

André Oboler

# Solving Antisemitic Hate Speech in Social Media through a Global Approach to Local Action

## Introduction

In 2008 the term “Antisemitism 2.0” was coined to describe the normalisation of antisemitism in society through the use of social media.<sup>1</sup> In the past decade, the impact of social media in daily life has grown dramatically as has its use as a medium for hate speech.<sup>2</sup> Antisemitism remains one of the most common forms of hate speech in social media along with the rise in anti-Muslim hate speech following the rise of Daesh (ISIS), the resulting refugee crisis, and the rise in global terrorism. Other groups in society are also targeted with misogyny, homophobia, and racism against Indigenous peoples making headlines around the world. The Jews have again been the canary in the coal mine with efforts to tackle Antisemitism 2.0 leading the way in the broader response to what has become known as Hate 2.0.<sup>3</sup> In anonymous platforms like 4chan and 8chan, a more extreme version of this hate was normalised and by 2019 had led to members of that community carrying out multiple deadly terrorist attacks.<sup>4</sup>

The first problem in tackling Antisemitism 2.0 is being able to identify antisemitic content in social media in an efficient and effective manner so it can be empirically measured. This problem was identified as a key challenge at the 2009 Global Forum for Combating Antisemitism, and a solution involving crowdsourcing of reports and automated verification was presented to a meeting of the Online Antisemitism Working Group of the Global Forum in 2011. The soft-

---

1 Cf. A. Oboler, “Online Antisemitism 2.0. ‘Social Antisemitism on the Social Web’,” *Jerusalem Center for Public Affairs Post-Holocaust and Antisemitism Series* 67 (April 2008), <https://jcpa.org/article/online-antisemitism-2-0-social-antisemitism-on-the-social-web/>.

2 Cf. M. Wendling, “2015: The Year that Angry Won the Internet,” *BBC News*, December 30, 2015, <https://www.bbc.com/news/blogs-trending-35111707>.

3 Cf. A. Oboler, *Aboriginal Memes and Online Hate* (Melbourne: Online Hate Prevention Institute, 2012), <http://ohpi.org.au/aboriginal-memes-and-online-hate/>.

4 Cf. A. Oboler, W. Allington, and P. Scolyer-Gray, *Hate and Violent Extremism from an Online Subculture: The Yom Kippur Terrorist Attack in Halle, Germany* (Melbourne: Online Hate Prevention Institute, 2019), ii, 1–3, <https://ohpi.org.au/hate-and-violent-extremism-from-an-online-subculture-the-yom-kippur-terrorist-attack-in-halle-germany/>.

ware was presented at the 2013 meeting and formally endorsed after a draft report based on the first 2,024 reported items was circulated at the 2015 meeting.<sup>5</sup> The final report was released on Holocaust Memorial Day in 2016.<sup>6</sup>

The new technical solution allows the problem to be redefined as a quality of service challenge where the level of hate must be constantly measured and kept below a threshold of acceptability.<sup>7</sup> As was foreshadowed in 2010, if platforms failed to keep the level of hate low enough, governments would step in with regulation.<sup>8</sup> This occurred in 2016 in Germany and the European Union with agreements between companies and governments.<sup>9</sup> In 2017, Germany passed regulatory laws targeting non-compliance.<sup>10</sup> Facebook itself was singled out on the question of measurement and fined two million Euros, when it only reported the number of complaints explicitly made under the German law rather than reporting all user reports of antisemitism and other forms of hate flagged by German users as the law required.<sup>11</sup> The result was underreporting by multiple orders of magnitude.

The solution to antisemitism in social media has two parts. The first is a global effort to create transparency and accountability through a sharing of real-time data about hate speech in social media. The second part is local action in response to this data which is in keeping with the values and norms of each society. For example, criminal sanctions for posters of hate speech; penalties for social media platforms; counter speech exposing hate speech; counter speech promoting alternative positive narratives; education; campaigns targeting hate promoters, social media platforms, or advertisers.

---

5 Cf. A. Oboler, *Measuring the Hate: The State of Antisemitism in Social Media* (Melbourne: Online Hate Prevention Institute, 2016), <http://ohpi.org.au/measuring-antisemitism/>.

6 Cf. *ibid.*

7 Cf. A. Oboler and K. Connelly, "Hate Speech: A Quality of Service Challenge," in *2014 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)* (Piscataway: IEEE, 2014), 117–21.

8 Cf. A. Oboler, "Time to Regulate Internet Hate with a New Approach?," *Internet Law Bulletin* 13, no. 6 (2010): 102–6; Oboler, "A Legal Model for Government Intervention to Combat Online Hate," *Internet Law Bulletin* 14, no. 2 (2011): 27–30.

9 Cf. "European Union agreement with Social Media Platforms on tackling Hate Speech," Online Hate Prevention Institute, last modified May 31, 2016, accessed December 19, 2019, <http://ohpi.org.au/european-union-agreement-with-social-media-platforms-on-hate-speech/>.

10 Cf. M. Connellan, "Germany Holds Social Media Companies to Account for Hate Speech," *SBS News*, April 6, 2017, <https://www.sbs.com.au/news/germany-holds-social-media-companies-to-account-for-hate-speech>.

11 Cf. "Germany Fines Facebook for Underreporting Hate Speech Complaints," *Deutsche Welle*, July 2, 2019, <https://www.dw.com/en/germany-fines-facebook-for-underreporting-hate-speech-complaints/a-49447820>.

## The Danger of Antisemitism 2.0

Antisemitism 2.0 is

the use of online social networking and content collaboration to share demonization, conspiracy theories, Holocaust denial, and classical antisemitic motifs with a view to creating social acceptability for such content.<sup>12</sup>

The paper describing the new phenomenon was pre-released at the Global Forum for Combating Antisemitism in Jerusalem in February 2008, while that week's *New York Jewish Week* carried a first-page story warning that the phenomena was "potentially more hazardous than the relatively straightforward smear campaigns and petitions of yesteryear."<sup>13</sup> Facebook at this time had just turned four, was slightly less popular than MySpace, and boasted around 100 million users.<sup>14</sup>

As social media's influence continued to grow, the danger of Antisemitism 2.0 was further explained in a hearing before the Italian Parliament's subcommittee on antisemitism. The hearing heard that,

the danger is not so much that people might read content inspired by anti-Semitism, but rather that they may be induced to accept it as a valid point of view, a fact of life, or something with which one may or may not agree, but not something whose dissemination one should oppose. This is where the risk lies. Some people will feel affected by it and will want to do something against anti-Semitism, but others will remain passive and consider it normal, humdrum, legitimate. And this gives rise to a culture in which hatred, racism and antisocial behaviour are able to spread, posing huge risks to law and order and to security.<sup>15</sup>

It is not just the online world that is threatened but the values of society as a whole.

---

<sup>12</sup> Oboler, "Online Antisemitism 2.0."

<sup>13</sup> T. Snyder, "Anti-Semitism 2.0 Going Largely Unchallenged," *The New York Jewish Week*, February 20, 2008, 1, <http://jewishweek.timesofisrael.com/anti-semitism-2-0-going-largely-unchallenged/>.

<sup>14</sup> Cf. M. Arlington, "Facebook Now Nearly Twice The Size Of MySpace Worldwide," *Tech Crunch*, January 22, 2009, <https://techcrunch.com/2009/01/22/facebook-now-nearly-twice-the-size-of-myspace-worldwide/>.

<sup>15</sup> Chamber of Deputies, "Presentazione del documento conclusivo del Comitato di Indagine sull'Antisemitismo = Presentation of the final document of the Sub-Committee of Inquiry into Anti-Semitism," *Convegni e conferenze* 164 (Rome: Chamber of Deputies, 2013), 304.

The alt-right in the United States is a manifestation of Antisemitism 2.0. It began in parts of Reddit and 4chan,<sup>16</sup> as an “obscure, largely online subculture” before entering the “very center of American politics.”<sup>17</sup> *Rolling Stone* describes it as a white supremacy movement linked by a “contempt for mainstream liberals, feminists, ‘social justice warriors’ and immigrants,”<sup>18</sup> but overlooked the most common target: the Jews.

The antisemitism of the alt-right and the subculture it emerged from can be seen in 4chan’s Politically Incorrect boards efforts to mainstream the antisemitism meme of the Jew.<sup>19</sup> It can be seen in the dedication of this subculture to white nationalism with antisemitism “at its theoretical core.”<sup>20</sup> It can be seen in the promotion of the idea of Jews as grand manipulators trying to destroy the white race.<sup>21</sup> DNA testing followed up by posts to prove one’s “whiteness” have taken off within these groups.<sup>22</sup> In the alt-right, antisemitism can be seen in Richard Spencer’s call to “Hail Trump, hail our people, hail victory!” followed by a Nazi salute at the “DeploraBall.”<sup>23</sup> It can be seen in the meme of Pepe the Frog as a Nazi with Trump’s face, spread virally by the alt-right and even by

---

16 “Alt Right,” Know Your Meme, accessed September 1, 2017, <http://knowyourmeme.com/memes/cultures/alt-right>.

17 G. Michael, “The Seeds of the Alt-right, America’s Emergent Right-wing Populist Movement,” *The Conversation*, November 23, 2016, <https://theconversation.com/the-seeds-of-the-alt-right-americas-emergent-right-wing-populist-movement-69036>.

18 C. Skutsch, “The History of White Supremacy in America,” *Rolling Stone*, August 19, 2017, <http://www.rollingstone.com/politics/features/the-history-of-white-supremacy-in-america-w498334>.

19 Cf. A. Oboler, *The Antisemitic Meme of the Jew* (Melbourne: Online Hate Prevention Institute, 2014), <http://ohpi.org.au/the-antisemitic-meme-of-the-jew/>.

20 Cf. E. K. Ward, “Skin in the Game: How Antisemitism Animates White Nationalism,” *The Public Eye*, June 29, 2017, <http://www.politicalresearch.org/2017/06/29/skin-in-the-game-how-an-tisemitism-animates-white-nationalism/>.

21 This idea of war between the Jews and the white race is the origin of the “Gas the Kikes, Race War Now” slogan common on these forums. The idea of the Jews as the enemy emerges from Nazi literature, including their use of the *Protocols of the Elders of Zion*, and continues to be portrayed in modern online memes and websites, for example, that of Holocaust denier Andrew Carrington Hitchcock, see <http://andrewcarringtonhitchcock.com/jewish-genocide-of-the-white-race-case-closed/>, accessed December 12, 2019 [no longer available].

22 Cf. E. Reeve, “Alt-Right Trolls Are Getting 23andme Genetic Tests to ‘Prove’ Their Whiteness,” *Vice News*, October 8, 2016, [https://www.vice.com/en\\_us/article/vbygqm/alt-right-trolls-are-getting-23andme-genetic-tests-to-prove-their-whiteness](https://www.vice.com/en_us/article/vbygqm/alt-right-trolls-are-getting-23andme-genetic-tests-to-prove-their-whiteness).

23 Cf. D. Lombroso and Y. Appelbaum, “‘Hail Trump!’: White Nationalists Salute the President-Elect,” *The Atlantic*, November 21, 2016, <https://www.theatlantic.com/politics/archive/2016/11/richard-spencer-speech-npi/508379/>.

Trump's son.<sup>24</sup> It can be seen in the resulting widespread use of Pepe as a symbol by the alt-right and the ADL's response listing it as a hate symbol.<sup>25</sup> It can also be seen in the alt-right's (((echoes))) targeting Jews on Twitter.<sup>26</sup>

The alt-right and those in the subculture from which it grew are not just promoters of antisemitism. They are deliberate promoters of the normalization of antisemitic messages in society through social media. They are deliberate promoters and accelerators of Antisemitism 2.0. This wasn't a natural progression but a result of efforts by hardcore antisemites from Stormfront who deliberately infiltrated and then reshaped the subculture on /pol/ to advance their own ideology in a younger generation.<sup>27</sup> Within the subculture, the brainwashing of people to accept the group's conspiracy theories is known as "red-pilling," named after the choice given to Neo in *The Matrix* movie where he had to choose between a blue pill that would put him back to sleep in an artificial reality, or the red pill which would break him out of this controlled environment and let him see the real world.<sup>28</sup> It is the use of social media to red-pill the public, opening their eyes to narratives that demonize Jews, propagate conspiracy theories, promote Holocaust denial and spreading classical antisemitic motifs in mainstream online spaces that drives Antisemitism 2.0 forward.

Take for example the triple parentheses identifying Jews on Twitter. Identifying Jews online is not new. The infamous JewWatch website is one of the most well-known and oldest antisemitic websites on the internet and is built on this concept. What the triple parentheses add is embedding the identification of Jews into the fabric of the Twittersphere. The markers would appear in the Twitter feeds of otherwise regular conversations and normalize the singling out of Jews.

The alt-right and others in this subculture formulate plans to manipulate the mainstream media, social media, and online culture to spread their narratives—

---

24 Cf. A. Ohlheiser, "Why Pepe the Frog's Nazi Phase Doesn't Worry his Creator," *The Washington Post*, September 14, 2016, [https://www.washingtonpost.com/news/the-intersect/wp/2016/09/14/why-pepe-the-frogs-nazi-phase-doesnt-worry-his-creator/?utm\\_term=.afc8b7441709](https://www.washingtonpost.com/news/the-intersect/wp/2016/09/14/why-pepe-the-frogs-nazi-phase-doesnt-worry-his-creator/?utm_term=.afc8b7441709).

25 Cf. S. Begley, "Anti-Defamation League Declares Pepe the Frog a Hate Symbol," *Time*, September 28, 2016, <http://time.com/4510849/pepe-the-frog-adl-hate-symbol/>.

26 Cf. "Triple Parentheses Echo," Know Your Meme, accessed September 1, 2017, <http://knowyourmeme.com/memes/triple-parentheses-echo>.

27 Cf. Oboler, Allington, and Scolyer-Gray, *Hate and Violent Extremism from an Online Subculture*, 70, <https://ohpi.org.au/hate-and-violent-extremism-from-an-online-subculture-the-yom-kippur-terrorist-attack-in-halle-germany/>.

28 Cf. A. Marwick and B. Lewis, "The Online Radicalization We're Not Talking About," *NY Magazine*, May 18, 2017. <https://nymag.com/intelligencer/2017/05/the-online-radicalization-were-not-talking-about.html>.

any “escape” of a meme or narrative into the mainstream is seen as victory.<sup>29</sup> The aim is to have the subculture’s narratives embedded as part of the fabric of the online world and daily life. This leaves the individual with the “choice” of becoming red-pilled or continuing life as what the subculture derogatorily call “normies.”

Following the election of President Trump and the emergence of the alt-right as a public force, members of the subculture were encouraged to red-pill their families enlarging the support base.<sup>30</sup> The alt-right marches, such as that in Charlottesville, further promote the message of white supremacy as a normal part of politics which people should accept. This message was reinforced by President Trump’s comment, “I think there is blame on both sides,” after the violence at the alt-right march in Charlottesville.<sup>31</sup> What started with an effort to normalize antisemitism in the online world has in 2017 shifted to an effort to normalize it on the streets of America. It is not just in America either: when a senior Google representative tells a UK Home Affairs Select Committee that a YouTube video titled “Jews admit organising white genocide” and featuring former KKK Grand Wizard David Duke “did not cross the line into hate speech” and therefore remains online (with over 91,000 views),<sup>32</sup> that too helps to normalize antisemitism online and in society.

Antisemitism has been, and remains, the canary in the coal mine for society. It is through the prism of the fight against antisemitism that both the new manifestations of hate and new efforts to tackle it emerge. We must continue to specifically tackle Antisemitism 2.0 even as we simultaneously use what we learn to also tackle the wider problem of Hate 2.0 affecting other groups in society. The creation of social acceptability for racism,<sup>33</sup> religious vilification (particularly against Muslims),<sup>34</sup> misogyny, homophobia, and other forms of hate weaken society and makes the fight against antisemitism that much harder.

---

<sup>29</sup> Cf. *Ibid.*

<sup>30</sup> Cf. M. Pearl, “How to Tell if Your Alt-Right Relative Is Trying to Redpill You at Thanksgiving,” *Vice*, November 24, 2016, [https://www.vice.com/en\\_au/article/nnk3bm/how-to-tell-if-your-alt-right-relative-is-trying-to-redpill-you-at-thanksgiving](https://www.vice.com/en_au/article/nnk3bm/how-to-tell-if-your-alt-right-relative-is-trying-to-redpill-you-at-thanksgiving).

<sup>31</sup> D. Merica, “Trump Says Both Sides to Blame amid Charlottesville Backlash,” *CNN*, August 16, 2017, <http://edition.cnn.com/2017/08/15/politics/trump-charlottesville-delay/index.html>.

<sup>32</sup> S. Oryszczuk, “Google Chief: Far Right Video Accusing Jews of ‘Organising Genocide’, Isn’t Hate Speech,” *The Times of Israel*, March 15, 2017, <http://jewishnews.timesofisrael.com/google-chief-far-right-video-accusing-jews-of-organising-genocide-isnt-hate-speech/>.

<sup>33</sup> Cf. Oboler, *Aboriginal Memes*.

<sup>34</sup> Cf. A. Oboler, “The Normalisation of Islamophobia through Social Media: Facebook,” in *Islamophobia in Cyberspace: Hate Crimes Go Viral*, ed. I. Awan (London: Routledge, 2016), 41–61.

## Accountability of Platforms and People

At the 2009 Global Forum for Combating Antisemitism, the working group on antisemitism on the internet and in the media identified the lack of metrics for measuring antisemitism on social media as a major challenge. The challenge remained open and was reaffirmed at a 2011 meeting of the working group and then in a report released at the 2013 Global Forum.<sup>35</sup> It noted the

lack of metrics on: a. The number of problem items in specific platforms e.g. reported groups in Facebook, reported Videos on YouTube; b. The number of items resolved on specific platforms e.g. groups shut down, videos removed, complaints reviewed or dismissed; ... d. The time delay between something being reported and action being taken in a specific platform.<sup>36</sup>

The reluctance of social media platforms to tackle antisemitism was clear. Facebook, for example, refused to recognise Holocaust denial as a form of hate speech and therefore as a breach of its community standards.<sup>37</sup> This was confirmed to the Global Forum's working group in a 2011 letter that stated in part, "the mere statement of denying the Holocaust is not a violation of our policies. We recognize people's right to be factually wrong about historical events."<sup>38</sup>

It was readily demonstrated that even obvious cases of antisemitism were being rejected when they were reported to Facebook, for example the picture of Anne Frank with the words "What's that Burning? Oh, it's my family" written across it.<sup>39</sup> Further work looking at forty-seven antisemitic Facebook pages showed how many remained online despite numerous reports.<sup>40</sup> The report led to formal complaints through the Australian Human Rights Commission in which a mediated solution, involving the removal of the listed content (at

---

<sup>35</sup> Cf. A. Oboler and D. Matas, "Online Antisemitism: A Systematic Review of the Problem, the Response and the need for Change," issued May 30, 2013 (Jerusalem: Israeli Ministry of Foreign Affairs, 2013), <http://mfa.gov.il/MFA/AboutTheMinistry/Conferences-Seminars/GFCA2013/Pages/Online-Antisemitism-A-systematic-review.aspx>.

<sup>36</sup> Ibid.

<sup>37</sup> Cf. A. Oboler, "Facebook, Holocaust Denial and Antisemitism 2.0," Jerusalem Center for Public Affairs, August 27, 2009, <http://jcpa.org/article/facebook-holocaust-denial-and-anti-semitism-2-0/>.

<sup>38</sup> The full text of the letter can be seen in Oboler and Matas, "Online Antisemitism," 50.

<sup>39</sup> "Facebook Fails Review," Online Hate Prevention Institute, issued September 14, 2012, accessed December 12, 2019, <http://ohpi.org.au/facebook-fails-review/>.

<sup>40</sup> Cf. A. Oboler, *Recognizing Hate Speech: Antisemitism on Facebook* (Melbourne: Online Hate Prevention Institute, 2013), [http://ohpi.org.au/reports/IR13-1\\_Recognizing\\_hate\\_speech\\_anti\\_semitism\\_on\\_Facebook.pdf](http://ohpi.org.au/reports/IR13-1_Recognizing_hate_speech_anti_semitism_on_Facebook.pdf).

least for Australian users) and a commitment to remove any identical content uploaded in the future was given by Facebook. The problem is not limited to Facebook, indeed later research has shown Facebook's response, while still far from acceptable, was the most effective response from the major social media platforms. YouTube, for example, took a similar approach to Facebook on Holocaust denial until June 2019, when, as part of a suite of policy changes to tackle hate, they did what Facebook still refuses to do and announced a ban on Holocaust denial.<sup>41</sup>

While the isolated examples and small samples in research and advocacy in the early 2010s demonstrate the problem, without detailed metrics there was no transparency and as a result there could be no accountability. Social media platforms were largely self-regulated and unresponsive. Inside the United States, hate speech enjoys first amendment protection, meaning laws seeking to restrict it would be deemed unconstitutional. Outside the United States, the platforms argued that they were mere carriers, and it is the users who should be prosecuted if illegal hate speech is uploaded. This argument is problematic as it is the platforms that decide what content is promoted to whom and profit from the existence of the content.

The 2009 Global Forum for Combating Antisemitism recommended that carrier immunity was "too broad and needs to be limited in the case of antisemitism and other forms of hate," more specifically, it recommended that

while real time communication may be immune, stored communication e.g. user published content, can be brought to a service providers attention and the provider can then do something about it. Opting not to do something about it after a reasonable time should in all cases open the service provided up to liability.<sup>42</sup>

The 2013 report to the Global Forum, which repeated this recommendation and presented the TEMPIS Taxonomy which outlined different types of online communication (defined according to factors such as timing, empowerment of users, moderation, publicness, identity and social impact) so that similar types of communication could have the same expectations applied to them regardless of the social media platform being used.<sup>43</sup> A 2010 article warned that

---

<sup>41</sup> Cf. "Our Ongoing Work to Tackle Hate," YouTube Official Blog, issued June 5, 2019, <https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>.

<sup>42</sup> Oboler and Matas, "Online Antisemitism," 30.

<sup>43</sup> Cf. *ibid.*, 5–10.



those who profit from user generated content need to be given the responsibility to take reasonable steps to ensure their platforms have a robust response to the posting of hateful material. The role of government, and the law, is to ensure reasonable steps are indeed taken.<sup>44</sup>

A draft report released at the 2015 Global Forum for Combating Antisemitism for the first time provided a large sample of data about antisemitism in social media. The final report, released on International Holocaust Remembrance Day, January 27, 2016, added statistics about platform responsiveness.<sup>45</sup> Based on a sample of 2,024 unique items of antisemitic content from across Facebook, YouTube, and Twitter, the report divided the content both by platform and across four different categories of antisemitism: traditional antisemitism (49%), New Antisemitism (34%), Holocaust denial (12%), and promoting violence (5%).<sup>46</sup> The percentage of content removed by the platforms, after ten months, varied greatly both by category and within each category. The high was Facebook removing 75% of content promoting violence and the low was YouTube removing just 4% of New Antisemitism.<sup>47</sup> The full spread can be seen in Table 1.

**Table 1:** Percent of Antisemitism Removed after 10 Months

	Traditional	New Antisemitism	Holocaust Denial	Violence
Facebook	42 %	27%	58 %	75 %
Twitter	25 %	20%	20 %	14 %
YouTube	9 %	4 %	10 %	30 %

In 2017, after efforts to resolve the problem of online hate through agreements between the government and the social media companies failed to deliver the desired results, Germany became the first country to legislate liability for platform providers. The law allows for fines of up to 500 million euros for platforms that systematically fail to remove obvious breaches of German hate-speech law within twenty-four hours.<sup>48</sup> This German approach was a positive step forward and long overdue but relies on the ability to track social media hate speech in order for it to be effectively applied. It is also, at present, a blunt tool without differentiation for the different types of communication which may be used on-

<sup>44</sup> Oboler, “Time to Regulate,” 105.

<sup>45</sup> Cf. Oboler, *Measuring the Hate*.

<sup>46</sup> Cf. *ibid.*, 7.

<sup>47</sup> Cf. *ibid.*

<sup>48</sup> Cf. Connellan, “Germany Holds Social Media Companies to Account for Hate Speech.”

line, as discussed in the TEMPIS Taxonomy. Not all hate requires the same priority of response. Content which incites violence may need a more rapid response than Holocaust denial, for example, while content shared publicly in a form that can go viral may need a more urgent response than the same content sent as a private message to a single person.

There is a need to track antisemitic content and content related to other forms of hate and extremism, both at the level of individual items and at the level of summary data showing the content which is impacting society. The challenge first presented at the Global Forum in 2009 is now more urgent than ever. Indeed, governments beyond Germany have now taken up this challenge and are working to develop plans to regulate transparency.<sup>49</sup>

## Approaches to Monitoring Antisemitism in Social Media

The gathering of data on online antisemitism can be approached in three ways: expert solicitation, automation through artificial intelligence (AI), or crowdsourcing. Each approach has both advantages and drawbacks. The best approach is one which synthesizes the contributions of all three approaches to triangulate a more accurate and complete picture.

Expert solicitation is the oldest approach and involves experts first finding and then assessing examples of antisemitism. The Simon Wiesenthal Centre's *Digital Terrorism and Hate Project*, for example, has been using this approach since the late 1990s.<sup>50</sup> The problem is the time involved, the limited number of experts, and the fact that using experts for anything but the most viral and high impact cases is not an effective use of resources, makes this an impractical approach to use a comprehensive understanding at scale. This is particularly true when close to real-time analysis is needed. As an ADL spokesperson told

---

<sup>49</sup> Cf. *Online Safety Legislative Reform: Discussion Paper*, Department of Communications and the Arts, Australian Government, issued December 11, 2019, <https://www.communications.gov.au/have-your-say/consultation-new-online-safety-act>, 23; *Online Harms White Paper*, UK Government, Presented to Parliament by the Secretary of State for Digital, Culture, Media & Sport and the Secretary of State for the Home Department by Command of Her Majesty, issued April 2019, [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/793360/Online\\_Harms\\_White\\_Paper.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf), 40.

<sup>50</sup> Cf. "Facebook, Youtube+: How Social Media Outlets Impact Digital Terrorism and Hate," Simon Wiesenthal Center, June 18, 2009, <http://www.wiesenthal.com/about/news/facebook-youtube-how.html>.

the *New York Jewish Week* when the issue of Antisemitism 2.0 was first raised back in 2008, “we can’t sit here all day monitoring YouTube and Facebook.”<sup>51</sup> There is a role for experts, but it needs to be reserved for analysis, investigating new phenomena of antisemitism, providing commentary on high impact cases, and perhaps most importantly for training both people and artificial intelligence systems. This includes training for civil society organisations, volunteers, the staff of both social media companies and the companies they outsource relevant functions to, as well as training people in government across the areas of human rights, public policy, law reform, and law enforcement.

Artificial Intelligence is promoted by some as a silver bullet. A web crawler can be used to capture information from websites which can then be analyzed. At a large scale, this is how data is gathered for search engines such as Google. A relevant page is chosen as a starting point, all outgoing links are explored and if those linked pages are classified as antisemitic, the process will repeat and their links will be followed as well. On social media, a similar approach can be applied using algorithms that either crawl through the content on social media sites, or with the consent of a platform provider directly access the content on the platform’s servers. An example can be seen in the work of Monika Schwarz-Friesel where a specially created web crawler for antisemitic discourse gathered more than 250,000 items of text.<sup>52</sup>

AI solves the problem of scale by applying raw machine power to the task. Given enough time, it is able to read through all the content that is publicly accessible. The problem is in understanding the content. There are two limitations: the first is an inability to process certain types of content, for example, an AI agent may be limited to parsing text and therefore ignore the vast amount of content in images and videos. The second is understanding the text it reads, processing it, and giving it context.

The first problem can in theory be addressed with more complex algorithms which can extract text written into an image, or transcribe sound. These tasks, however, significantly increase the complexity and the cost of the processing. With the huge volume of new content being uploaded all the time, these approaches are not practicable. Nor are they likely to be practicable in the future as the quality of the content continues to increase, requiring greater processing power whenever such additional processing power becomes available. What about messages delivered through the images themselves? The Anne Frank

---

51 Snyder, “Anti-Semitism 2.0 Going Largely Unchallenged.”

52 Cf. M. Schwarz-Friesel, “‘Antisemitism 2.0’—The Spreading of Jew-hatred on the World Wide Web,” in *Comprehending and Confronting Antisemitism: A Multi-Faceted Approach*, ed. A. Lange et al. (Berlin: De Gruyter, 2019), 317.

meme previously discussed requires the graphical representation of text in the image to be converted into real text that can be processed by the computer, then requires an ability to recognize Anne Frank. Both these tasks are possible today, but they are costly to do at scale. Once done, the second problem, that of context and understanding, needs to be solved. In this case it requires a knowledge of Anne Frank's connection to the Holocaust, then an association between the word "burning" and crematoria and the extermination of Jewish families. To further confuse the AI, the statement of her family burning is inaccurate as it was her father who survived and published her diary. For a human with the appropriate background knowledge, the message mocking the victims of the Holocaust is clear. For AI, extracting this message from the many possible messages is beyond what is currently possible. A system could be built to deal with this specific meme, but it would not recognise other memes that promote the same antisemitic ideas, for example the image on Instagram saying "when you meet a Jewish girl and remember you just bought an oven that can fit a person."<sup>53</sup>

The second problem, that of context and understanding, occurs even when the processed content originates as regular text. Simple AI approaches use keywords to identify hate speech. Imagine running a Google search for "kike" on Facebook (to do this, enter the search in Google as "site:facebook.com kike" without the quotes). The problem with this approach is immediately obvious. There are 1,750,000 results for this found on Google, but a quick glance shows that many of them would be false positives. One word is not enough to accurately find such content. What about kike and gas together? The first result is a page called "Kike Gas" which repeatedly posts pictures of an oven with gas cylinders of various sizes, a refill price, and a phone number.<sup>54</sup> The page has an address in Puerto Rico and a picture of a real-world gas cylinder storage area with a sign saying Gas Kike. The page appears to be related to a real business, and there is no overt antisemitic content. Searching for kikes (plural) and gas gives better results: a blank Facebook page with the name "Gas the Kikes" claiming to be a university,<sup>55</sup> a post from a group combating antisemitism which quotes antisemitic phrases,<sup>56</sup> an alt-right like page dedicated to "Aryan Argentine Memes,"<sup>57</sup> a

---

53 J. Butler, "Far-Right Extremist Groups Are Using Instagram As A Recruiting Ground," *10 Daily*, March 3, 2020 <https://10daily.com.au/news/tech/a200302fyqto/far-right-extremist-groups-are-using-instagram-as-a-recruiting-ground-20200302>.

54 <https://www.facebook.com/KIKE-GAS-294941390704685/> [no longer available].

55 <https://www.facebook.com/pages/Gas-the-Kikes/920193234669031>.

56 Documenting Anti-Semitism, "USA, April 5, 2017 AMCHA 2016 annual study," Facebook, April 6, 2017, <https://www.facebook.com/Documenting.Anti.Semitism/posts/1356674271077720:0>.

page from Facebook’s Help Community titled “I am sick and tired of the antisemitism that is allowed on this site” in which people have posted examples of antisemitism some of which have been linked to from that page for months with no resulting action,<sup>58</sup> and a page titled “Right Wing Death Squad,”<sup>59</sup> among many other false positives.

A more specific search for “gas the kikes” in October 2017 gave 594 results from YouTube but none on Facebook (despite the presence of the “gas the kikes” page we know exists).<sup>60</sup> The same search in March 2020 showed that this has reduced to 31 results.<sup>61</sup> On YouTube itself a search for “gas the kikes” gave 176 results in October 2017,<sup>62</sup> but had dropped to 48 by March 2020.<sup>63</sup> On Facebook’s internal search looking for “gas the kikes” brings up a post with the lyrics of an antisemitic music video with phrases such as “I wanna gas all the kikes until they become zero,”<sup>64</sup> another post says “Gas the kikes, race war now!” over a background of rainbow and multi-coloured heart balloons,<sup>65</sup> many others use the phrase to describe what antisemites are saying as they comment on news stories. The results show that major platforms like Facebook and YouTube are getting better for at least some types of antisemitic expression, but even for these easy to identify cases, problems persist.

Text analysis is not much more advanced than a search, though by examining the raw data such approaches may find examples which are being filtered out by search engines or platforms’ internal search functions. Text analysis can find specific phrases, or combinations of words and phrases related to antisemitism through AI,<sup>66</sup> but the AI cannot accurately differentiate between antisemitic content and discourse about antisemitism (i. e., discussing and seeking to counter antisemitism) so the capacity to use AI to measure antisemitism directly is limited. However well it does, there will always be many false positives. Without human review, the results are likely to be misleading—just as any search

---

57 <https://www.facebook.com/AryanArgieMemes/?nr> [no longer available].

58 <https://www.facebook.com/help/community/question/?id=881307688690518> [no longer available].

59 <https://www.facebook.com/Right-Wing-Death-Squad-591174651085238/> [no longer available].

60 Search as at October 15, 2017.

61 Search as at March 15, 2020.

62 Search as at October 15, 2017.

63 Search as at March 15, 2020.

64 <https://www.facebook.com/enix.sho/posts/1491580350918573> [no longer available].

65 <https://www.facebook.com/myron.dawson.965/posts/130188527710071> [no longer available].

66 Cf. S. Zannettou et al., “A Quantitative Approach to Understanding Online Antisemitism,” Network Contagion Research Institute: *arXiv e-prints*, issued September 2018, accessed December 12, 2019, <https://arxiv.org/abs/1809.01644>.

is likely to provide some irrelevant results. As can be seen, the results also depend on the search tool. Facebook clearly allows its internal search far more access than Google is able to get. YouTube only searches the titles and descriptions of videos while a Google search for content on YouTube also picks up phrases appearing in the first few comments. These technical limitations may be built into software, or may be inherited unknowingly by software due to dependencies on other libraries, online services, or APIs (application programming interfaces). More advanced text analysis can be seen in “Conversation AI” a tool launched by a Google subsidiary in September 2016, and

designed to use machine learning to automatically spot the language of abuse and harassment—with ... an accuracy far better than any keyword filter and far faster than any team of human moderators.<sup>67</sup>

The tool is still based on text analysis and the use of phrases and the presence or absence of other words to determine if content is antisemitic. While it may be an improvement over simple keyword identification, it is far from a robust solution. Even if it were, 4chan’s /pol/ quickly responded with “Operation Google” which sought to use “Googles” as a code word for African Americans, “Skypes” as code word for Jews, “Yahoos” as code for Mexicans and “Skittles” as code for Muslims.<sup>68</sup> More recently, /pol/ started a campaign trying to get the hash symbol to be used as a replacement for the swastika, arguing that if it was done successfully, they would either make many regular uses of the hash, such as hashtags, seem like an endorsement of neo-Nazism, or forcing platforms like Twitter to break their core systems.<sup>69</sup> The use of code words greatly complicates the process of text analysis, particularly if the code words are regularly changed. In a fight between regulation by Artificial Intelligence and circumvention through human creativity, the computers have little chance of winning. This is especially true when the message seeking to be regulated is an idea which can be rephrased and adapted, rather than a repetition of a known item of content such as song or movie.

---

<sup>67</sup> A. Greenberg, “Inside Google’s Internet Justice League and its AI-Powered War on Trolls,” *Wired*, September 19, 2016, <https://www.wired.com/2016/09/inside-googles-internet-justice-league-ai-powered-war-trolls/>.

<sup>68</sup> Cf. P. Tambutto, “4chan Aims to Fill Google with Racism Following ‘War on Trolls,’” *Crave*, September 23, 2016, <https://www.mandatory.com/living/1125143-4chan-aims-fill-google-racism-following-war-trolls>.

<sup>69</sup> Cf. M. Thalen, “4chan’s New Troll Campaign Aims to Make the Hashtag a White Supremacist Symbol,” *Daily Dot*, May 26, 2019, <https://www.dailydot.com/unclick/4chan-troll-campaign-bash-the-hash/>.

Another approach, focused on images, has been to track the spread of a specific antisemitic meme across different sites. The single meme is represented by a collection of different images, each of which can be recognised through image analysis. An experiment tracked the spread the antisemitic meme of the Jew and also noted the spread of variants of this meme as it infected other memes.<sup>70</sup>

The last approach is that of crowdsourcing. This is the method used in the 2016 report previously discussed. Through a custom built advanced online reporting tool called Fight Against Hate, members of the public reported and categorized items of antisemitism they found across Facebook, YouTube, and Twitter.<sup>71</sup> The volume of data collected is far higher than could be gathered by experts alone but far lower than what an AI tool could find in an automated search. The main weakness of this system is the potential for users, through ignorance or malice, to incorrectly make reports. For example, organized groups could seek to use the tool to report their opponents (be it businesses, sporting teams, schools, political parties, etc.) or to report content they disagree with but which is not hate.<sup>72</sup> The solution has two parts, one part sees users offering to review content others have reported, and the system determines which items each person gets in order to ensure an independent judgement, and the other involves the limited use of experts to validate items allowing a model of trust to be developed so the system is aware which non-experts tend to agree with the experts judgements.<sup>73</sup> This approach uses people to identify and review the content and artificial intelligence approaches to ensure quality.<sup>74</sup>

Mixed approaches where experts are supported by AI give better results. For example, the use of AI to gather examples of discourse from online conversations and publications can provide a rich starting point for analysis using corpus linguistics. In corpus linguistics, a sample is gathered and then manually coded by experts, usually with each sample being coded by more than one expert to avoid both bias and error. This form of coding is not related to programming but rather to the association of a piece of text with a particular concept. This approach to manual coding in relation to Wikipedia back in 2010 exposed a prob-

---

<sup>70</sup> Cf. S. Zannettou et al., “A Quantitative Approach to Understanding Online Antisemitism Part 3: Meme Analysis,” Network Contagion Research Institute: *arXiv e-prints*, issued September 2018, accessed December 12, 2019, <https://ncri.io/2018/09/06/a-quantitative-approach-to-understanding-online-antisemitism-part-3-meme-analysis/>.

<sup>71</sup> Cf. Oboler, *Measuring the Hate*.

<sup>72</sup> Cf. Oboler and Connelly, “Hate Speech.”

<sup>73</sup> Cf. *ibid.*

<sup>74</sup> Cf. *ibid.*

lem of systematic bias related to New Antisemitism.<sup>75</sup> The combination of a crawler to gather a sample of likely antisemitic discourse, combined with expert coding to classify it, can be seen in the work of Monika Schwarz-Friesel.<sup>76</sup> The use of an API to gather a corpus and extract patterns relating to evolving antisemitic messaging was demonstrated in the work of Paul Iganski and Abe Sweiry on Twitter.<sup>77</sup> The work of Gunther Jikeli, Damir Cavar, Daniel Miehling, also looking at antisemitism on Twitter, highlights how “annotators have to be highly trained and knowledgeable about current events to understand each tweet’s underlying message within its context.”<sup>78</sup> Their work used students in a course on antisemitism in social media who were applying the IHRA Working Definition of Antisemitism to code the data and found some significant variance in application even with a common understanding of the definition. While these approaches all use real world data, they are limited to the sample their tools have gathered. The last of these studies, for example, extracted Tweets that included the word “Jew\*” (e.g., Jew, Jews, Jewish, etc.) and “Israel.” If the “Just bought an oven” and “what’s that burning” memes discussed above were (in text form) on Twitter during the time period this data was gathered, the first would be included while the second would not. This means that these approaches have the potential to focus the study of antisemitism on those posts which use words or phrases that have been previously identified. They will not provide a complete picture of the level of antisemitism, and even data on the relative level over time may be skewed by efforts from the platforms which target those same words. Research on the topic also alters the environment as it may suggest words or phrases to which platforms’ attention could be applied to reduce the count they need to report as governments increase their demands for transparency and systemic improvement. This is a benefit for real-world action tackling antisemitism, even as it disrupts research into manifestations of antisemitism.

---

75 Cf. A. Oboler, G. Steinberg, and R. Stern, “The Framing of Political NGOs in Wikipedia through Criticism Elimination,” *Journal of Information Technology and Politics* 7, no. 4 (2010): 284–99.

76 Cf. M. Schwarz-Friesel, *Antisemitism 2.0 and the Cyberculture of Hate* (Berlin: Technische Universität Berlin, 2018), [https://www.linguistik.tu-berlin.de/fileadmin/fg72/Antisemitism\\_2.0\\_short\\_version\\_final.pdf](https://www.linguistik.tu-berlin.de/fileadmin/fg72/Antisemitism_2.0_short_version_final.pdf).

77 Cf. P. Iganski and A. Sweiry, “Antisemitism on Social Media,” lecture for the Panel Internet and Antisemitism, Conference “An End to Antisemitism!,” University of Vienna, Vienna, February 21, 2018.

78 Cf. G. Jikeli, D. Cavar, and D. Miehling, “Annotating Antisemitic Online Content. Towards an Applicable Definition of Antisemitism,” submitted to Cornell University, September 29, 2019, <https://arxiv.org/abs/1910.01214>.



The best solution would be to use experts to calibrate a crowdsourced system. It would use the crowdsourcing system to review content collected by the AI tools, as occurs with items people have reported. This would give statistics on the level of false positives in the AI system. It would also review how many items in crowdsourced systems, gathered initially from human reports, were not found by the AI. This would provide data on the AI's blind spots and the degree of false negatives. The three approaches together would allow a triangulation on the real nature of antisemitism in social media.

## Transparency through Global Cooperation

Creating real transparency around the problem of Antisemitism 2.0 requires a global approach. Social media platforms often block content in particular countries rather than globally. Differences can occur between the treatment of content in different languages. A crowdsourced approach fundamentally needs the support not only of a large crowd but of one seeing social media through the lens of different countries, languages, and cultures. It needs to be calibrated to relevant experts. For a truly global picture, the tools and methodology also need to be consistent.

The work of the 2016 “Measuring the Hate: The State of Antisemitism in Social Media” report based on crowdsourcing is limited to the English language. The data was not easily accessible to other researchers and was managed by a single organisation. A 2017 report from the World Jewish Congress and Vigo Social Intelligence looks at antisemitism across many countries;<sup>79</sup> it is based on automated text analysis (with some manual review) and skewed toward expressions of antisemitism that the software was able to easily identify which results, for example, in overreporting on Twitter and underreporting on YouTube. This problem also exists in academic research where research is more likely to follow the path with the lowest implementation barriers. Some platforms, and some content on certain platforms, are simply harder to include in AI gathered data and therefore less likely to be included in research.

The Fight Against Hate reporting tool has been rebuilt so it can operate in any language (subject to a list of translations being provided) and can be embedded on the websites of many organisations and configured to their needs. Exam-

---

<sup>79</sup> Cf. World Jewish Congress and Vigo Social Intelligence, *The Rise of Anti-Semitism on Social Media: Summary of 2016* (New York: World Jewish Congress, 2017), <http://www.crif.org/sites/default/fichiers/images/documents/antisemitismreport.pdf>.

ples of antisemitism are classified into categories and then subcategories as they are reported. There are thirty-one categories in all, and together they cover all the aspects of the *IHRA Working Definition of Antisemitism* as well as the *IHRA Working Definition of Holocaust Denial and Distortion*.

There is also a review capability where people can check and classify items others have reported. Organisations can now see a summary report and a list of the items reported via their website, enabling them to respond in real time and monitor trends. The next step is for organisations to partner in the project, help translate the tool to their language, and to use it to gather data locally for themselves while also contributing to the global pool of data. The first such project is taking place with Australia and Italy, led by Australia as part of its engagement with the International Holocaust Remembrance Alliance.

An additional tool, CSI-CHAT (Crowd Sourced Intelligence—Cyber Hate and Threats), has been developed to enable organisations and experts to access and work with the data from Fight Against Hate. Data can be sorted, searched, classified, annotated, and compiled. The tool also produces statistical reports on the data including trend analysis and dataset comparisons.

Organisations will have access to work with the data reported via the gateway on their own website. Each organisation will be able to choose whether they share this data with other organisations, or whether they will release it into the common data pool. Access to additional items from the common data pool will be available in return for reciprocity. As the same item may be reported to many organisations and only the first to release the item to the common data pool will be credited with it, there is an inbuilt incentive to pool data. Even where the specific items of data are not shared, the system can account for duplicates in the private data sets of multiple organisations and provide a true picture of the total numbers of items reports and their nature.

Future work will allow organisations to upload additional datasets gathered through automated approaches and assess their coverage for different types of antisemitism. Access to the human-collected data and the support of experts will allow automated tools to be further improved.

## Community Resilience through Local Cooperation

The benefits of the Fight Against Hate reporting tool combined with the CSI-CHAT analysis tool go well beyond transparency. With the involvement of different stakeholders in society, the tools can help create communities of resistance and solidarity against antisemitism, racism, and other forms of bigotry while strengthening the resilience of communities, like the Jewish community, which

are the targets of online hate. The SMARTER (Solidarity in Moving Against Racism Together while Enabling Resilience) Approach outlines how local adoption of the tool can build resilience by empowering individuals who encounter hate, their local communities whose websites feature the reporting tool, and those who are able to undertake positive action once data is made available to them through CSI-CHAT.<sup>80</sup>

To fully engage the model, a Jewish community, which would be an “Affiliated Civil Society Organisation” under the SMARTER Approach, would need to also engage other Affiliated Civil Society Organisations, General Civil Society Organisations, Government Agencies and Academic Researchers ensuring that all have access to the data the Jewish community is gathering.<sup>81</sup> Resilience is increased when it is clear there is wide engagement on the problem of online antisemitism, and this is not simply left to the Jewish community. At the same time, through reporting, members of the community can take positive action to counter online antisemitism.

With some small adaptations the system can also empower community organisations and others in civil society so they can assist during crisis situations to triage data which may provide real-time intelligence on physical threats to the community. This has been presented in the context of far-right extremism targeting the Muslim community and inciting violence through social media following mass casualty events that are blamed on the Muslim community.<sup>82</sup> It could work equally well in situations which cause a significant increase in the level of threat faced by the Jewish community, for example, an extreme escalation in the Israeli-Palestinian conflict, and particular situations where some actions of violence are occurring and social media is being used to fan the flames and incite further attacks. In such a situation, multiple organisations inside and outside the Jewish community could review online threats and escalate those that look credible for attention by police.

---

**80** Cf. A. Oboler and K. Connelly, “Building SMARTER Communities of Resistance and Solidarity,” *Cosmopolitan Civil Societies* 10, no. 2 (2018): 100–18.

**81** Cf. *ibid.*

**82** A. Oboler, “Open Source Intelligence During High Risk Situations,” lecture, Australia and Asia Pacific Launch of Tech Against Terrorism and the Global Internet Forum to Counter Terrorism, Sydney, Australia, June 20, 2018, <https://www.facebook.com/onlinehate/videos/1800284873382060/>.

## Accountability through Local Action

Reports on the hate speech that real people have seen and reported within a particular country, and evidence of any failure of social media companies to appropriately respond, will enable national governments to hold social media platforms accountable. This data from people within the jurisdiction, with commentary from local experts in line with the values and norms of the society, can support legal schemes like the one created in Germany in 2017 and ensure that they are practically and routinely applied.

Local organisations can also work with the data in a practical way. Knowing the common narratives of antisemitism, including any new narratives or symbols, can assist with the development of responses and counter narratives. Being able to monitor trends can also help to assess the effectiveness of strategies seeking to combat antisemitism. Tracking individual items can allow repeat abuses to be identified and the efforts of law enforcement in this regard can be supported by civil society. They can also form the basis for education in schools and society more generally in responding to antisemitism. Based on greater information campaigns highlighting areas of weakness in the platforms, response can be initiated. While the local action may be different around the world, with access to a common platform best practices can be created and shared.

## Conclusions

The problem of Antisemitism 2.0 is growing globally. The normalisation of hate that started online is increasingly being manifested offline. We cannot tackle the problem of antisemitism without paying specific attention to its normalisation through Antisemitism 2.0 in social media.

To meet the challenge of Antisemitism 2.0, we need data, and we need to be able to share it. There are positive developments in automatically gathering samples of potentially antisemitic data,<sup>83</sup> or discussions about Jews or Israel, to manually evaluate them for antisemitism,<sup>84</sup> but we need to move to a real-time picture of what is occurring, and how it is changing. We need to move to a model of trust which allows us to use input from more people, with varying levels of expertise, to assess the content. We need this picture to stretch across countries and

---

<sup>83</sup> Cf., e.g., Schwarz-Friesel, *Antisemitism 2.0 and the Cyberculture of Hate*; World Jewish Congress and Vigo Social Intelligence, *The Rise of Anti-Semitism on Social Media*.

<sup>84</sup> Cf. Jikeli, Cavar, and Miehl, “Annotating Antisemitic Online Content.”

across language barriers. We need to draw on experts, artificial intelligence, and the resources of the public through crowdsourcing. We need to know not just what is out there, but what people are seeing, and what is having an impact. We need enough examples of what is being seen in the wild, whether or not it contains the words and phrases we are expecting and actively looking for, to know how antisemitism is adapting to both local and global changes.

We also need information at the national level to support action against antisemitism by civil society and governments. From transparency we can create accountability, but creating that accountability is a national responsibility, and it must take place within the norms and culture of each society. Global transparency reports and common standards are a start, but some local tailoring of such reporting will become essential.

We must work together and provide the expertise and cooperation needed to bridge the technological gap, the language barriers, and the cultural differences and to empower both civil society and governments to tackle the rising, global problem of Antisemitism 2.0. That also means overcoming practical barriers such as ensuring governments are funding both civil society monitoring work and new research.<sup>85</sup>

*Dr. André Oboler is CEO of the Online Hate Prevention Institute (OHPI). As part of La Trobe's Master of Cybersecurity, a Vice-President of the IEEE Computer Society, an honorary associate in the La Trobe Law School, and a former member of Australia's delegation to the International Holocaust Remembrance Alliance. He also serves as co-chair of the Global Forum for Combating Antisemitism's working group on antisemitism on the internet and in the media.*

## References

- Arlington, Michael. "Facebook Now Nearly Twice The Size Of MySpace Worldwide." *Tech Crunch*, January 22, 2009. <https://techcrunch.com/2009/01/22/facebook-now-nearly-twice-the-size-of-myspace-worldwide/>.
- Begley, Sarah. "Anti-Defamation League Declares Pepe the Frog a Hate Symbol." *Time*, September 28, 2016. <http://time.com/4510849/pepe-the-frog-adl-hate-symbol/>.
- Butler, Josh. "Far-Right Extremist Groups Are Using Instagram As A Recruiting Ground." *10 Daily*, March 3, 2020. <https://10daily.com.au/news/tech/a200302fyqto/far-right-extremist-groups-are-using-instagram-as-a-recruiting-ground-20200302>.

---

<sup>85</sup> A. Lange et al., "Recommendations Regarding the Internet, its Influencers and its Users," in *Comprehending and Confronting Antisemitism: A Multi-Faceted Approach*, ed. A. Lange et al. (Berlin: De Gruyter, 2019), 382.

- Chamber of Deputies. "Presentazione del documento conclusivo del Comitato di Indagine sull'Antisemitismo = Presentation of the final document of the Sub-Committee of Inquiry into Anti-Semitism." *Convegni e conferenze* 164. Rome: Chamber of Deputies, 2013.
- Connellan, Matt. "Germany holds Social Media Companies to Account for Hate Speech." *SBS News*, April 6, 2017. <https://www.sbs.com.au/news/germany-holds-social-media-companies-to-account-for-hate-speech>.
- Department of Communications and the Arts, Australian Government. *Online Safety Legislative Reform: Discussion Paper*. Issued December 11, 2019. <https://www.communications.gov.au/have-your-say/consultation-new-online-safety-act>.
- Greenberg, Andy. "Inside Google's Internet Justice League and its AI-Powered War on Trolls." *Wired*, September 19, 2016. <https://www.wired.com/2016/09/inside-googles-internet-justice-league-ai-powered-war-trolls/>.
- Iganski, Paul, and Abe Sweiry. "Antisemitism on Social Media." Lecture for the Panel *Internet and Antisemitism*, Conference "An End to Antisemitism!," University of Vienna, Vienna, February 21, 2018.
- Jikeli, Gunther, Damir Cavar, and Daniel Miehling. "Annotating Antisemitic Online Content. Towards an Applicable Definition of Antisemitism." Submitted to Cornell University, September 29, 2019. <https://arxiv.org/abs/1910.01214>.
- Lange, Armin, Kerstin Mayerhofer, Dina Porat, and Lawrence H. Schiffman. "Recommendations Regarding the Internet, its Influencers and its Users." In *Comprehending and Confronting Antisemitism: A Multi-Faceted Approach*, edited by Armin Lange, Kerstin Mayerhofer, Dina Porat, and Lawrence H. Schiffman, 379–86. Berlin: De Gruyter, 2019.
- Lombroso, Daniel, and Yoni Appelbaum. "'Hail Trump!': White Nationalists Salute the President-Elect." *The Atlantic*, November 21, 2016. <https://www.theatlantic.com/politics/archive/2016/11/richard-spencer-speech-npi/508379/>.
- Marwick, Alice, and Becca Lewis. "The Online Radicalization We're Not Talking About." *NY Magazine*, May 18, 2017. <https://nymag.com/intelligencer/2017/05/the-online-radicalization-were-not-talking-about.html>.
- Merica, Dan. "Trump Says Both Sides to Blame amid Charlottesville Backlash." *CNN*, August 16, 2017. <https://edition.cnn.com/2017/08/15/politics/trump-charlottesville-delay/index.html>.
- Michael, George. "The Seeds of the Alt-right, America's Emergent Right-wing Populist Movement." *The Conversation*, November 23, 2016. <https://theconversation.com/the-seeds-of-the-alt-right-americas-emergent-right-wing-populist-movement-69036>.
- N. N. "Alt Right." Know Your Meme. Accessed September 1, 2017. <http://knowyourmeme.com/memes/cultures/alt-right>.
- N. N. "Germany Fines Facebook for Underreporting Hate Speech Complaints." *Deutsche Welle*, July 2, 2019. <https://www.dw.com/en/germany-fines-facebook-for-underreporting-hate-speech-complaints/a-49447820>.
- N. N. "Triple Parentheses Echo." Know Your Meme. Accessed September 1, 2017. <http://knowyourmeme.com/memes/triple-parentheses-echo>.
- Oboler, André. *Aboriginal Memes and Online Hate*. Melbourne: Online Hate Prevention Institute, 2012. <http://ohpi.org.au/aboriginal-memes-and-online-hate/>.
- Oboler, André. *The Antisemitic Meme of the Jew*. Melbourne: Online Hate Prevention Institute, 2014. <http://ohpi.org.au/the-antisemitic-meme-of-the-jew/>.

- Oboler, André. "Facebook, Holocaust Denial and Antisemitism 2.0." *Jerusalem Center for Public Affairs*, August 27, 2009. <http://jcpa.org/article/facebook-holocaust-denial-and-anti-semitism-2-0/>.
- Oboler, André. "A Legal Model for Government Intervention to Combat Online Hate." *Internet Law Bulletin* 14, no. 2 (2011): 27–30.
- Oboler, André. *Measuring the Hate: The State of Antisemitism in Social Media*. Melbourne: Online Hate Prevention Institute, 2016. <http://ohpi.org.au/measuring-antisemitism/>.
- Oboler, André. "The Normalisation of Islamophobia through Social Media: Facebook." In *Islamophobia in Cyberspace: Hate Crimes Go Viral*, edited by Imran Awan, 41–61. London: Routledge, 2016.
- Oboler, André. "Online Antisemitism 2.0. 'Social Antisemitism on the Social Web'." *Jerusalem Center for Public Affairs Post-Holocaust and Antisemitism Series* 67 (April 2008). <http://jcpa.org/article/online-antisemitism-2-0-social-antisemitism-on-the-social-web/>.
- Oboler, André. "Open Source Intelligence During High Risk Situations." Lecture, Australia and Asia Pacific Launch of Tech Against Terrorism and the Global Internet Forum to Counter Terrorism, Sydney, Australia. June 20, 2018. <https://www.facebook.com/onlinehate/videos/1800284873382060/>.
- Oboler, André. *Recognizing Hate Speech: Antisemitism on Facebook*. Melbourne: Online Hate Prevention Institute, 2013. [http://ohpi.org.au/reports/IR13-1\\_Recognizing\\_hate\\_speech\\_antisemitism\\_on\\_Facebook.pdf](http://ohpi.org.au/reports/IR13-1_Recognizing_hate_speech_antisemitism_on_Facebook.pdf).
- Oboler, André. "Time to Regulate Internet Hate with a New Approach?" *Internet Law Bulletin* 13, no. 6 (2010): 102–6.
- Oboler, André, and David Matas. "Online Antisemitism: A Systematic Review of the Problem, the Response and the Need for Change." Issued May 30, 2013. Jerusalem: Israeli Ministry of Foreign Affairs, 2013. <http://mfa.gov.il/MFA/AboutTheMinistry/Conferences-Seminars/GFCA2013/Pages/Online-Antisemitism-A-systematic-review.aspx>.
- Oboler, André, Gerald Steinberg, and Raphael Stern. "The Framing of Political NGOs in Wikipedia through Criticism Elimination." *Journal of Information Technology and Politics* 7, no. 4 (2010): 284–99.
- Oboler, André, and Karen Connelly. "Building SMARTER Communities of Resistance and Solidarity." *Cosmopolitan Civil Societies* 10, no. 2 (2018): 100–18.
- Oboler, André, and Karen Connelly. "Hate Speech: A Quality of Service Challenge." In *2014 IEEE Conference on e-Learning, e-Management and e-Services (IC3e)*, 117–21. Piscataway: IEEE, 2014.
- Oboler, André, William Allington, and Patrick Scolyer-Gray. *Hate and Violent Extremism from an Online Subculture: The Yom Kippur Terrorist Attack in Halle, Germany*. Melbourne: Online Hate Prevention Institute, 2019. <https://ohpi.org.au/hate-and-violent-extremism-from-an-online-subculture-the-yom-kippur-terrorist-attack-in-halle-germany/>.
- Ohlheiser, Abby. "Why Pepe the Frog's Nazi Phase Doesn't Worry his Creator." *The Washington Post*, September 14, 2016. [https://www.washingtonpost.com/news/the-intersect/wp/2016/09/14/why-pepe-the-frogs-nazi-phase-doesnt-worry-his-creator/?utm\\_term=.afc8b7441709](https://www.washingtonpost.com/news/the-intersect/wp/2016/09/14/why-pepe-the-frogs-nazi-phase-doesnt-worry-his-creator/?utm_term=.afc8b7441709).
- Online Hate Prevention Institute. "European Union agreement with Social Media Platforms on tackling Hate Speech." Last modified May 31, 2016. Accessed December 19, 2019. <http://ohpi.org.au/european-union-agreement-with-social-media-platforms-on-hate-speech/>.

- Online Hate Prevention Institute. "Facebook Fails Review." Issued September 14, 2012. Accessed December 19, 2019. <http://ohpi.org.au/facebook-fails-review/>.
- Oryszczuk, Stephen. "Google Chief: Far Right Video Accusing Jews of 'Organising Genocide', Isn't Hate Speech." *The Times of Israel*, March 15, 2017. <http://jewishnews.timesofisrael.com/google-chief-far-right-video-accusing-jews-of-organising-genocide-isnt-hate-speech/>.
- Pearl, Mike. "How to Tell if Your Alt-Right Relative Is Trying to Redpill You at Thanksgiving." *Vice*, November 24, 2016. [https://www.vice.com/en\\_au/article/nnk3bm/how-to-tell-if-your-alt-right-relative-is-trying-to-redpill-you-at-thanksgiving](https://www.vice.com/en_au/article/nnk3bm/how-to-tell-if-your-alt-right-relative-is-trying-to-redpill-you-at-thanksgiving).
- Reeve, Elspeth. "Alt-Right Trolls Are Getting 23andme Genetic Tests to 'Prove' Their Whiteness." *Vice News*, October 8, 2016. [https://www.vice.com/en\\_us/article/vbygqm/alt-right-trolls-are-getting-23andme-genetic-tests-to-prove-their-whiteness](https://www.vice.com/en_us/article/vbygqm/alt-right-trolls-are-getting-23andme-genetic-tests-to-prove-their-whiteness).
- Schwarz-Friesel, Monika. "'Antisemitism 2.0'—The Spreading of Jew-hatred on the World Wide Web." In *Comprehending and Confronting Antisemitism: A Multi-Faceted Approach*, edited by Armin Lange, Kerstin Mayerhofer, Dina Porat, and Lawrence H. Schiffman, 311–38. Berlin: De Gruyter, 2019.
- Schwarz-Friesel, Monika. *Antisemitism 2.0 and the Cyberculture of Hate*. Berlin: Technische Universität Berlin, 2018. [https://www.linguistik.tu-berlin.de/fileadmin/fg72/Antisemitism\\_2.0\\_short\\_version\\_final.pdf](https://www.linguistik.tu-berlin.de/fileadmin/fg72/Antisemitism_2.0_short_version_final.pdf).
- Simon Wiesenthal Center. "Facebook, Youtube+: How Social Media Outlets Impact Digital Terrorism and Hate." Issued June 18, 2009. Accessed December 19, 2019. <http://www.wiesenthal.com/about/news/facebook-youtube-how.html>.
- Skutsch, Carl. "The History of White Supremacy in America." *Rolling Stone*, August 19, 2017. <http://www.rollingstone.com/politics/features/the-history-of-white-supremacy-in-america-w498334>.
- Snyder, Tamar. "Anti-Semitism 2.0 Going Largely Unchallenged." *The New York Jewish Week*, February 20, 2008. <http://jewishweek.timesofisrael.com/anti-semitism-2-0-going-largely-unchallenged/>.
- Tambutto, Paul. "4chan Aims to Fill Google with Racism Following 'War on Trolls'." *Crave*, September 23, 2016. <https://www.mandatory.com/living/1125143-4chan-aims-fill-google-racism-following-war-trolls>.
- Thalen, Mikael. "4chan's New Troll Campaign Aims to Make the Hashtag a White Supremacist Symbol." *Daily Dot*, May 26, 2019. <https://www.dailydot.com/unclick/4chan-troll-campaign-bash-the-hash/>.
- UK Government. *Online Harms White Paper*. Presented to Parliament by the Secretary of State for Digital, Culture, Media & Sport and the Secretary of State for the Home Department by Command of Her Majesty. Issued April 2019. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/793360/Online\\_Harms\\_White\\_Paper.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/793360/Online_Harms_White_Paper.pdf).
- YouTube. "Our Ongoing Work to Tackle Hate." YouTube Official Blog. Issued June 5, 2019. Accessed December 19, 2019. <https://youtube.googleblog.com/2019/06/our-ongoing-work-to-tackle-hate.html>.
- Ward, Eric K. "Skin in the Game: How Antisemitism Animates White Nationalism." *The Public Eye*, June 29, 2017. <http://www.politicalresearch.org/2017/06/29/skin-in-the-game-how-antisemitism-animates-white-nationalism/>.
- Wendling, Mike. "2015: The Year that Angry Won the Internet." *BBC News*, December 30, 2015. <https://www.bbc.com/news/blogs-trending-35111707>.



- World Jewish Congress and Vigo Social Intelligence. *The Rise of Anti-Semitism on Social Media: Summary of 2016*. New York: World Jewish Congress, 2017. <http://www.crif.org/sites/default/fichiers/images/documents/antisemitismreport.pdf>.
- Zannettou, Savvas, Joel Finkelstein, Barry Bradlyn, and Jeremy Blackburn. "A Quantitative Approach to Understanding Online Antisemitism." Network Contagion Research Institute: *arXiv e-prints*. Issued September 2018. <https://arxiv.org/abs/1809.01644>.
- Zannettou, Savvas, Joel Finkelstein, Barry Bradlyn, and Jeremy Blackburn. "A Quantitative Approach to Understanding Online Antisemitism Part 3: Meme Analysis." Network Contagion Research Institute: *arXiv e-prints*. Issued September 2018. <https://ncri.io/2018/09/06/a-quantitative-approach-to-understanding-online-antisemitism-part-3-meme-analysis/>.

