

International Liability for Social Media Companies in Artsakh

Rajika L. Shah
Clinical Director, Loyola Genocide Justice Clinic
Deputy Director, Center for the Study of Law & Genocide
LMU Loyola Law School

In the Shadow of Territorial Conflict: Artsakh and Other
Legacies of Soviet-era Media Control and Speech Norms
Southwestern Law School
February 4, 2023

Social Media in 2020 War

- Before, during, and after the 2020 war, accounts on social media platforms were used as a primary means of communicating and disseminating conflict-related information
- Misinformation, disinformation, propaganda → inflame emotions and provoke further violence
- Examples:
 - Posts included hate speech referring to Armenians as “dogs,” “cockroaches,” and “vermin”
 - Extremely graphic images and videos of Armenians being tortured and executed
 - Justified violence by claiming that Armenians deserved ill-treatment as a result of alleged past wrongs
 - Largely as a result of posts whipping up Turkish and Azerbaijani nationalist sentiment, hate crimes against Armenian communities around the world began to increase
- Social media response:
 - Ineffective enforcement of policies against inauthentic behavior, harassment, and the promotion of violence
 - Failure to prevent or quickly shut down “ephemeral astroturfing”: a quick and coordinated campaign in which an account is created with the sole purpose of promoting a keyword or hashtag, then deleting the account and its content once that goal is achieved
 - Banned accounts and content often re-emerge a short time after receiving disciplinary action



LGJC Report

- Research carried out at request of Armenian Bar Association
- Goal: Examine potential liability in international law for social media companies and executives
- Understand existing precedent, focus on international criminal law
 - Nuremberg and ICTR: individual criminal liability of mass media publishers
 - Nuremberg Control Council: corporate liability for aiding and abetting
 - Myanmar: consequences of Facebook's action and inaction
- Important developments relating to accountability of social media companies:
 - *Jane Doe v. Facebook*, N.D. Cal. (Dec. 14, 2022): motion to dismiss granted with leave to amend; amended complaint to be filed mid-February 2023
 - *Gonzalez v. Google, Twitter v. Tamneh*: US Supreme Court to hear oral argument later this month

IMT Nuremberg: Julius Streicher

- Julius Streicher, publisher of Der Stürmer, a violently antisemitic, semi-official weekly newspaper from 1923-1945; editor until 1933
- One of the most influential and widely circulated Nazi-affiliated publications
- Widely known as the “Jew-Baiter Number One,” in his capacity at the publication, Streicher heralded a “call for the annihilation of the Jewish race”
- Many articles explicitly called for the “root and branch” extermination of Jewish people
- Commonly referred to Jewish people as “germ,” “pest,” and “parasite...who must be destroyed in the interest of mankind”



IMT Nuremberg: Julius Streicher

- Following his capture in May 1945, Streicher was charged with crimes against humanity
- Evidence showed that Streicher knew of and “continually received current information on the progress of the ‘final solution,’” and that this only intensified his campaign against the Jewish people
- The Nuremberg Tribunal found Streicher **guilty** of crimes against humanity for his incitement to murder and extermination at a time when he knew Jews were being systematically killed
- Hanged in Nuremberg on October 16, 1946
- → 1951 Genocide Convention, Article III(c): “Direct and public incitement to genocide”



ICTR: Media Case



- Ferdinand Nahimana and Jean Bosco Barayagwiza founded Radio-Television Libre des Mille Collines (RTLM) and acted as its official representatives
- RTLM was widely broadcast, particularly among younger generations who played it on the streets, at work, and in bars, taxis and markets
- Leading up to April 6, 1994, RTLM broadcasts increasingly distorted facts to portray the Tutsi as unjustifiably wealthy in a country of enormous poverty, as well “ethnic stereotyping in reference to physical characteristics”
- After April 6, RTLM escalated its anti-Tutsi rhetoric by unequivocally defining the enemy as the Tutsi and explicitly calling for the extermination of the Tutsi
- Broadcasts published lists of Tutsi names and asked people to provide information on those listed, look for those who were in hiding, describe the hiding places, and to kill them

ICTR: Direct and Public Incitement to Genocide

- Only the most inflammatory broadcasts – those made after the genocide began on April 6 – contributed substantially to the killing of numerous Tutsi and could incur liability
- Relationship with hate speech:
 - Hate speech that does not directly incite genocide, or that is made without the requisite genocidal intent, is not actionable
 - Nevertheless, hate speech that can appear ambiguous at first sight may amount to direct incitement to commit genocide in a particular context
 - The fact that genocide took place may be considered as one of the indicia that, in a given context, the speech was perceived as an incitement
- Individual criminal responsibility:
 - Nahimana: **Guilty**, because he knew or had reasons to know that his subordinates at RTLM over whom he had effective control were about to broadcast or had already broadcast utterances inciting the killing of Tutsis and failed to prevent or punish them
 - Barayagwiza: **Acquitted**, because he did not exercise effective control over RTLM staff

ICTR: Persecution as a Crime against Humanity

- RTLM broadcasts before April 6, 1994 could not constitute acts of persecution because they were not part of a systematic and widespread attack -- the predicate elements of crimes against humanity – nor did they substantially contribute to the killing of Tutsis
- RTLM broadcasts after April 6 which substantially contributed to the killing of Tutsis constituted acts of persecution
- Relationship between hate speech and persecution as a crime against humanity:
 - Hate speech targeting a population on ethnic grounds or on any other discriminatory grounds violates the right to the respect of human dignity of members of the group
 - Hate speech cannot on its own amount to a violation of the right to life, freedom, and physical integrity
 - In Rwanda, hate speech went beyond advocating hatred – it was accompanied by calls for genocide against the Tutsi, and was part of a massive campaign characterized by acts of violence and destruction of property directed against the Tutsi population → acts of persecution
- Individual criminal responsibility:
 - Nahimana: Convicted
 - Barayagwiza: Acquitted

IIFFM: Facebook in Myanmar



- Beginning August 2017 in Rakhine state, Myanmar military and local police initiated a brutal campaign targeting Rohingya Muslim civilians, inflamed by exclusionary and discriminatory rhetoric
- Rohingya were subject to large-scale massacres of men, women, and children with execution-style killings, burning of entire villages, and sexual violence
- Within a matter of weeks, nearly 1 million people fled across the border to Bangladesh
- UN Independent International Fact-Finding Mission on Myanmar (IIFFM) was created in 2017 to establish the facts and circumstances of the alleged human rights violations by military and security forces

IIFFM: 2018 Report Findings



- Facebook played a key role in the incitement of violence against the Rohingya
- Facebook was most common social media platform in use in Myanmar and the main mode of communication – for many, Facebook *was* the internet
- Government and military officials regularly used Facebook to release news and information, causing many to believe information posted on Facebook could be trusted
- Facebook’s platform contained inflammatory messages portraying Rohingya as violent, dishonest, anti-Buddhist, illegal immigrants and/or terrorists
- Death threats, incitement to violence and discrimination, and online harassment were common features both against the Rohingya themselves and against moderate commentators, human rights defenders and ordinary people with differing views

IIFFM: 2018 Report Recommendations



- Facebook’s efforts to eliminate hate speech and halt the spread of misinformation fell far short of what was necessary:
 - Facebook community standards and user agreement policies -- including rules relating to hate speech and violence -- did nothing to halt the use of hate speech on the platform
 - Facebook failed to sufficiently address “fake accounts and false news” with intervention and prevention
 - Facebook’s content moderation was inadequate, lacked transparency, and lacked sufficient local experts that could interpret and contextually understand Myanmar language and culture
 - Facebook failed to undertake a comprehensive human rights impact assessment in Myanmar
 - Facebook lacked any country-specific data regarding the spread of hate speech on its platform

IIFFM: Facebook's Response

- Facebook has since admitted it was “too slow” in responding to events in Myanmar and that its platform was used to incite violence
- A Facebook-commissioned human rights impact assessment report (the HRIA) largely placed blame on a “low digital literacy rate” among the local population, general lack of government accountability, and local, deep-rooted systemic discrimination
- The HRIA Report also appeared to paint Facebook as a helpless victim of both its own success and “bad actors” who sought to manipulate Facebook’s “neutral” communication tools for their own nefarious ends
- Noting that Facebook had increased its number of Myanmar language experts to 60 as of August 2018 — a full year after the mass atrocities occurred — the HRIA Report also cautioned that *“[i]nvestment in local staff could raise expectations that Facebook will take a similar approach in other countries”*
- → The IIFFM turned over its evidence in 2019 to the Independent Investigative Mechanism for Myanmar (IIMM), which is preparing dossiers for eventual prosecution
- → Facebook is cooperating with the IIMM

Conclusions

- There is a decades-long history of international criminal accountability for mass media publishers who promote and disseminate discriminatory hate speech that incites violence
- Those convicted to date have themselves espoused and endorsed such hate speech, actively utilizing the mass media they controlled for that purpose
- Social media companies were well aware of (1) the potential for serious, widespread violence as a result of speech disseminated on their platforms, and (2) steps they could and should have taken to intervene quickly and prevent such violence in Artsakh
- The IFFM's focus on Facebook's central role in the Rohingya massacres may signal a turning point in international attempts to hold social media platforms accountable