



Cyber Security Research Center – Cyber Law Program

Inadvertently Appointing Digital Judges? A Canadian Perspective on Restricting Speech and Social Media

Prof. Karen Eltis

Published: October 26th, 2017

On the heels of Charlottesville, questions respecting platforms' responsibility for hateful content abound. Presumably recognizing that the internet has effectively undermined the formerly sacred "marketplace of ideas" paradigm, Facebook and similar platforms have now committed to suppressing "pernicious form[s] of harassment" and so-called "fake news" sites. And so it came to be that private American companies are reluctantly but surely stepping into the role of international arbiters of free expression, increasingly employing artificial intelligence to contain content that is potentially "offensive" to advertisers or otherwise disturbing.

Without a doubt, we are witness to a culture of instantaneous sharing, enabling anyone to communicate random thoughts potentially with everyone worldwide, without any ability to correct, retract, control or contextualize subsequent dissemination. They may further – to a certain degree – do so anonymously. What is more, a culture of this ilk may, in Cass Sunstein's words, create a "situation in which thousands or perhaps millions or even tens of millions of people are mainly listening to louder echoes of their own voices." These refers to the confirming and reinforcing of biases and half-truths on issues as diverse as climate change and vaccination.

The Internet has certainly and laudably democratized expression, shifting the narrative from freedom *from* government muzzling to freedom *to* express oneself to a far broader audience, irrespective of financial means or social status. Prior to the digital age, only the moneyed or otherwise influential could dream of reaching a significant audience, thereby arguably emasculating any purposive construction of free expression enshrined in a constitution prizing equality. The internet has become the proverbial "soapbox" of the 21st century, providing everyone and anyone with what may be deemed an optimal platform for expression. But again, unlike their predecessors, these new media sources are subject to little, if any, oversight, scrutiny or accountability (save AI, it now appears).

In contradistinction to the institutional press, which typically boasts the use of built-in safeguards (editorial oversight, and a civil if not purportedly neutral tone), posters and tweeters can share information in a manner that radically compounds the difficulties inherent in traditional defamation, including punitive shaming. This may ultimately force us to recalibrate the balance between free speech – regarded as the "very life blood of our freedom and free institutions" – and competing considerations.

It is indeed the medium's very structure that tends to bestow the appearance of legitimacy and veracity on even the most mendacious of racist sites, in the absence of gatekeepers or other traditional controls. Therefore, as a medium, it may help legitimate the most pernicious forms of hate and incitement, if only due to the arduous task of distinguishing between reliable, authoritative cyber sources, and those peddling racism and fabrications under the unique guise of respectability imparted by the net environment.

The post-war human rights legal framework, which includes Canada's Charter of Rights, was devised with government action in mind, as a bulwark against government abuses. Yet "government action," however generously interpreted, requires at least a measure of just such abuse.

It is apparent, however – and increasingly so of late – that the power to infringe on constitutional values, however inadvertently, including (but not limited to) freedom of expression and privacy, or due process as presumably in the Loomis case, does not lie principally – or at the very least exclusively – with the state. Rather, in a post-Costeja (ECJ Right to be Forgotten decision) world, private parties, namely "data controllers" with global influence, using Artificial Intelligence, have become the unwitting arbiters of the global public interest, a role they presumably neither covet nor properly fit, especially in the absence of guidance or criteria.

Facebook and other platforms are thus incomprehensibly saddled with the gargantuan task of determining how to **"balance the need for transparency with the need to protect people's identities."** This inevitably leads to *ad hoc* approaches by these companies. In addition, transparency and accountability are notoriously difficult to cultivate when balancing delicate constitutional values, such as freedom of expression and privacy. Even the constitutional courts and policy makers who typically perform this balancing struggle with it, as for example in the context of the controversy associated with so-called "judicial activism." This difficulty soars when the balancers are instead inexperienced and reticent corporate actors, who presumably lack the requisite public legitimacy for such matters, especially when dealing with foreign (non-U.S.) nationals.

This leads in turn to absurd results, such as the suppression of the picture of a 1972 depiction of a Canadian-Vietnamese child war victim by Facebook's algorithms, subject only to the immediate oversight of outsourced corporate actors abroad. Another example is PayPal's algorithm marginalizing cookbooks featuring the terms "Syria" or "Cuba" in a misguided effort to comply with security regulations.

Mindful of the above, the notion of "government action" might be purposefully revisited with a view to its adaption to the exigencies of this digital age. If this does not take place, the ultimate arbiters of the proper limits on fundamental rights may be algorithms or other forms of Artificial intelligence deployed by platforms that can be assumed to lack judicial training, not to mention cross-border accountability. As we struggle with delineating the proper limits on speech in a post-Charlottesville world, let us be cognizant of the importance of maintaining courts' oversight on constitutional values and the proper limits on expression, rather than letting them become the province of the unknown.

