



Task Force on Internet Hate

Report and Recommendations of the Co-chairs*

May 29, 2013

-
- Yuli Edelstein, MK and former Member of the Steering Committee and Christopher Wolf, Partner at Hogan Lovells US and Chair of the National Civil Rights Committee of the Anti-Defamation League.
 - Special thanks to Steven Sheinberg and Robert Trestan, professionals at the Anti-Defamation League for their substantial contributions to the work of the Task Force and to this Report and Recommendations. Thanks also to our colleagues at Hogan Lovells who contributed pro bono services to assist in the fact-finding and legal research associated with the work the Task Force, and to Jardena Lande, Director of the ICCA. And, of course, thanks to MP John Mann, whose vision for the creation of the ICCA and the appointment of the ICCA Task Force made our work possible.

Note from the Chair of the Inter-parliamentary Coalition for Combating Antisemitism, John Mann MP

In November 2010 participants at the second Conference of the Inter-parliamentary Coalition for Combating Antisemitism (ICCA) in Ottawa unanimously adopted the Ottawa Protocol on Combating Antisemitism. In it, the parliamentarians from around the world tasked the ICCA with

'Establishing an International Task Force of Internet specialists comprised of parliamentarians and experts to create common indicators to identify and monitor antisemitism and other manifestations of hate online and to develop policy recommendations for Governments and international frameworks to address these problems.'

Internet Hate is a grave development of recent times, and while the advancement of technology has brought many advantages and made the world a smaller place, the simplified spreading of hate and incitement is at the same time a seriously dangerous phenomenon which must be addressed.

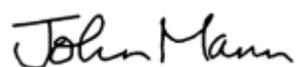
ICCA Steering Committee Member Yuli Edelstein, Speaker of the Israeli Parliament, and Christopher Wolf, Partner at Hogan Lovells US and ADL National Civil Rights Committee Chair, were appointed co-chairs of this new Task Force and together with the ICCA they recruited members for this new initiative. The group was made up of selected experts, parliamentarians and most uniquely representatives of the Internet industry and was commissioned to write a report which will be presented to the ICCA leadership. After several meetings and much hard work from the Task Force members and the co-chairs, we now welcome the outcome of their deliberations.

On behalf of the ICCA I would like to thank the co-chairs for their dedication and initiative and in particular the Anti-Defamation League (ADL), a Task Force member, for the leading role they took on in the editing of this report and all the hard work they invested in it.

John Mann

Member of Parliament, United Kingdom

Chair of the Inter-parliamentary Coalition for Combating Antisemitism



Background: The Task Force

The Inter-Parliamentary Coalition for Combating Anti-Semitism (ICCA) is an organization comprised of parliamentarians from around the world working to combat resurgent global anti-Semitism. Following the November 2010 conference of the ICCA in Ottawa, Canada, the parliamentarians adopted a Protocol that included a commitment to establishing an International Task Force of Internet Specialists comprised of parliamentarians and experts. The goal of this task force is to create common indicators to identify and monitor anti-Semitism and other manifestations of hate online and to develop policy recommendations for governments and international frameworks to address these problems.

Accordingly, task force chairs Yuli Edelstein (Member of Knesset, Israel) and Christopher Wolf (Hogan Lovells LLP, USA) assembled a task force with a wide range of parliamentarians, NGOs and experts (Exhibit A).

The first meeting of the Task Force was held at the Parliament in London, United Kingdom, on 27 October 2011 on the topic *The Scope, Nature and Effect of Online Hate Speech*. In addition to a full roster of witnesses and expert testimony (Exhibit B), the Task Force received extensive evidence prepared pro-bono by the international law firm Hogan Lovells (Exhibit C). The Task Force also created the following subcommittees: *Counter-speech: How can we explore opportunities, how can we promote it; Education; Intermediaries; and Law: Common understanding of commonalities, recommendations on how to enforce laws*.

In May 2012, the Task Force met again at the Center for Internet and Society at Stanford Law School in Palo Alto, California, USA. The proximity to Silicon Valley enabled significant participation by content hosts, search providers, and application service providers ("intermediaries"). The Task Force received subcommittee reports (several of which are incorporated by reference in this Task Force report). The Task Force also adopted a voluntary statement of principles for companies that routinely deal with hate speech and hate content and adopted a resolution concerning the creation of a permanent and standing working group to address multiple approaches to counter online hate. Given the participation and cooperation of intermediaries in this working group alongside parliamentarians, scholars, free speech advocates and civil rights advocates, the creation of this working group is an unprecedented and enormously productive advance.

This report on the work of the Task Force will:

- Summarize the Task Force's observations on the nature of cyberhate
- Discuss the testimony received concerning the targets and geography of cyberhate
- Provide recommendations on how to address the problem of cyberhate.

The Nature of Cyberhate

Over the last 20 years, the Internet has revolutionized the way we communicate, connect, and spread information, and with the Internet now easily available on cell phones and other mobile devices, it is usually easily accessible from almost any location. Consequently, the Internet has also become a vehicle for hate and intolerance—cyberhate. Specifically, cyberhate is defined as the use of any electronic technology to spread bigoted, discriminatory, terrorist, and extremist information.

Given the Internet's role as a powerful and virulent platform for anti-Semitism, racism, misogyny, homophobia and bigotry of all kinds, it is important that policy makers focus their attention on the dangers of cyberhate.

Although NGOs dedicated to combatting hate have focused their attention on cyberhate, many policy leaders have yet to focus substantially on the issues of cyberhate and methods of combatting it. Given the Internet's role as a powerful and virulent platform for anti-Semitism, racism, misogyny, homophobia and bigotry of all kinds, it is important that policy makers focus their attention on the dangers of cyberhate.

The media has always been a tool for disseminating hate, as epitomized by Hitler and the Nazis who used propaganda to spread anti-Semitism throughout the Holocaust. As pointed out in my new book with Abraham Foxman, *Viral Hate: Containing its Spread on the Internet*, the Internet is a vehicle of hate that Hitler and the Nazis could never have dreamed of:

Instead of being under the central control of a political party or group, the power of the Internet lies in its viral nature. Everyone can be a publisher even the most vicious anti-Semite, racist, bigot, homophobe, sexist or purveyor of hatred." Hateful content can spread to millions of Internet user in seconds and can often go unchallenged, especially in online communities of bigots.

Turning to specifics, in addition to its positive uses, information technology has become a tool to spread incredible amounts of bigotry and hatred, so-called "cyberhate."

The Internet is used nearly universally as a tool of spreading hateful content by hatemongers, including:

- Right Wing Extremists

- Left Wing Extremists
- Extreme Anti-Government Groups
- Terrorists
- Bigots, including anti-Semites, misogynists, racists, homophobes, etc.

In these environments, even the most bizarre and bigoted views are shared, discussed, and move toward the mainstream.¹

As technology evolves, so do the ways in which hate and terrorism spread. With technology's help, bigotry, hate and extremism have moved beyond physical, political, and social boundaries. In order to address the problem of cyberhate, it is critical to understand how hatemongers use information technology to spread hate. Key uses by hatemongers include:

1. **Web 1.0. Standard websites.** These play a key role in allowing hatemongers to easily and effectively post searchable information, share ideas, recruit, coordinate, and encourage participation. Examples include the early (but still current) "supermarket of white supremacist hate" <http://www.Stormfront.org>;² <http://www.martinlutherking.org/>, which presents a fictional and racist account of the American civil rights leader, Dr. Martin Luther King, Jr., targeted toward unsuspecting researchers and children; "Why Christians Ought to "Hate" Homosexuals" at Craig for [United States] Congress, <http://kevincraig.us/homophobia.htm> 3/15/2013; and Neo-Nazi websites which allow users to maneuver virtual nooses over digital images of black men.
2. **User-submitted content.** Services such as video sharing sites, where hate videos can be put up at will (often in violation of a site's terms of service or guidelines), allowing users to share their own multimedia productions and convey their messages with a degree of sophistication never seen before. The effects of this information sharing have serious real world

¹ For instance In the aftermath of the mass shooting tragedy at Sandy Hook Elementary in the United States, Internet posts blaming Jews for the killings began to spring up on various online message boards and conspiracy theory websites. One theory, for example, promotes the notion that a supposedly Jewish-controlled Hollywood encourages killing sprees with subliminal messages glorifying gun violence. Another attributes the violence to the State of Israel, claiming Israel perpetrated a "false flag" terrorist attack on American soil as alleged punishment for American foreign policy on Israel and the Palestinians. See <http://blog.adl.org/anti-semitism/conspiracy-theorists-blame-jews-for-sandy-hook-massacre>, 12/18/12.

² See, e.g., <http://blog.adl.org/civil-rights/white-supremacists-plan-anti-immigration-rallies-nationwide-in-february-2013>

consequences.³ In addition YouTube clones, home-made video hosting sites specifically for hate videos, offer a platform for hatemongers who wish to spread their bigotry online. Although hateful content hosted on websites such as YouTube may be removed for terms of service violations, the content can easily spread before its removal from server to server, country to country.⁴

3. **Social networking.** The use of social networking as a vehicle for spreading hate is especially pernicious because hateful content can be spread to wide audiences whose relationship with a hatemonger may be simply tangential.

The greatest increase in digital hate has occurred on social media sites. For example, the Facebook group *Kill a Jew Day* declared July 4, 2010 as the start of an eighteen-day period of violence "anywhere you see a Jew." The group's profile featured a swastika and images of corpses piled on top of one another and group members commented that they could not "wait to rape the dead baby Jews."

Additional examples include the *How to Kill a Beaner* video posted on YouTube, which allowed players to kill Latinos while shouting racial slurs and the Facebook group *Kick a Ginger Day*, which inspired physical attacks on students with red hair. Facebook has hosted groups such as *Hitting Women*, *Holocaust Is a Holofoax*, and *Join if you hate homosexuals*.⁵

³ In the months before his arrest for allegedly plotting to attack the Military Entrance Processing Station in Seattle, Abu Khalid Abdul-Latif shared and uploaded various videos on YouTube. He subscribed to a number of radical YouTube channels, including an [Anwar al-Awlaki](#) channel that features several sermons by the U.S. born Muslim cleric whose online sermons have influenced scores of extremists in the West. In a comment posted in response to one of these sermons, Abdul-Latif made reference to the [Fort Hood shooting](#), writing "hopefully there will be more soldiers who come out of the woodwork to serve Allah." See http://archive.adl.org/main_Terrorism/latif_seattle_plot.htm, 6/30/2011.

⁴ See also "Anti-Semitism right on your ipod," <http://blog.adl.org/international/anti-semitism-right-on-your-ipod>, 12/28/01. ("A variety of anti-Semitic pod-casts by scholars and religious leaders in the Arab world are available for download through iTunes.")

⁵ Danielle Keats Citron & Helen Norton, *Intermediaries and Hate Speech: Fostering Digital Citizenship* 91 B.U.L. Rev. 1435 at 1436-1437.

4. **Revenue generation.** Violent organizations, terrorists, and hatemongers use online revenue to fund their hateful messages and acts. As we consider online hate we have to consider online revenue streams.

5. **Gaming.** Fully interactive gaming has become a home for online hate and bigotry. By allowing users to “connect” directly to each other, such online games have allowed haters easy access to our living rooms. Said one leading online gamer:

At [video game] tournaments...I've been called a dyke, a butch, a slut, a bitch... I was even called a black bitch to my face along with being called a lesbian, a gorilla, and a monkey.

“At tournaments players talk [crap] to each other. That’s just the way tournaments are. People get hyped. Players get salty when they lose, which is fine. But there is a difference between trash talking and calling other players disrespectful names. For me, I’ve been called a dyke, a butch, a slut, a bitch... I was even called a black bitch to my face along with being called a lesbian, a gorilla, and a monkey. Now I know people are going to say that as a player in the community, you have to have a thick skin. I do, but that doesn’t mean that I won’t speak up about the names I’ve been called. Because these names refer to my sex, but most of them refer to my race; which to me is racist. I think some of these people are blurring the lines between trash talk and disrespectful trash talk. And again, this is just my experience on the matter. I don’t know if anyone else has had this experience. So I wrote a note on my Facebook, made it private, and got the opinions of several other black female gamers. They all have had somewhat the same type of experience as I, some have seen it and others have heard of it.”⁶

6. **Texting and smart phones.** Texting (and other smart phone apps) allow 24/7 communications in a way that bypasses any gatekeeper, access control or even meaningful record keeping.⁷ Consequently, such

⁶ <http://www.racialicious.com/2011/04/21/quoted-burn-your-bra-on-racism-and-body-image-in-gaming/>, 4/21/2011.

⁷ According to CNET, as of 2010, AT&T, T-Mobile, and Sprint did not store the contents of text messages. Verizon did for up to five days, a change from its earlier no-logs-at-all position, and Virgin Mobile kept them for 90 days. The carriers generally kept metadata such as the phone numbers associated with the text for 90 days to 18 months; AT&T was an outlier, keeping it for as long as seven years, according to the chart. CNET, “Cops to Congress: We Need Logs of Americans’ Text Messages,” http://news.cnet.com/8301-13578_3-57556704-38/cops-to-congress-we-need-logs-of-americans-text-messages/, 12/12/2012.

ubiquitous and persistent messaging allows constant access to children and provides unparalleled opportunities for violent bigots to coordinate and connect.

7. **Force multiplication.** Hatemongers who use information technology can do so in aggressive and violent fashion, from enlisting others to commit crimes to using technological means to attack and damage critical infrastructure. For instance, The Department of Homeland Security's Office of Intelligence and Analysis (I&A) judges that a number of emerging trends point to leftwing extremists maturing and expanding their cyber attack capabilities over the next decade with the aim of attacking targets in the United States.⁸

With this in mind, the task force heard testimony that three key trends are emerging:

- New ways of using the web (such as social networking sites like Facebook and user-generated content sites such as YouTube) have led to an explosion of online bullying and hate. Social networking sites are used to promulgate hate and extremist content, increasing the depth and breadth of hate material that confronts non-extremist users.
- Electronic bullying, hate and extremism have evolved off the Web. Ubiquitous technologies, such as smartphones, have allowed youth to readily engage in bullying while remaining anonymous. Ultra-realistic, real-time, and fully interactive video games have put hate content directly into the hands of children and teenagers. These technologies also allow haters and terrorists to reach anyone—including children—anytime and anywhere.
- Haters exploit the same technologies that more benign users use, creating a dark shadow web of hate content, commerce, and money laundering. By using sophisticated online storefronts, easy-to-program video games, and software that lets any user create sophisticated websites, bigots and terrorists have increased their fundraising, information-sharing, and recruitment capabilities.

⁸ See DHS, “Leftwing Extremists Likely to Increase Use of Cyber Attacks over the Coming Decade” at <http://www.fas.org/irp/eprint/leftwing.pdf>; See also White, J.R. (2006). *Terrorism and Homeland Security*. (5th ed.). Belmont: Thomson-Wadsworth (Technology as a key force multiplier for terrorism)

PART ONE: The Geography and Targets of Hate Speech

The Task Force investigated the nature and extent of cyberhate. Cyberhate, online hate speech, which has proliferated as the open platform of the Internet has lent itself to increasing online incivility, the targeting of minorities, and the spread of lies and inaccurate information meant to mislead and spur violence.⁹

Possible categories of cyberhate

- Racism
- anti-Semitism
- religious bigotry
- homophobia
- bigotry aimed at the disabled
- political hatred
- rumor-mongering
- misogyny and violent pornography
- promotion of terrorism
- cyberbullying
- harassment and stalking
- speech that silences counter-speech, like slurs, insults and epithets
- speech that defames an entire group.

As reported in *Viral Hate*, the following categories of speech could be (and often times have been) considered hate speech: Racism; anti-Semitism; religious bigotry; homophobia; bigotry aimed at the disabled; political hatred; rumor-mongering; misogyny and violent pornography; promotion of terrorism; cyberbullying, harassment, and stalking; and the sale and promotion of hate-oriented products.¹⁰ Other possible definitions include speech that silences counter-speech, like slurs, insults and epithets, or speech that defames an entire group.¹¹

Hate speech in electronic communications takes many forms. Internet hate encompasses websites, social media (e.g. Facebook, Twitter, YouTube), and e-commerce. Extremist and hate groups host their own websites with impunity. They use their websites to spread propaganda, offer lessons on how to implement guerrilla warfare techniques, allow like-minded people to interact with one another and to recruit additional members.¹² The danger is exacerbated by the fact that many extremist groups meet in password-protected or otherwise closed forums.¹³ A

leader of the Ku Klux Klan has said that the Internet has been the biggest breakthrough in recruiting in the 30 years since he was involved with the group.¹⁴

⁹ [Viral Hate, p. 7]

¹⁰ Id, 49-50.

¹¹ [Citron and Keats, BU Law Rev 1466-67]

¹² [Viral Hate, p. 11]

¹³ [Testimony of Jamie Bartlett to ICCA Task Force, October 27, 2011]

¹⁴ [Viral Hate, p. 12]

Social media and Internet intermediaries can be even more effective at spreading hate. Sites like Facebook, Twitter, YouTube, and Tumblr make it easy for hateful videos, songs, links, comments, and images to spread virally.¹⁵ Indeed, social networking sites are especially risky because they air more subtle expressions of hate and are aired to a wider and younger audience than traditional hate-based websites.¹⁶ They also allow like-minded people to find each other even more easily. Facebook, for example, has hosted such groups as *Hitting Women* and *Join if you hate homosexuals*.¹⁷

Search engines like Google, Bing and Yahoo!, as well as browsers like Mozilla, also play a role in spreading hate.¹⁸

Internet hate includes the less obvious worlds of online gaming and e-commerce, as well. Game creators spread their propaganda through gaming by making minorities the targets of aggression and celebrating and encouraging violence against them, which can make them targets of real-world violence.¹⁹ Regarding e-commerce, the Internet has made it easy for people to bypass reputable retailers and sell racist T-shirts, books, movies, posters, and Nazi memorabilia.²⁰

Internet hate can include a broad range of secondary media as well, such as links to websites, viral emails, and re-tweets of Twitter posts.²¹ In all, there are now thousands of websites, videos, and social network groups devoted to spreading hate.²²

The Danger of Internet Hate

The Task Force heard testimony and received evidence as to the danger of cyberhate.

There is no doubt that words are powerful. One need only look at the Nazi propaganda machine as an example of how words can create a dangerous climate of hatred.²³ The Internet has only magnified the danger.

¹⁵ [Viral Hate, p.19]

¹⁶ [Testimony of Dr. Nathan Hall to Task Force, October 27, 2011, at 2]

¹⁷ [BU Law Rev, 1437]

¹⁸ [BU Law Rev, 1439]

¹⁹ [Viral Hate, p. 19 - 20]

²⁰ [Viral Hate, p. 20-22]

²¹ [Viral Hate, p. 8]

²² [Citron & Norton, BU Law Review, 1436]

²³ [Viral Hate, p. 9]

People who before the Internet age may have lived on the fringe and kept their radical and minority views to themselves now find validation on the Internet because they can easily find and interact with like-minded people.²⁴ These people may begin to believe that the viewpoints they harbor are normal and mainstream.²⁵ Moreover, the Internet provides a cover of anonymity that empowers people to express and share ideas that they would be too ashamed to express in real life.²⁶ The Internet also allows for lies to quickly spread, so that even if the truth later comes out, it is often too late and the damage has been done.²⁷

The danger is real because Internet hate easily translates into real-world violence.²⁸ The Internet allows haters to draw inspiration from online sources, to engage with fellow like-minded people, to become emboldened, to act on their rage, and to coordinate attacks. The recent bombings in Boston Massachusetts starkly emphasize that point: Reports are emerging that Tamerlan and Dzhokhar Tsarnaev, the brothers allegedly responsible for the April 15 Boston Marathon bombings, were radicalized, at least in part, by radical cleric Anwar al-Awlaki – whose primary platform was online video – a platform that survives his death.²⁹ ADL and others report that Awlaki's influence can be seen in at least nine other plots.³⁰

²⁴ [Viral Hate, p. 10]

²⁵ Viral Hate, p. 17] ; *see also* Ronald Eissens, *The Geography of Online Hate Speech*, Testimony presented at the Hearing of the Inter-Parliamentary Coalition for Combatting Anti-semitism Task Force on Internet Hate (Oct. 27, 2011) at 3 (commenting that “enemies lists” in Russia can now be found in the mainstream media); *Eissens* at 4 (citing the European Commission on Racism and Intolerance’s report on Poland that “a particularly worrying aspect is the tacit acceptance of anti-Semitism by an influential media group belonging to a Catholic organization and sometimes even by mainstream political parties”).

²⁶ [Viral Hate, p. 15] *see also* Citron & Norton, *BU Law Review*, 1447 (“The Internet facilitates anonymous and pseudonyms discourse, which can just as easily accelerate destructive behavior as it can fuel public discourse.”).

²⁷ [Viral Hate, p. 16].

²⁸ *See* Edgar Burch, *Comment: Censoring Hate Speech In Cyberspace: A New Debate in a New America*, 3 *N.C. J.L. & Tech.* 175, 178 (2002) (“There is a fine line between the exchange of values and ideas and the perpetuation of hate in the form of degradation and violence.”); *see also* Lyrissa Barnett Lidsky, *Fifth Annual Criminal Law Symposium: Criminal Law & the First Amendment: The First Amendment, the Internet, and the Criminal Law: Incendiary Speech and Social Media*, 44 *Tex. Tech. L. Rev.* 147, 148 (2011) (noting that the “Internet in general and social media in particular amplify the potential for speech to cause violence”).

²⁹ *See* Anti-Defamation League “Boston Marathon Bombers Inspired By Anwar al-Awlaki,” May 3, 2013, <http://blog.adl.org/extremism/boston-marathon-bombers-anwar-al-awlaki>.

³⁰ *Id.* *See also* Chris Montgomery, *Can Brandenburg v. Ohio Survive the Internet and the Age of Terrorism?: The Secret Weakening of a Venerable Doctrine*, 70 *Ohio St. L.J.* 141, 161-62 (noting that government

Scholars suggest that Internet hate can cause violence because it follows the model of mob dynamics.³¹ In mob dynamics, groups with homogeneous views become more extreme as the members' interactions reinforce preexisting views; people in groups start to lose their individuality, which encourages them to act on "destructive impulses"; groups encourage their members to view their victims as lacking humanity and personal identity, leaving them to become more destructive toward their targets; and group members become more aggressive if they perceive that authority figures are supporting their efforts.³² As one scholar noted:

[T]he sheer size of prospective audiences also increases the potential for violent audience reactions. Audience size matters: the bigger the audience, the greater the chance at least one audience member will respond with violence to speech that is offensive or advocates violence...[I]nteractions with these subcommittees may serve to foster group violence or to 'normalize' individual violence.³³

Investigations of rampage shooters, for example, often find that the shooter visited or frequented racist online sites and forums.³⁴ The following are just a few examples:

- In 1999, a member of a white supremacist group admitted after shooting African-Americans, Asian Americans, and Jews in suburban Chicago that "It wasn't really 'til I got on the Internet, read some literature of these groups that...it really all came together."³⁵
- The terrorist who detonated a bomb and executed 79 people gathered at a political summer camp in July 2011 in Norway had been radicalized through exposure to Islamophobic Internet hate speech. There was evidence that the terrorist had contact via Facebook and email with far-

reports have indicated that the radicalization of young Muslims is "occurring more quickly, more widely, and more anonymously in the Internet age, raising the likelihood of surprise attacks by unknown groups whose members and supporters may be difficult to pinpoint.") (citing a National Intelligence Estimate).

³¹ See Danielle Keats Citron, *Cyber Civil Rights*, 89 B.U. L. Rev 61, 81-82 (2009).

³² *Id.*

³³ Lidsky, 44 Tex. Tech. L. Rev. at 149.

³⁴ [Viral Hate, p. 14]

³⁵ Citron & Norton, 91 B.U. L. Rev. at 1448.

right extremist groups, and he had gained inspiration from Internet hate websites.³⁶

- In January 2009, a lone white supremacist raped and shot one West African woman and shot and killed her sister in Brockton, MA. Investigators say that the suspect told arresting officers about white supremacist Internet sites and that the "people on these sites spoke the truth about the demise of the white race."³⁷
- In 1999, Mathew Williams and his brother murdered a gay couple and set fire to three Sacramento-area synagogues. Williams had frequented radical-right websites since his days as a student at University of Idaho, and, while investigating the crimes, police discovered boxes of hate literature at his home.³⁸
- In 2004, a Pittsburgh gunman murdered a Jewish woman, a Black man, two Asian-Americans, and two Indian men after visiting supremacist websites (such as <http://www.Stormfront.org>), joining the e-mail mailing list of the hate rock band *Aggressive Force* and downloading material from hate sites created by members of the neo-Nazi National Alliance.³⁹
- In November 2009, Nidal Malik Hasan, an army psychiatrist, opened fire at the Fort Hood soldier readiness center, killing thirteen and injuring 32. According to the authorities, Hasan had visited websites "espousing radical Islamist ideas."⁴⁰

In addition to merely inspiring individuals to commit violence against targeted groups, many Internet sites are actually devoted to detailing specific instructions on how to construct and detonate bombs.⁴¹ Versions of books and online books,

³⁶ Sindree Bangstad Testimony presented at the Hearing of the Inter-Parliamentary Coalition for Combatting Anti-Semitism Task Force on Internet Hate (Oct. 27, 2011).

³⁷ See *White Supremacist Shooting Spree Leaves Bloody Trail in Massachusetts*, Anti-Defamation League (January 2009), available at http://archive.adl.org/learn/extremism_in_the_news/White_Supremacy/Keith+Luke.htm.

³⁸ See Lorraine Tiven, *Hate on the Internet: A Response Guide for Educators and Families* (December 2003), available at http://www.partnersagainsthate.org/publications/hoi_full.pdf.

³⁹ See Christopher Wolf, *Needed: Diagnostic Tools to Gauge the Full Effect of Online Anti-Semitism and Hate*, OSCE Meeting on the Relationship Between Racist, Xenophobic and Anti-Semitic Propaganda on the Internet and Hate Crimes (Paris, June 16, 2004), at 5.

⁴⁰ See *Fort Hood and the Growing Muslim Extremist Threat*, ADL (Dec. 28, 2009), available at http://archive.adl.org/main_Terrorism/fort_hood_shooting.htm.

⁴¹ See the now-dated *Bomb-Making Manuals: Explosive Content*, ADL, *Poisoning the Web: Hatred Online*, available at http://archive.adl.org/poisoning_web/bomb_making.asp.

like the *Anarchist's Cookbook* and the *Terrorist Handbook*, circulate all over the Internet. According to United States Department of Justice reports, when investigating bombings or attempted bombings, Federal agents often find bomb-making literature that suspects had obtained from the Internet. One example is when a British neo-Nazi planted nail bombs in a Black neighborhood, an Indian area and a gay pub, killing three and injuring more than a hundred people.⁴² He had learned how to build his bombs by downloading from the Internet *The Terrorist Handbook* and *How to Make Bombs: Book Two*. Indeed, the Boston Marathon Bombers also used online instruction.⁴³

Even without causing any immediate violence, Internet hate can be extremely dangerous. Over time, the steady proliferation of hateful content on the Internet serves to desensitize the overwhelmingly well-meaning public and lulls the general population into tolerating this propaganda.⁴⁴ Long and sustained exposure to Internet hate speech can lead to ideological shifts in the general mainstream population.⁴⁵ This is especially true because young people are disproportionate users of the Internet, especially social media, and they are particularly vulnerable when it comes to being influenced by misinformation and conspiracy theories.⁴⁶

Targeted groups

The Task Force heard testimony that virtually every minority has been targeted by Internet hate sites, with particularly dangerous and harmful results. Some examples are set forth immediately below:

African Americans and Other Non-whites

Internet hate directed against African Americans dates back to the beginning of the Internet. With the Internet, white nationalists can reach a wider audience and recruit new members more effectively. White nationalist Don Black, founder of the racist Stormfront website, has commented that “The Internet is that

⁴² See Christopher Wolf, *Needed: Diagnostic Tools to Gauge the Full Effect of Online Anti-Semitism and Hate*, OSCE Meeting on the Relationship Between Racist, Xenophobic and Anti-Semitic Propaganda on the Internet and Hate Crimes, at 5.

⁴³ Anti-Defamation League “Inspire Magazine a Staple of Domestic Terrorism,” <http://blog.adl.org/extremism/inspire-magazine-a-staple-of-domestic>, April 22, 2013.

⁴⁴ See Alexander Tsisis, *Destructive Messages: How Hate Speech Paves the Way for Harmful Social Movements* (N.Y. Univ. Press 2002), at 5 (“Propaganda is essential for eliciting widespread cultural acceptance of exclusionary and supremacist ideologies.”).

⁴⁵ Bangstad testimony.

⁴⁶ Jamie Bartlett, Testimony presented at the Hearing of the Inter-Parliamentary Coalition for Combatting Anti-Semitism`1 Task Force on Internet Hate (Oct. 27, 2011), at 2.

opportunity we've been looking for...We never were able to reach the audience that we can now so easily and inexpensively." Black and other racist leaders see the Internet as a way to "burnish the reputation" of the KKK.⁴⁷

The Internet is also used to deliberately mislead young people—for instance, another site operated by Don Black, [Martinlutherking.org](http://martinlutherking.org), appears to be a legitimate documentary-style site, but rather than containing historical information it contains defamatory material about the civil-rights leader.⁴⁸

Some of the anti-Black sites are also anti-immigration (especially Latino) and anti-Muslim.⁴⁹

Jews

There are thousands of hate sites on the Internet, many of them targeting Jews.⁵⁰ While anti-Semitism is obviously not new, the Internet provides the opportunity for instant coordination and planning of real-world violence.⁵¹ For instance, following the Israeli operation in Gaza in 2009, Jews were threatened and beaten on the streets, and synagogues were firebombed all over Europe and South America, in large part due to the Internet's ability to quickly mobilize perpetrators.⁵²

Holocaust denial is another form of anti-Semitism that is very prevalent on the Internet, and, like other forms of hate speech, can lead to physical violence.⁵³ The shooter who opened fire on the U.S. Holocaust Museum in Washington, D.C. in 2009, ran a hate website on which he posted essays denying the Holocaust and had a long history of associations with neo-Nazis and holocaust deniers and in his self-published e-book, "Kill the Best Gentiles," he railed against a Jewish conspiracy to "destroy the white gene pool."⁵⁴ Perhaps even more troubling,

⁴⁷ See Raphael Cohen-Almagor, *Fighting Hate and Bigotry on the Internet*, Policy & Internet Vol. 3: Iss. 3, Article 6 (2011), at 6.

⁴⁸ *Id.* at 7; see also <http://martinlutherking.org/>.

⁴⁹ *Id.*; see, e.g. National Socialist Movement, <http://www.nsm88.org/>; The Creativity Movement, <http://creativitymovement.net/>.

⁵⁰ See *Hate Crimes Against Jews*, Civilrights.org, available at <http://www.civilrights.org/publications/hatecrimes/jews.html>.

⁵¹ See Christopher Wolf, *The Internet is Making Anti-Semitism Socially Acceptable*, Remarks To the London Conference on Combating Anti-Semitism Houses of Parliament, London (Feb. 16, 2009), available at <http://www.adl.org/combating-hate/cyber-safety/c/the-Internet-is-making.html>.

⁵² See *id.*

⁵³ See Raphael Cohen-Almagor, *Holocaust Denial is a Form of Hate Speech*, Amsterdam Law Forum Vol. 2:1, at 33 (2009).

⁵⁴ Richard Sisk & Richard Shapiro, "'Obama created by Jews': Holocaust Museum shooting suspect

Holocaust deniers seek to demonize Jews as the aggressors and fabricators of an elaborate myth, with their ultimate purpose to legitimize another Holocaust against Jews.⁵⁵ This kind of hate speech, even if originally found only on the Internet, does not always stay on the fringe, but can be adopted and incorporated by more mainstream movements.⁵⁶ For example, many anti-Semitic hate groups in the United States publish anti-Israel and anti-Zionist materials with anti-Semitic overtones online, and in some cases these materials have been adopted by rallies of groups holding anti-Israel (but not necessarily anti-Semitic) views. This adoption of anti-Semitic images and ideas has led to an increase in expressions of anti-Semitism in forums and events targeting Israeli political decisions.

LGBT

Particularly devastating real-world effects can be felt in the arena of Internet hate speech in the lesbian, gay, bisexual, and transgender (LGBT) community.⁵⁷ Much of the harassment centers around “outing” someone’s sexuality or gender identity, with a secondary level of harassment—often threats of violence or actual violence—resulting. The outing over the Internet is particularly hateful because the communication spreads quickly and widely and therefore allows the secondary victimization to occur quickly. Right-wing extremist propaganda directed against the LGBT community on the Internet can also create a feeling of uneasiness among that community.⁵⁸

James von Brunn's chilling racist note". *New York Daily News*, 6/11/2006..

⁵⁵ See Cohen-Almagor, *Holocaust Denial is a Form of Hate Speech*, Amsterdam Law Forum, at 36.

⁵⁶ See *Antisemitism on the Internet: An Overview*, International Network Against Cyberhate, OSCE Conference on Anti-Semitism (Berlin 2004), at 17.

⁵⁷ See Deborah Gold, *Internet-based Hate Targeted at Lesbian, Gay, Bisexual, and Trans People Briefing Document*, Testimony presented at the Hearing of the Inter-Parliamentary Coalition for Combatting Anti-Semitism Task Force on Internet Hate (Oct. 27, 2011).

⁵⁸ See Suzette Bronkhorst & Ronald Eissens, Eds. *Hate on the Net: Virtual Nursery for In Real Life Crime*, International Network Against Cyberhate (June 2004), at 29-30 (noting that after anti-gay rhetoric increased on the Internet, one café frequented by homosexuals was set on fire and another had its address posted on a neo-Nazi website, indicating a threat).

Women

Similar harassment, with similarly unfortunate effects, occurs within the arena of misogynistic Internet hate speech.⁵⁹ Internet hate is disproportionately directed against women and can include sexual threats, privacy invasions, technical sabotage, photographs, and defamation. It can also include “Google bombing,” where cyber haters push a site’s threads to the first position on a search engine. This kind of cyber harassment can lead to real physical violence, or at the very least ruin women’s lives by shaking their personal safety and causing them to alter their careers and other life decisions or, in some extreme cases, commit suicide.⁶⁰ It can also keep women from being able to participate actively online, thereby depriving women of access to the benefits of technology. For example, in 2007, anonymous posters began targeting Kathy Sierra, a technology blogger, by threatening violence against her and stating that she deserved to be raped. Fearing for her safety, Ms. Sierra canceled speaking engagements and stayed away from public life. A recent *New York Times* article highlighted the dangers of the misogynist Internet subculture with a discussion of the case of Gilberto Vallo.⁶¹ Vallo “inhabited an Internet subculture populated by men who think about raping and torturing women, roasting them on spits and eating them,” which caused him to plot “actual crimes that would begin with abduction and result in the cannibalization of female victims,” including his wife. Valle’s wife testified at his trial that she felt terrorized.⁶²

Misogynist cyber harassment can lead to real physical violence, or at the very least ruin women’s lives by shaking their personal safety and causing them to alter their careers and other life decisions or, in some extreme cases, commit suicide.¹ It can also keep women from being able to actively participate online, thereby depriving women of access to the benefits of technology.

■ **Professor
Danielle Citron**

⁵⁹ See Danielle Keats Citron, *Misogynistic Cyber Hate Speech*, Testimony presented at the Hearing of the Inter-Parliamentary Coalition for Combatting Antisemitism Task Force on Internet Hate (Oct. 27, 2011).

⁶⁰ see also Nathan Hall, The Submission of The Cross-Government Hate Crime Independent Advisory Group to the Inter-Parliamentary Coalition for Combating Anti-Semitism Task Force on Internet Hate (Oct. 27, 2011), at 3 (“[O]ur evidence shows that direct targets of Internet hate report significant and ongoing feelings of fear, threat, and intimidation”).

⁶¹ Citron, citing, Ginia Bellafante, *Remember Misogyny?* The New York Times (Mar. 1 2003), available at <http://www.nytimes.com/2013/03/03/nyregion/remember-misogyny-its-on-the-web.html>.

⁶² *Id.*

Other forms of Internet misogyny, such as pro-rape groups and groups called *Hitting Women* hosted on Facebook, can also be dangerous. Citron notes that these forms of Internet hate create a climate of fear, intimidation, and subordination.

Ethnic Minorities

Nationalist motivations can also underpin harassment that begins online but then translates to the real world. In Russia, for example, a nationalist blogger published names and contact details of students from the Caucasus attending Russia's top universities and attached inflammatory videos.⁶³ Other blogs reposted the story, creating a dangerous situation for the students who were targeted.

Hate Speech Researchers and Academics

Another less obvious group targeted by Internet haters—but with equally dangerous harmful effect—is researchers of Internet hate. Researchers often publish their work online, which means that insulting and threatening comments are often left on the Internet. For example, after a recording of an interview with Internet hate researcher Jamie Bartlett was made available on YouTube, comments left on the site included such vitriolic sentiments as “jamie bartlett's head will be taken by the public when we overthrow the nest of poisonous vipers. I look forward to the day when you beg for mercy Bartlett.” These kinds of threats can chill research in this area as researchers begin to feel threatened and abandon their work out of their own personal safety.⁶⁴

Other Targets

Extremists in support of, or against, any cause can use the Internet to galvanize individuals to take harmful and dangerous action. For example, *The Nuremberg Files*, a virulently anti-abortion website that was later shut down, published the names and detailed personal information of abortion providers.⁶⁵ The list of abortion providers read like a list of targets for assassination, with names listed in black “still working,” those in grey “wounded” and those with a line through their names “fatality,” or “murdered.” Indeed, when abortion providers on the list

⁶³ See *The Brave New World of E-hatred: Social Networks and Video-sharing Sites Don't Always Bring People Closer Together*, *The Economist* (Jul. 24, 2008), available at <http://www.economist.com/node/11792535/print>.

⁶⁴ Testimony of Jamie Bartlett, *Supra*.

⁶⁵ See *Anti-Abortion Extremism in Cyberspace: The Creator's Rights Party, ADL, Poisoning the Web: Hatred Online*, available at http://archive.adl.org/poisoning_web/anti_abortion.asp.

were killed, they were crossed off the list—as when the name of Dr. Barnett Slepian, a Buffalo abortion-provider murdered in his home by a sniper, was crossed off hours after he was murdered. Similarly, in 2010, a blogger posted on his blog that three federal judges “deserved to be killed” because of their decisions regarding a handgun ban.⁶⁶ The blog post, which included the judges’ photographs, phone numbers and work addresses, led to the blogger’s conviction of threatening to assault or murder the judges.

As Lidsky explained:

“[The blogger’s] ambiguous call for violence against three federal judges capitalized on the size, anonymity, and ideology of his blog’s audience, knowing that telling an unknown number of committed radicals that the judges deserved to die and providing the addresses necessary to track them down and murder them would generate terror without Turner himself having to lift a finger other than to type on his keyboard.”⁶⁷

⁶⁶ See Lidsky, 44 Tex. Tech. L. Rev at 157

⁶⁷ *Id.*, at 163.

PART TWO: RESPONDING TO CYBERHATE

Turning to the second part of its work to identify and recommend responses to the problem of cyberhate, the Task Force first examined legal responses to cyberhate. The Task Force then took a very unique approach: It sought to address responses a collaborative rather than a confrontational view. That is, recognizing that intermediaries are often excluded from the discussion on Internet hate, the Task Force specifically sought to include intermediaries in its work.

That approach yielded significant fruit, including the critical conclusion that responding to online hate, even by the most sympathetic and willing intermediaries, is extremely difficult to fight. These are not excuses; rather, they are realities that any program that is to successfully combat cyberhate must address.

Limits to the law

The Task force considered the legal questions inherent in governing cyberhate and notes that conflicting legal regimes make it very difficult for intermediaries *and* users to know their rights—and makes it very impractical for legal regimes to resolve issues of cyberhate.

Online hate occurs in a global context and is thus subject to laws of different countries. Law exists on a spectrum; at one end, the U.S. offers strong jurisprudential protection of speech with a few narrow, historically recognized exceptions such as true threats, incitement of imminent violence, intentional infliction of emotional distress, obscenity, child pornography, etc. Other countries are situated in the middle, regulating hate speech such as bigoted harassment or incitement to violence. On the other end of the spectrum, some countries, such as India, embrace a capacious definition of hate speech, such as promoting hatred, and prohibiting it via criminal or civil law.

The reason why it is important to understand the American free speech tradition is that the...executives who regulate speech are marinated in the American free speech tradition; they embrace it and accept it. In applying community guidelines ...which allow the regulation of hate speech and prohibit speech that promotes terrorism these executives are actually construing this through the narrow lens that there has to be some threat of imminent violence for the speech to be suppressed.

■ **Professor Jeff Rosen**

However, the American perspective is particularly relevant both jurisdictionally and philosophically.

Philosophically, as Professor Rosen testified to the Task Force, the American tradition dominates executive-level thinking at the major intermediaries. He said:

The reason why it is important to understand the American free speech tradition is that the...executives who regulate speech are marinated in the American free speech tradition; they embrace it and accept it. In applying community guidelines...which allow the regulation of hate speech and prohibit speech that promotes terrorism these executives are actually construing this through the narrow lens that there has to be some threat of imminent violence for the speech to be suppressed.

Moreover, the task force heard testimony that because most online hate content is stored on U.S. servers, the American legal perspective is particularly influential in the arena. U.S. law insulates most providers from liability for much of the content that passes over their services—much in the same way it protects telephone companies from liability for crimes committed over their wires. Specifically, the Telecommunications Act of 1996 (which applies to Internet companies) states:

No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider. 18 U.S.C. 230 § (c)(1).

Commenting on this provision, Leslie Harris of the Center for Democracy and Technology (CDT) said:

The interesting thing about our laws section 230 they are intended to do two things. It's an elegant bargain. We did not want those sites like Facebook and Google, to find themselves in constant lawsuits for content that was posted by third party users. Part of our law was intended to provide protection for the intermediaries. The other part of the law was to protect them against liability for actions they took voluntarily, the Good Samaritan, so they could take action without risks of liability on the other side.

Despite the latitude in United States law, an intermediary must not knowingly provide material support for terrorists by providing a communications gateway for their activity. 18 U.S.C. §§ 2339A, 2339B. This American view of the law reflects the position held by many U.S. individuals and courts (and many US

institutions and intermediaries) that the right to freedom of expression is vested in the individual but should be interfered with where interests are most compelling.

Nonetheless, Internet intermediaries face demands from non-U.S. countries to remove hate speech that would not be subject to regulation in the U.S. As the case against Google's executives in Italy makes clear, companies need to comply with those requests or face criminal prosecution. Importantly, intermediaries face conflicting directions on several fronts. First, U.S. law advises that they can make voluntary choices about the content they host as Good Samaritans yet the U.N. Special Report on Free Speech and Human Rights provides that only governments have the right to decide what content must be taken down. Second, laws of many nations pull intermediaries in opposite directions, with one country prohibiting particular content as hate speech, such as France banning Holocaust denial, and the other, such as the U.S., providing broad protection from government regulation for political speech no matter how distasteful or offensive. Third, intermediaries sometimes face unreasonable demands to remove speech, demands that are antithetical to their values and to the policies underlying other countries' laws. As a recent blog post at CDT succinctly puts it:

Moreover, many governments have enlisted, or are considering proposals that would enlist, Internet intermediaries – search engines, social networks, ISPs – in controlling expressive content. By holding intermediaries responsible for content that they did not create, governments seek to deputize these companies into monitoring their networks and censoring their users. Placing liability on intermediaries often stifles legitimate speech: Fearful of liability, intermediaries often over-react, blocking or taking down even lawful content. In addition, placing liability on intermediaries can discourage them from offering of innovative new services that would expand opportunities for economic activity and personal or community development.⁶⁸

Indeed, increasing pressure is coming from the European Union that threatens this dominance of the US view. Professor Rosen writes in the New Republic:

Unless Google, Facebook, Twitter, and other Internet giants draw a hard line on free speech, they will find it more difficult to resist European efforts

⁶⁸ See Erica Newland, *Shielding the Messengers: Global Threats to Free Expression*, Center for Democracy and Technology (April 2012), available at <https://www.cdt.org/blogs/erica-newland/0304shielding-messengers-global-threats-free-expression>

to transform them from neutral platforms to censors-in-chief for the entire globe. Along with tougher rules on hate speech, the European regulators are weighing a sweeping new privacy right called “the right to be forgotten.” If adopted, it would allow users to demand the deletion from the Internet of photos they’ve posted themselves but come to regret—as well as photos of them that have been widely shared by others and even truthful but embarrassing blog comments others have posted about them. The onus would be on Google or Facebook or Yahoo or Twitter to take down the material as soon as a user makes the request or make the bet that a European privacy commissioner—to whom requests could be appealed—would determine that keeping the material online serves the public interest or provides journalistic, literary, or scientific value. If the companies guess wrong, they could be liable in each case for up to 2 percent of their annual incomes. A European Commission press officer stresses that each member country would choose how to implement the penalties, but for Google, the fines could hit \$1 billion per incident.⁶⁹

While there is not enough space within this report to discuss how non-US countries deal with cyberhate, it is clear that these perspectives are increasing in influence. For a detailed look at various national laws against cyberhate, and for more information about international cyberhate, please see the Web site of the International Network Against CyberHate (INACH).⁷⁰

However, at the end of the day, the American view rejects the idea that the community has a stake and rejects the position that free expression can be limited where it conflicts with larger democratic ideals. Understanding this rejection is a key to understanding the dominant American perspective on hate online—and key to understanding the politics and policies behind the decision making of many American intermediaries. It is here that Americans often cite United States Supreme Court Justice Louis Brandeis’ statement that:

⁶⁹ Jeffrey Rosen, “The Delete Squad: Google, Twitter, Facebook and the new global battle over the future of free speech,” *The New Republic*, April 29, 2013.

⁷⁰ INACH’s legislation page can be found here: <http://www.inach.net/legislation.html>. Please also see the Center on Technology And Democracy’s Intermediary Liability reports, found here: <https://www.cdt.org/toolkit/intermediary-liability>

Publicity is justly commended as a remedy for social and industrial diseases. Sunlight is said to be the best of disinfectants; electric light the most efficient policeman.⁷¹

Limits to the American View

While sacrosanct, rights to speech in the U.S. are not absolutely inviolate, and even the individualistic view has important contours. The U.S. Constitution's First Amendment applies to government entities only, and it does not require private actors to make or allow speech they do not wish to. An intermediary can post or take down what it wishes within the limits of the private contractual agreements it may have with users. Of course, in the U.S., free speech does not prevent access to courts to prosecute private lawsuits over issues such as defamation and workplace- and school-based harassment and it does not permit expression that would be criminal in other forms, such as child pornography or treason. Finally, as noted above, free speech does not permit individuals to provide material aid to terrorist organizations, which may include publishing training material or providing web hosting services or incitement to imminent lawlessness and on speech that presents other clear and present dangers to law and order.

Limits to other views

Indeed, even if one looks toward non-American contexts for legal guidance, the landscape is murky. For instance, one must ask if there an inherent conflict between Article 9 (the freedom of thought, conscience and religion) and Article 10 (freedom of expression) of the European Convention on Human Rights (ECHR). European case law strongly supports peoples' rights to express themselves so long as it does not contravene another law or right under the ECHR; often the line is very grey and difficult to interpret and judge, as the laws criminalizing hate speech differ in every country. What can be done to preserve peoples' right to freedom of expression when that freedom results in hate and violence or spreads potentially false, libelous and defamatory information? Case law at both a national and European level have repeatedly concluded that while hate speech may be immoral and wrong, this alone cannot be criminalized as it conflicts with one's basic civil liberties and rights to freedom of expression, particularly if such speech does not incite violence or result in other illegal activity criminal or otherwise.

With this in mind, and to turn at last to the purely practical, the Task Force heard testimony concerning the reality that U.S. based servers will continue to keep

⁷¹ L. Brandeis, *Other People's Money* 92 (Stokes Publishing 1914).

serving up hate speech. Cross border litigation is expensive, has very limited impact, and is also likely to be ineffective. As said, the vast majority of Internet hate resides in the U.S. on U.S. servers. While an American citizen who posts material on the Internet that is illegal in another country could be prosecuted in that country if he/she subjects him/herself to the jurisdiction of that country, a person within the United States will likely not be extradited to a foreign country based on Internet speech, as most contemporary treaties setting out the duties of the signatory countries require that the extraditable conduct be criminal and punishable as a felony in both countries. Thus, the United States will often not extradite a person for engaging in a constitutionally protected activity even if that activity violates a criminal law elsewhere.

Challenges Facing Intermediaries

Specifically, the Task Force heard testimony that intermediaries face a number of key challenges, some of which present significant obstacles to dealing with cyberhate:

- **Challenge of scale (the amount of materials).** There are vast quantities of materials on most intermediaries' pages—amounts that are so vast that it is simply impossible to proactively police it all while maintaining a business model that will allow companies to enter or stay in the online market. Recent statistics posted by the companies show that Facebook has 1.11 billion monthly active users as of March 2013, there are 72 hours of video uploaded to YouTube every minute, and Twitter has approximately 400 million tweets each day. The vast amount of data does not excuse action, but it must inform policy and business decision-makers at every turn.
- **Challenge of knowledge.** It is not always obvious to an intermediary or its employees what constitutes hate speech—and how to adjudicate amongst competing claims. Furthermore, the problem of scale (see above) and the volume of complaints received means that reviewers may have mere seconds to make a historical, contextual, political and social determination about the message before them. Policy and business leaders have rarely been able to identify such kinds of speech and when they do it, it is either very specific kinds of speech (e.g., holocaust denial) or it is very general and can be significantly overreaching.⁷²

⁷² This is not to say that hate speech is free from definition. Quite the contrary, certain kinds of speech fall four-square into the ambit of hate speech. For instance, in the area of anti-Semitism, comments that deny the holocaust, that Jews are cheap, that Jews control banks or the media are unquestionably anti-Semitic. The question of who intermediaries trust to help them see these types of hate speech and who they trust to make “hard calls” feed into this problem of knowledge.

- **Challenge of politics.** Accusations of hate are rampant between competing political groups—and often very unclear to the reviewing intermediary. Again, policy and business leaders may find it impossible to identify and discern amongst the competing claims in this arena.
- **Challenge of terms of service.** Given the foregoing, terms of service and community guidelines are very difficult to write and manage—and legislative responses even more difficult. And they give rise to key questions: Are they adequate to ensure that there is a test to determine whether a particular word/concept, as applied, violates the terms of service. How is hate applied? Does it target words directed at an individual only? Or calls to violence against a group or a people? In this area standardization and intercompany working groups may be helpful.
- **Challenge of process and access.** Although steadily improving amongst the bigger players, the reporting mechanisms for offended parties are rarely sufficient, in terms of process or access. In addition to navigating difficult interfaces, users need to be instructed in how to efficiently file reports of hateful content that are cogent and understandable to the intermediary reviewing them.
- **Challenge of consistency.** Even if willing to do so, intermediaries are rarely in a position to coherently and consistently describe their decisions with regard to cyberhate. This unmoors the process from any sense of *stare decisis*, provides no guidance as to acceptable behavior and leaves users wondering about their filed reports (see below). Compiling “law reports” of decisions seems like one solution to this problem: publicly announcing at least some representative decisions will allow intra- and inter-site comparisons of how hate is treated and what standards apply.
- **Challenge of transparency.** The sheer volume of complaints means that it is very difficult for intermediaries to respond to most complaints filed by users—and will leave those users frustrated that their complaints are not taken seriously. In short, to most users, the experience of filing hate-related reports is a “black box” process with no clear output.
- **Challenge of user and intermediary education.** Given the difficulty of knowing what counts as hate speech that violates an intermediary’s terms of service, it seems very difficult to educate users and the intermediary’s own staff as to the nature and extent of hate speech.

PART THREE: SOLUTIONS

Role of Education

While some countries endorse organizations that combat hate speech, most countries' education departments do not explicitly include lessons on countering hate speech in their curriculum guidelines. In addition, most countries do not yet have national laws in place that require schools to educate youth on how to counter hate speech. Three examples:

- *United Kingdom.* The UK Department of Education explicitly endorses the UK Council for Child Internet Safety (UKCCIS).⁷³ UKCCIS is a voluntary organisation that develops and promotes effective tools and information for children and parents in order to protect children from risks including cyberbullying, harmful content such as pro-anorexia or suicidal information, sexual images, grooming, loss of privacy and scams. It is jointly chaired by Edward Timpson MP, Parliamentary Under-Secretary of State for Children and Families; Jeremy Browne MP, Minister of State for Crime Prevention; and Ed Vaizey, MP, Parliamentary Under-Secretary of State for Culture, Communications and the Creative Industries. Although this endorsement demonstrates that the UK government is aware of the importance of protecting children from the dangers of the Internet, the government has yet to institute an explicit policy that requires educating children about Internet hate.
- *Australia.* The position in Australia reflects the trend of having voluntary organisations primarily deal with issue of educating children on Internet hate. Here, the B'nai B'rith Anti-Discrimination Commission, a human rights arm of B'nai B'rith Australia/New Zealand, has launched *Click Against Hate*.⁷⁴ This program aims to provide teachers and students with the necessary skills to identify and respond to anti-Semitism and other forms of racism they encounter online. It offers workshops and courses for both students and teachers, providing an educational framework for responding to anti-Semitism and racism online. Developed by social media expert, Dr Andre Oboler, initial workshops were being held at Jewish day schools in Melbourne as of 2010. These workshops serve as a

⁷³ <http://www.education.gov.uk/ukccis>

⁷⁴ <http://www.antidef.org.au/click-against-hate/w1/i1011350/>

useful template for future educational programs concerning Internet hate.

- *Germany.* One of the most advanced countries in terms of concrete, explicit government support for combatting Internet hate in the classroom is Germany. The German government finances the organisation *Jugendschutz*⁷⁵ to investigate and counter online hate speech with education. Setting up and financing an organisation is a helpful and proactive move taken by the German government that symbolises the final hurdle before governments and departments of education finally inscribe the educational countering of Internet hate in their policies and curriculums.

Education and Protecting Civil Liberties

Countering cyberhate in educational environments has raised significant civil liberties issues. Schools are often required to install computer programs and software that blocks hate sites and encourage parents to introduce similar programs at home. In addition, some schools and universities adopt codes of practice or policies that prohibit speech that offends any group based on race, gender, ethnicity, religion sexual orientation, or other fundamental identities. However, these methods are not necessarily a viable solution. Children can easily access the Internet in unregulated environment, such as at a friend's house or an Internet café. And many would argue that censorship is not the correct solution as it is contrary to individuals' right to free speech. It is imperative that children are taught to understand that online information is not necessarily right or accurate. The key issue is whether more speech, not less, is the best tactic to combat hate speech.

The Task Force heard testimony that education systems can play a pivotal role in countering Internet hate speech by educating young people on how to *identify* Internet hate speech. Educating youth to identify Internet hate, the perils of it and the impact it has, can act as a deterrent against promulgating it further.

Studies have shown that much of today's youth do not appreciate and understand the boundaries and differences between reputable and inaccurate information online. For example, most newspapers and magazine in the UK sign up to the voluntary Editors' Code of Practice published by the Press Complaints Commission.⁷⁶ This Code requires press providers (such as online news websites)

⁷⁵ <http://www.inach.net/content/Annual%20Report%20jugendschutz.pdf>

⁷⁶ Press Complaints Commission, Code of Practice:
http://www.pcc.org.uk/assets/111/Code_of_Practice_2011_A4.pdf

to maintain high professional standards where all signatories promote common interests such as avoiding discrimination and prejudicial references regarding race, religion, gender, sexual orientation, and disability. Children that can easily identify reliable websites may be more inclined to avoid potentially harmful websites, including those that spread hate speech.

Once children can identify hate speech, they can be empowered to combat it. One suggested method for encouraging dialogue is to introduce 'cyber space walks'⁷⁷, a means of presenting conflicting views on various issues on the same website. For example, if a website maintains or supports certain opinions they deem to be factually based, opposing views could also be presented, or at least links to them, to enable the reader to independently opine. To the extent that it is possible to contribute opinions and opposing views on a topic, this should be encouraged in reputable forums. Websites such as Facebook also promote initiatives to combat hate and host different forums to discuss the hate challenge, such as *United Against Hatred* and *United Against Hate*.

It is imperative to educate children to be able to identify reputable sources and distinguish them from non-reputable websites. It should also be encouraged to educate people to research differing views on the same topic to ensure they obtain a well-balanced and accurate understanding. Counter-speech can then be used after teaching that it should be supported by fact and research and only posted on a reliable source, to better enable readers to form their own views.

⁷⁷ Adam Thierer, "*Do We Need a Ministry of Truth?*", Forbes Magazine Online, (January 2012) available at <http://www.forbes.com/sites/adamthierer/2012/01/29/do-we-need-a-ministry-of-truth-for-the-internet/>

AVENUE FOR FUTURE EXPLORATION

One area for future exploration is the intersection of anonymity and hate speech. The task Force heard that people who are able to post anonymously (or pseudonymously) are far more likely to engage in cyberhate. Anonymously posted hate speech can extend from bigoted posts in online forums to hate-filled and inflammatory comments appended to online newspaper articles.

As previously stated:

The Internet, in large part because of the shield of online anonymity, has become the medium through which hate groups plot and promote real-world violence, recruit and indoctrinate like-minded haters, and mislead and distort information for those—like students—who innocently link to their content. There are, of course, notorious hate mongers who use their real identities and revel in the limelight. But the vast majority of hate spewed online is done so anonymously. The Internet content of hate mongers—words, videos, music, and social network postings—serve to offend the human dignity of the intended victims, minorities, and those who hate groups identify as “the other.” The Chief Commissioner of the Canadian Human Rights Commission, Jennifer Lynch, recently commented: “Freedom of expression is a fundamental right...[s]o is the right to be treated with equality, dignity and respect.” The balance between free expression and the right to human dignity is way out of whack online. The Internet has become the launching pad for mean-spirited, hateful, and harmful attacks on people.⁷⁸

One event in recent history illustrates this point: the Madoff scandal. After the Madoff scandal broke, hundreds of people posted viciously anti-Semitic comments on newspaper web sites such as the Palm Beach Post and on financial-oriented websites such as Yahoo! Finance. For instance, using screen names that hid their identity, comments such as the following were posted on Yahoo! Finance:

- "This is what happens when you let Jews run amok in a country for too long."
- "You are guaranteed to get screwed when trusting a JEW."

⁷⁸ Christopher Wolf, “Accountability For Online Hate Speech: What Are The Lessons From “Unmasking” Laws?,” 87 *Denver Law Review Online* 30 (2010).

- "He is a jew, remember GODS CHOSEN PEOPLE , so he gets away with s**t cuz the f*****g dumb evangelicals are totally brainwashed !"
- "Jews should be banned from working on Wall Street. All they do is steal our money."

The *Palm Beach Post* articles relating to the Madoff scandal had comments such as these:

- Just another jew money changer thief. It's been happening for 3,000 years. Trust a jew and this is what will happen. History has proven it over and over and over. Jews have only one god - money."
- "Jews ripping off Jews, the economy must really be tough. The gov't will probably bail out these loses like it does to all other Jewish loses."

Notably, the Palm Beach Post shut down its user comment section in response to this hate speech.

At the same time, anonymity has many virtues and supporters: It allows for unfettered testing of ideas in the free marketplace of ideas and it permits the expression of minority/disfavored viewpoints free from fear of discovery by governments, social contacts, bosses and others. Credited for empowering the Arab Spring and empowering corporate and government whistleblowers, anonymity also has played a traditional role in United States politics, such as with the great U.S. Revolutionary War writer Thomas Paine and the authors of *The Federalist Papers* (Alexander Hamilton, James Madison, and John Jay), who chose to remain anonymous with their publications.

SOME INNOVATIONS IN COMBATING CYBERHATE

This report has identified the numerous challenges facing intermediaries in combating cyberhate. However, a number of intermediaries have taken interesting and creative steps in that fight worth setting out here. While much more can and must be done—the Task Force's recommendations show several critical paths—it is important to identify some public-facing responses to cyberhate by intermediaries.

Counter-speech in the Corporate Voice

One approach has been to use an intermediary's corporate voice where a platform is used to promote cyberhate. For instance, rather than tinker with its sacrosanct Page Rankings, Google responded to searches for the term "Jew" and the high-ranking offensive results it returned by using its own message and voice: it used its "Keyword" advertising program to serve up an advertisement called "An Explanation of Search Results." The advertisement—now earning a high Page Rank in its own right—disclaims association with the offensive results and offers users additional resources.

Intermediaries, especially those with advertising capabilities, can and should speak in their own voices to counter offensive speech or to at least point readers to meaningful counter-information elsewhere on the web. This is perfectly consistent with a rich "American View" of free speech—a corporate voice, like any other, and (and with cyberhate it has enabled, should) participate in the marketplace of ideas. No serious theory supporting the promotion of free speech requires a platform provider to maintain neutrality and silence in the face of hate. In fact, the very opposite may be true. The power of an intermediary to create an environment where, for instance, misogyny or homophobia are met with an indication that the intermediary disagrees (and provides both corrective information and an avenue to voice complaints) will empower more users to lend their voices and perspectives to the web. And if those intermediary voices become the target of discussion—all the better for the marketplace of ideas. Thus, if Facebook wishes, for corporate reasons, to permit Holocaust denial pages, then it may consider serving ads on those pages that assure users that such offensive and anti-Semitic pages are not the view of Facebook or its employees and that resources exist to counter such pernicious views. Facebook's powerful voice can counter the hateful rhetoric and remind users that the company does not support bigotry. Indeed, online retailers, such as Amazon and Barnes and Noble, have allowed trusted NGOs to place

discussions and highest-ranking comments near books that are offensive (such as the anti-semitic forgery "The Protocols of the Learned Elders of Zion"), allowing them to live up to their own corporate ideals of selling a wide range of books and not acquiescing to hateful speech.

Of course, it must be said that even on the richest conception of free speech, nothing should stop companies from exercising their own free speech rights and refuse to carry hateful items or content in the first place.

Real World Identities, Anonymity, and Unmasking

Another technique, employed by Facebook in its new *Pages* product, is to allow anonymous users to change content that violated Facebook's Community Standards—and then threaten to take down content or expose their real names if they fail to comply. Apparently, this has been successful in removing hateful content.

This technique is a more specific version of a technique Facebook has used for a long time: requiring real identities to use most of Facebook's features. Newspapers, such as the New York Times, prioritize "real name" submissions to its comments sections, pushing anonymous content far down the page.

It is well-established that unmasking haters silences many hateful voices—bigots are often afraid of being known to their family, community, and coworkers as such. Thus, online public identification has stopped much cyberhate as anonymity is mask from behind which much hate is spewed. It is worth noting here that this is a double-edged sword—anonymity has been very important to political dissidents, religious minorities and radical thinkers.

Branding and Monetization

Another technique, used by Myspace to some success, is to inculcate a sense that, as a company, hate content materially interferes with monetization efforts and with its own corporate brand.

The idea here is the belief that efforts to sell ad space are hampered when there is the risk that an advertisement containing a carefully curated and cultivated brand identity is served up on a page containing cyberhate.

Relatedly, we infer that a reputation as a platform for cyberhate will damage a company's own prospects in the market (especially the IPO market) both because of the advertising issue mentioned above and because new customer acquisition may be hampered if the environment is perceived to be one that is

inaccessible to targeted groups (put positively, a platform that is seen in a welcoming light will acquire more users and thus increase valuation).

One hybrid example of the above is Godaddy, a leading online registrar and hosting service, has recently began removing anonymity from the registrations of certain especially pernicious sites. Godaddy recognizes that it need not provide the full benefit of its capabilities to support cyberhate—and that cybehate is business not worth having in the context of a multibillion dollar company with a very public and very expensive advertising campaign.

Again, this addresses only public-facing work and not the work companies do "behind the scenes." In short, there is much work to be done, but there are interesting examples of, and solid practical and philosophical reasons for, creative action.

CONCLUSION AND RECOMMENDATIONS

The problem of cyberhate is pervasive and, given the difficulty of responding to it, persistent.

However, it is clear from the testimony received that a number of factors make a legislative response to cyberhate both inappropriate and likely to fail:

- The core value and benefits of free speech
- The location of most hate content on U.S. servers
- The extreme difficulty of responding to cyberhate (even by willing intermediaries) including scale and definition
- The failure of cross-border law enforcement and civil actions to produce any meaningful change in the amount and intensity of cyberhate
- The ever-changing technology which makes cross-border law enforcement and civil actions significantly more difficult

Accordingly, the Task Force determined that continued work between NGOs, academics, and intermediaries would be the most meaningful way to approach this issue. Therefore, in order to expand upon the work started under the auspices of the Task Force, the task Force resolved as follows:

The ICCA Task Force on Internet Hate endorses the formation of the Anti-Cyberhate Working Group (ACWG) to include industry, academics, NGOs, and other interested parties to work together to build best practices for understanding, reporting upon and responding to Internet hate, and requests ICCA Task Force member Anti-Defamation League (ADL) to convene and move forward with the ACWG. The ACWG should meet regularly and provide reports to the public on the steps being taken by Internet companies to address Internet hate.

In guiding this group, the Task Force set out a number of recommendations which, while not endorsed by the entire Task Force, represent guidance to the ACWG derived from the testimony on how to approach cyberhate:

Principles for Responding to Internet Hate Speech

1. Create clear policies on hate speech and include them within the terms of service

Intermediaries' terms of service should clearly define prohibited speech on the website or service.

2. Create mechanisms for enforcing hate speech policies

While it is not possible for websites or social networking services to filter all of the content that is posted, intermediaries should put resources and mechanisms in place to monitor, respond to complaints about, and remove hate speech in a manner that preserves free expression and innovation while respecting the terms of service.

3. Establish a clear, user-friendly process for allowing users to report hate speech

Intermediaries should make it easy to utilize the collective efforts of their users to identify and limit hate speech. Intermediaries should have easy-to-find and easy-to-use reporting mechanisms that allow users to report hate speech that violates the terms of service. Further, intermediaries should have policies and procedures in place that allow for quick responses to these reports.

4. Increase transparency about terms of service enforcement decisions

Intermediaries should provide users with information about how decisions are made with respect to the removal or non-removal of hate speech that violates the terms of service. While it is probable that individualized responses to complaints would be difficult, intermediaries can increase transparency by publishing case studies and/or general examples of speech that has been deemed unacceptable under the terms of service.

5. Actively encourage counter-speech and education to address hate speech

Intermediaries should not remain silent about hate: they can use their own corporate voice (in press releases, advertisements that run next to the hate

content, or elsewhere) to respond to hateful content they choose not to take down. In addition, intermediaries should empower users to counter hate speech by including explanations and counter-speech where appropriate.

6. Unite industry to confront the issue of hate speech

Intermediaries should form an industry working group with the goal of meeting regularly to discuss and consult with each other on issues they face responding to Internet hate speech. On a yearly basis, this working group should publish a report detailing how intermediaries have been responding to hate speech and providing examples and case studies, which the larger Internet industry can use to combat hate speech.

In addition, as co-chair I recommend to the ICCA that:

- No new legislative action on cyberhate should be introduced, excluding national educational initiatives. Moreover, Countries with speech codes should use discretion in enforcing the laws against Internet hate speech so as not to trivialize the law. The laws should be reserved for particularly egregious cases.
- The ICCA endorse the Anti-Cyberhate Working Group and recognizes it as a leading effort to unite all interested parties, parliamentarians, intermediaries, NGOs and users in a common effort to address cyberhate in a collaborative working environment. To that end, the Task Force recommends that the ACWG be invited to present reports at ICCA meetings and participate in ICCA-convened events.
- Governments should ensure that laws and policing agency policy are sufficiently robust to ensure that they can respond to those actions that move beyond words into real world criminal behavior, such as true threats, stalking, and violence, etc.
- NGOs and Parliamentarians recognize the difficulty of policing online hate and work with and recognize efforts to combat it, looking at real-world solutions to the problem.
- The question of online anonymity and privacy be studied for its role in contributing to online hate and that this study be undertaken under either the auspices of ICCA or ACWG.

