International Approaches to Regulating Hate Speech Online: Brief Submitted to the Standing Committee on Justice and Human Rights (May 20, 2019)

Dr. Heidi Tworek, Assistant Professor, University of British Columbia, Vancouver

Three years ago, the question was *if* social media would be regulated. Now the questions are *how* and *when* they will be regulated. Even Facebook CEO Mark Zuckerberg called for greater regulation of "harmful content" on March 30, 2019. A key component of these debates is dealing with hate speech online. Democracies around the world are only just grappling with how to enforce existing hate speech law online or whether they need to update existing laws or whether they should be taking non-judicial regulatory approaches. Even in Europe, governments have proposed different solutions. Here, I briefly survey a representative sample of initiatives/proposals and relevant analyses. I also include proposals and initiatives dealing with terrorist/violent extremist content as they often overlap with hate speech.

Germany: Netzwerkdurchsetzungsgesetz (NetzDG, Network Enforcement Law)

In force since January 2018, this law enforces 22 extant statutes of German speech law online. It requires social media companies with more than two million unique users in Germany to respond to posts flagged by users under NetzDG within 24 hours or face fines of up to 50 million Euros per post. The law requires regular transparency reports from companies that receive more than 100 complaints (currently just Facebook, YouTube, Twitter, and Change.org). This law is arguably the most ambitious and first major attempt by a democracy to push social media companies to enforce domestic law.

Analysis of NetzDG by Heidi Tworek, Assistant Professor at University of British Columbia, and Paddy Leerssen, PhD candidate at the Institute for Information Law, University of Amsterdam (April 2019).

UK: White Paper on Online Harms

In April 2019, the UK released a white paper rethinking the British approach to social media companies. It argues that social media companies hold a "duty of care" and must design their systems to prevent online harms. The white paper proposes that a new regulator oversee if social media companies are meeting the standards of duty of care. The paper is in the consultation period until July 2019.

Online Harms White Paper (April 2019). For a critical analysis, see Graham Smith's blog post (April 2019).

France: Accountability by Design

In May 2019, France issued a first report that proposed a regulator who would enforce transparency and accountability from the largest social media companies. The idea is to create an *ex ante* regulator who will enable greater transparency from the companies and more involvement from civil society. The proposal followed a unique experiment where French civil servants were embedded at Facebook for several months.

Creating a French Framework to Make Social Media Platforms More Accountable (in French and in English, May 2019).

Australia: Sharing of Abhorrent Violent Material Bill

Passed swiftly after the Christchurch shooting, this is an amendment that now holds content and hosting service providers liable if they do not remove videos showing "abhorrent violent conduct" expeditiously or if they fail to notify the Australian federal police about them. Abhorrent violent conduct is defined as terrorist acts, murders, attempted murders, torture, rape, or kidnapping. The eSafety Commissioner also received new powers to notify companies, if they are hosting "abhorrent violent" material, thus triggering a take-down requirement. The Australian attorney-general described the law as "most likely a world first." There are concerns about the speed and process of passage, safeguards for freedom of expression, the penalties, and speed requirements for take-downs.

Sharing of Abhorrent Violent Material Bill (April 2019). For a critical analysis of the bill, see a Lawfare post by Evelyn Douek, SJD candidate at Harvard Law School and former clerk for the Chief Justice of the High Court of Australia (April 2019).

European Union: Terrorist Content Regulation (TERREG)

In September 2018, the European Commission proposed this legislative measure to require a one-hour removal time for "terrorist content" as well as proactive monitoring of content by social media companies. The proposal is currently under debate in the European Parliament and the Council.

Analysis of TERREG by Joris van Hoboken, Professor of Law at the Vrije Universiteit Brussels (VUB) and a Senior Researcher at the Institute for Information Law, University of Amsterdam (April 2019).

Global Internet Forum for Counter-Terrorism (GIFCT)

GIFCT is a private, information-sharing collaboration between social media companies to share "hashes" of terrorist imagery and videos. The collaboration played a key role after the Christchurch shooting in trying to enable companies to remove the millions of videos reproducing the livestream of the shooting.

Analysis of GIFCT by Brittan Heller, Fellow at the Carr Center for Human Rights Policy, Harvard Kennedy School (April 2019).

Social Media Councils

This suggestion has emerged from civil society and researchers, including myself, in various formats. Social media councils would be multi-stakeholder fora, convened to address online content moderation. The exact format and geographical scope remain up for debate, though my co-authored report from November 2018 suggested some specific paths forward for Canada.

Social Media Councils: From Concept to Reality, Stanford Global Digital Policy Incubator, ARTICLE 19, and David Kaye, UN Special Rapporteur on the Right to Freedom of Opinion and Expression (February 2019).

Poisoning Democracy? How Canada Can Address Harmful Speech Online by Dr. Chris Tenove, postdoctoral research fellow at University of British Columbia, Dr. Heidi Tworek, Assistant Professor at University of British Columbia, and Dr. Fenwick McKelvey, Associate Professor at Concordia University (November 2018).