example, he wrote, "every minute and every second is wasted if you're not out there building the Islamic Caliphate (a reference to ISIS). Go out and make *hijrah* (migration to a Muslim land) from the east and the west and join *jihad* (the fighting). Let your blood be the water for the tree of *Khilafah* (caliphate, a reference to ISIS).

Many ISIS supporters also take advantage of the websites Justpaste.it and its Arabic-language counterpart Manbar.me, which enable them to quickly publish content to unique URLs online that can then be shared on social media. ISIS supporters have used these sites to publish links to downloadable propaganda materials, instructions for traveling to Syria and Iraq, manifestos encouraging lone wolf attacks, and more.

A number of ISIS supporters maintain blogs on which they detail their extremist ideology and narratives of an idealized day-to-day life which they hope will appeal to potential recruits. There have also been instances of ISIS supporters creating new websites to make ISIS propaganda even more accessible. In February 2015, an ISIS supporter created a website called IS-Tube that featured a searchable archive of ISIS propaganda videos, including videos depicting beheadings. The site was hosted on a Google-owned IP-bloc, and was removed after ADL alerted Google to its presence.

Online repositories of terrorist propaganda are not unique to Google, yet some platforms do not have clear or effective policies regarding terrorist content, enabling terrorists and their supporters to exploit their services more easily and uninterrupted. For example, WordPress hosts a website that features hundreds of ISIS propaganda videos, statements and publications. Among the hundreds of items on the site are beheading and execution videos, as well as videos and articles encouraging Westerners to travel to join ISIS or to commit attacks on its behalf in their home countries. The site remains online despite efforts to flag the material.

## Other Terrorist Groups Using the Internet

Other terrorist organizations use social media as well, and many have learned from ISIS's techniques. During the 2014 conflict between Israel and Hamas, for example, ADL documented multiple social media profiles that could be considered official Hamas accounts.

Like ISIS followers, Hamas supporters utilized hashtag campaigns to promote terror attacks against Israelis and posted videos and images to social media that both applauded and encouraged killing Israelis and Jews with hatchets and by running them over. Indeed, instructional videos on stabbing, clips of preachers calling for attacks on Jews, graphic images are all going viral. Such incitement on social media is widely understood as having a significant link to the stabbing attacks against Israelis, and the online approbation of each attack further spreads the message and encourages would-be attackers.

The increase in small arms attacks in both the U.S. and abroad serves as a testimony to the potential power of social media. Spurred at least in part by extortions by ISIS propaganda on social media to undertake attacks by any means

possible, including with knives, the U.S. has seen an increase in small arms attacks by apparent terrorist sympathizers. These have been directed at law enforcement in particular, but pose a more general threat as well.

Advances in technology have enabled terrorist video production to rival high quality films. ISIS even released a feature-film length video, titled "Flames of War," that portrayed the group as part of an apocalyptic struggle of good versus evil. Other terrorist groups – including Al Qaeda, Al Shabaab (Al Qaeda in Somalia), Boko Haram, Taliban affiliates, the Caucasus Emirates and more – have also distributed propaganda videos via Twitter in recent years.

Perhaps the most infamous English-language terrorist magazine, Inspire, is now distributed via Twitter instead of on extremist forums. An online English-language propaganda magazine produced by Al Qaeda in the Arabian Peninsula (AQAP), Inspire provides articles about terrorist ideology, recruitment information, and encouragement and instructions for homegrown attacks, including the very bomb-making instructions that the Tsarnaev brothers allegedly utilized in the 2013 Boston Marathon bombing.

An article in the magazine's second issue encouraged "brothers and sisters coming from the West to consider attacking the West in its own backyard. The effect is much greater, it always embarrasses the enemy, and these types of individual attacks are nearly impossible for them to contain." Its 2014 editions contained directions for making car bombs and bombs designed to evade airport security measures, as well as instructions regarding the best places to detonate them.

Outside the sphere of social media, terrorist groups and sympathizers have also attempted to create applications promoting their organizations and propaganda on iTunes and Google Play.

Hezbollah, for example, has launched a number of applications that provide streaming access to the group's propaganda-based television station, Al Manar. Google Play and iTunes have been quick to remove them, but Hezbollah, having blamed "the Jewish Anti-Defamation League" for launching a "campaign" to remove the original application, has created the applications so users can download them directly from the Hezbollah website, without going through iTunes or Google Play. Hezbollah has also created several video games on its website with the explicit intent of indoctrinating young players.

Other applications are created by terrorist supporters. The Anwar al-Awlaki application, for example, enabled users to listen to Awlaki's sermons directly from their mobile devices. Awlaki, the creator of Inspire magazine, was the primary English-language spokesman for AQAP until he was killed by a drone strike in 2011. Awlaki remains tremendously influential. Many of his lectures are still available on YouTube, and supporters regularly create Facebook and Twitter profiles dedicated to sharing his quotes. When these profiles are removed, they are quickly replaced by new ones. A significant number of domestic Islamic extremists, including the Tsarnaev brothers, have accessed his propaganda and cited him as an inspiration.

## Another Challenge: Islamic Extremists' Hacking Activity

Perhaps the newest frontier of online extremism comes in the form of Islamic extremist hackings. Politically motivated hackers from the Arab world have begun targeting the websites of perceived supporters of Israel, including synagogues, Jewish institutions, and individuals. These attacks are increasingly undertaken in the name of terrorist organizations, particularly ISIS. There are signs that ISIS is beginning to attempt to harness the hackers and hacker groups into supporting its own mission and expanding the hacks to target websites and government institutions in the U.S.

In March 2015, for example, hacker(s) identifying as "ISIS cyber army" claimed responsibility for hacking 51 American

websites on March 24. Each of the hacked websites was defaced with the ISIS flag, a statement that the website was "Hacked by Islamic State" and an e-mail address for the ISIS cyber army, the unit believed to be behind the cyber activities of ISIS. In the past, the ISIS cyber unit claimed responsibility for involvement in a series of attacks against a number of Israeli websites.

This capability to engage in cyber-attacks may be a reflection of ISIS's calls for support from individuals with various skills, from media experts to doctors, to join and contribute to the group and its mission of gaining strength and territory however they can.

In April 2015, as international hackers once again set their sights on Jewish and Israeli targets as part of "OpIsrael," an annual anti-Israel cyber-attack campaign, there were strong indications that AnonGhost, an international hacker group that supports terrorist groups and frequently employs anti-Semitism as part of its cyber activity, had replaced Anonymous as the main organizer of the campaign.

Groups such as AnonGhost express unambiguous support for the Palestinian terrorist group Hamas and the Islamic State (ISIS) and have carried out cyber-attacks in their names, bringing an Islamic extremist element into an already virulently anti-Israel and anti-Semitic campaign.

AnonGhost threatened individual Israelis with violence through mobile devices, claiming to have obtained personal information on more than 200 Israelis. One threatening text the group claims to have sent to an Israeli included an image of an infamous ISIS fighter with the caption, "We are coming O Jews to kill you." A text sent to another Israeli man included an image of his family with the threat, "I'll stick a knife in their throats."

While anti-Semitic themes existed in previous OpIsrael campaigns, it had been primarily billed as a response to the Israeli-Palestinian conflict. AnonGhost's participation and tactics thus far speak to the centrality of anti-Semitism in this year's campaign, which serves as an extension of AnonGhost's pro-terror activism around the world.

## Anti-Semitism: A Pillar of Islamic Extremist Ideology

As new technology and social media continue to alter the nature of global communications, terrorist groups have quickly adapted to these tools in their efforts to reach an ever-widening pool of potential adherents. As a result, anti-Semitism in its most dangerous form is easily accessible by a worldwide audience.

In a video message in August 2015, Osama bin Laden's son, Hamza bin Laden, utilized a range of anti-Semitic and anti-Israel narratives in his effort to rally Al Qaeda supporters and incite violence against Americans and Jews.

Bin Laden described Jews and Israel as having a disproportionate role in world events and the oppression of Muslims. He compared the "Zio-Crusader alliance led by America" to a bird: "Its head is America, one wing is NATO and the other is the State of the Jews in occupied Palestine, and the legs are the tyrant rulers that sit on the chests of the peoples of the Muslim Ummah [global community]."

Bin Laden then called for attacks worldwide and demanded that Muslims "support their brothers in Palestine by fighting the Jews and the Americans... not in America and occupied Palestine and Afghanistan alone, but all over the world…. take it to all the American, Jewish, and Western interests in the world."

While such violent expressions of anti-Semitism have been at the core of Al Qaeda's ideology for decades, terrorist groups motivated by Islamic extremist ideology, from Al Qaeda to ISIS, continue to rely on depictions of a Jewish enemy – often combined with violent opposition to the State of Israel – to recruit followers, motivate adherents and draw attention to their cause. Anti-Israel sentiment is not the same as anti-Semitism. However, terrorist groups often

link the two, exploiting hatred of Israel to further encourage attacks against Jews worldwide and as an additional means of diverting attention to their cause.

And they have more tools at their disposal than ever before.

Recent terrorist attacks against Jewish institutions in Europe, and the spike in incitement materials encouraging stabbing and other attacks against Jews and Israelis around the world, not only speak to the global reach provided by these new technologies, but also to the pervasive nature of anti-Semitism in terrorist propaganda that encourages violence directed at Jews.

In September 2015, ADL issued a report examining the nature and function of anti-Semitism in terrorist propaganda today. It focused on ISIS, Al Qaeda Central, and two of Al Qaeda's largest affiliates, Al Qaeda in the Arabian Peninsula (AQAP) in Yemen and Al Shabaab in Somalia, as well as the prevalence of anti-Semitism among supporters of Palestinian terrorist organizations. It also provides examples of individuals linked to terrorist plots and other activity in the U.S. that were influenced, at least to some degree, by anti-Semitic and anti-Israel messages.

In light of the growing sophistication and reach of terrorist use of social media, developing and implementing strategies to respond to this challenge must be a shared priority, not just for the Internet industry and counterterrorism experts, but for all of us.

# BEST PRACTICES FOR RESPONDING TO CYBERHATE

In May 2012, the Inter-Parliamentary Coalition for Combating Anti-Semitism, an organization comprised of parliamentarians from around the world working to combat resurgent anti-Semitism, asked the Anti-Defamation League (ADL) to convene a Working Group on Cyberhate. The mandate of the Working Group was to develop recommendations for the most effective responses to manifestations of hate and bigotry online. The Working Group includes representatives of the Internet industry, civil society, the legal community, and academia.

The Working Group has met five times, and its members have graciously shared their experiences and perspectives, bringing many new insights and ideas to the table. Their input and guidance have been invaluable, and are reflected in the following Best Practices. Obviously, the challenges are different for social networks, search engines, companies engaged in e-commerce, and others. Nevertheless, we believe that these Best Practices could contribute significantly to countering cyberhate.

## PROVIDERS

1. Providers should take reports about cyberhate seriously, mindful of the fundamental principles of free expression, human dignity, personal safety and respect for the rule of law.
2. Providers that feature user-generated content should offer users a clear explanation of their approach to evaluating and resolving reports of hateful content, highlighting their relevant terms of service.
3. Providers should offer user-friendly mechanisms and procedures for reporting hateful content.
4. Providers should respond to user reports in a timely manner.
5. Providers should enforce whatever sanctions their terms of service contemplate in a consistent and fair manner.

## THE INTERNET COMMUNITY

6. The Internet Community should work together to address the harmful consequences of online hatred.
7. The Internet Community should identify, implement and/or encourage effective strategies of counter-speech – including direct response; comedy and satire when appropriate; or simply setting the record straight.
8. The Internet Community should share knowledge and help develop educational materials and programs that encourage critical thinking in both proactive and reactive online activity.
9. The Internet Community should encourage other interested parties to help raise awareness of the problem of cyberhate and the urgent need to address it.
10. The Internet Community should welcome new thinking and new initiatives to promote a civil online environment.

In addition to the above, ADL would offer the following recommendations for responding to terrorist use of social media.

# Recommended Responses to Terrorist Use of Social Media

1. Providers should give priority attention to how their platforms are being used by terrorists and terrorist groups to promote terrorism, to recruit potential new terrorists, and to foster self-radicalization.
2. Providers should make their expertise available to those looking to generate and promote counter-narratives.
3. Providers should work with interested stakeholders to analyze the impact of counter-narratives in terms of their reach, scope, and effectiveness.
4. Providers should consider creating a specific new terrorism category for users seeking to flag terrorism-related content.
5. Providers should use their corporate voices to condemn terrorist use of their platforms and to explain why terrorist activity and advocacy is inconsistent with their goals of connecting the world.

Underlying all of these recommendations is the understanding that rules on hate speech may be written and applied too broadly so as to encumber free expression. Thus, an underlying principle for these recommendations is that care should be taken to respect free expression and not to encumber legitimate debate and free speech.

# APPENDIX: ADL CYBER-SAFETY ACTION GUIDE

This Appendix features ADL's Cyber-Safety Action Guide, which in its online form brings together in one place the relevant Terms of Service addressing hate speech of major Internet companies. Individuals and groups seeking to respond to various manifestations of hate online have found it to be a unique and very useful tool, and the list of participating companies continues to grow.

amazon.com
PayPal
ask.fm
Pinterest
at&t
Quizlet
comcast
reddit
ebay
topix
facebook
tumblr.
Go Daddy.com
Google
vimeo
Google Play
Instagram
YAHOO!
LinkedIn
YouTube
myspace

## ANTI-DEFAMATION LEAGUE

Marvin D. Nathan
*National Chair*

Jonathan A. Greenblatt
*CEO & National Director*

Kenneth Jacobson
*Deputy National Director*

Deborah M. Lauter
*Senior Vice President, Policy and Programs*

Steven M. Freeman
*Deputy Director, Policy and Programs*

Glen S. Lewy
*President, Anti-Defamation League Foundation*

## CIVIL RIGHTS DIVISION

Elizabeth A. Price
*Chair*

Christopher Wolf
*Chair, Internet Task Force*

Jonathan Vick
*Assistant Director, Cyberhate Response*

## CENTER ON EXTREMISM

Jared Blum
*Chair*

David Friedman
*Vice President, Law Enforcement, Extremism and Community Security*

Oren Segal
*Director*

This work is made possible in part by the generous support of:
William and Naomi Gorowitz Institute on Extremism and Terrorism
Marlene Nathan Meyerson Family Foundation
Charles and Mildred Schnurmacher Foundation, Inc.

For additional and updated resources please see: www.adl.org
Copies of this publication are available in the Rita and Leo Greenland Library and Research Center.



Anti-Defamation League®

Anti-Defamation League
605 Third Avenue, New York, NY 10158-3560
www.adl.org


605 Third Avenue
New York, NY 10158-3560
www.adl.org