

TIME TO REGULATE INTERNET HATE WITH A NEW APPROACH

- **Cyber-racism is more than a civil wrong, it is a public wrong against the community**
- **Corporations who seek to profit from social media assume public obligations. This includes an obligation to take reasonable steps to discourage online racism.**
- **Both racism itself and the failure to sufficiently regulate against it must have legal consequences. Regulation is both legally and technically possible.**

Andre Oboler* - Co-Chair, Online Antisemitism Working Group, Global Forum to Combat Antisemitism and Director of the Community Internet Engagement Project at the Zionist Federation of Australia.

INTRODUCTION

The rise of social media over the last seven years¹ has revolutionised communication. Applications like Facebook, YouTube, Twitter, and MySpace, have given the online public a means of mass communication. Behind these platforms are large corporations² who profit from the communication they facilitate, yet take little responsibility over the content. A new approach is needed in which corporations that seek to profit from social media assume public obligations. This includes an obligation to take reasonable steps to discourage online hate, a public wrong against the community.

In a recent issue of this bulletin, Dilan Thampapillai argued that Facebook should be responsible for policing the nature of the material it hosts but that given the volume of users, "internal policing will be imperfect" and competing business priorities would interfere.³ Ignoring for a moment the logistical issue (a technical and not a legal impediment), both arguments hold equally true for the sale of alcohol in licensed venues. The solution there has been an obligation on licensees to police their sales internally, even to the detriment of their core business of further sales. This is not an unusual position, and in the contest between personal interests and interests of others as protected by law, the law must prevail. For this to happen we need a new law, as we have with the sale of alcohol, that places a level of responsibility on the commercial operator.

I start by discussing why Online Hate should be considered a public wrong and show that it is deemed unlawful under Commonwealth legislation and criminal under Victorian legislation. I then discuss the value of the law in setting standards. I examine two key developments internationally, one involving Yahoo in France and the other Google in Italy. In discussing the need for a new approach I highlight how regulation of the platform providers can be a

* Director of the Community Internet Engagement Project at the Zionist Federation of Australia and Co-Chair of the Online Antisemitism Working Group at the Global Forum to Combat Antisemitism. PhD in Computer Science (*Lancaster University*), B. Comp. Sci. (Hons) (*Monash University*). Dr Oboler is currently a JD candidate at Monash University. Website: www.oboler.com, e-mail: andre@oboler.com

¹ Myspace was founded in August 2003, Facebook in February 2004, YouTube in 2005, and Twitter in 2006

² Facebook and Twitter are independently owned and each worth billions. YouTube is owned by Google, and My Space is owned by News Corporation.

³ Dilan Thampapillai, 'Facebook or Racebook? Time for a Different Approach?' (2010) 13(2) *Internet Law Bulletin* 30.

significant contribution. Finally I provide concrete suggestions to aid both regulation and implementation. The deficiencies of current legal definitions of various forms of hate speech are beyond the scope of this paper, I rather focus upon the consequences of this hate and the need for new approaches.

ONLINE HATE AS A PUBLIC WRONG

Hate propaganda is speech which “seeks to incite and encourage hatred and tension between different social and cultural groups in society”.⁴ It aims to provide a “foundation for the mistreatment of members of the victimised group”.⁵ Members of the victim group are humiliated, denigrated, have their self-worth undermined, and may ultimately withdraw from full participation in society out of fear.⁶ Prof Kathleen Mahoney explains that society suffers as hate propaganda “undermines freedom and core democratic values by creating discord between groups and an atmosphere conducive to discrimination and violence”.⁷

Mahoney rejects the absolutist position for free speech as “seriously and fundamentally flawed”. In questioning their underlying theory and practical value, she highlights that they are based on outdated theories, ignore the harm caused to the victim group, show both gender and class bias and seek to uphold free speech to the detriment of other fundamental democratic values.⁸

In the specific case of antisemitism, one of the steps taken to aid law enforcement agencies in identifying this hate was the creation of the “Working Definition of Antisemitism” by what is now the European Union Agency for Fundamental Rights. The document provides examples of activities that cross the line, such as “Holding Jews collectively responsible for actions of the state of Israel” or “Accusing the Jews as a people, or Israel as a state, of inventing or exaggerating the Holocaust.”⁹ At the same time it states “criticism of Israel similar to that levelled against any other country cannot be regarded as antisemitic.”¹⁰ Such tools help law enforcement evaluate materials presented to them, and can also aid the public understanding of particular forms of hate.

The publication and spread of hate messages is a public wrong, not simply a matter for reconciliation between private persons. A broadly targeted wrong deserves a public response, including through education and law enforcement.

AUSTRALIA LEGAL RESPONSES TO VILIFICATION AND HATRED

⁴ NKS Banks, ‘Could Mom be Wrong? The Hurt Of Names And Words: Hate Propaganda And Freedom Of Expression’ (1999) 6(2) *Murdoch University Electronic Journal of Law* 2. Available online <<http://www.austlii.edu.au/au/journals/MurUEJL/1999/20.html>>

⁵ Kathleen Mahoney, *Hate Vilification Legislation with Freedom of Expression: Where is the Balance?* (1994) 9.

⁶ *Ibid.*

⁷ *Ibid.*

⁸ *Ibid.*, 11.

⁹ Working Definition of Antisemitism, The European Union Agency for Fundamental Rights. Available online <<http://fra.europa.eu/fraWebsite/material/pub/AS/AS-WorkingDefinition-draft.pdf>>

¹⁰ *Ibid.*

The most developed area of vilification and hatred law relates to race. Other areas such as homophobia and vilification based on disability have yet to receive the same level of protection. The regulatory models to counter racial vilification and hatred include criminal prosecution, statutory tort civil suit, and civil human rights complaint.¹¹ Even in this more developed area, the existing approaches suffer from flaws limiting their capacity to provide protection and legal remedy.¹²

The Commonwealth position

Australia's first attempt to legislate against racial hatred was cl 28 of the *Racial Discrimination Bill* 1974 (Cth). Clause 28 made publishing, broadcasting or public utterance promoting ideas of racial hatred a crime punishable by a \$5,000 penalty.¹³ Approved by the House of Representatives, this clause was rejected by the Senate, and then deleted prior to enactment of the *Racial Discrimination Act* 1975 (Cth).¹⁴ A far weaker provision against racial hatred was eventually inserted into the Act by the *Racial Hatred Act* 1995 (Cth).

The *Racial Hatred Act* 1995 (Cth) makes it illegal to do an act, other than in private, if

- (a) the act is likely, in all the circumstances, to offend, insult, humiliate or intimidate another person or a group of people; and
- (b) the act is done because of race, colour, or national or ethnic origin of the other person or of some or all of the people in the group".¹⁵

A number of exemptions exist for acts done reasonably and in good faith.¹⁶ Ultimately the Act provides for civil sanction, not criminal prosecution.

The provision lacks force as s 26 of the Act deems these unlawful acts not to be offences, a significant downgrade from the position of an offence with a statutory penalty as proposed in clause 28. The Human Rights and Equal Opportunities Commission has recommended a strengthening of the provisions, proposing in 2001 that criminal sanctions be provided for racial discrimination, vilification and hatred in all Australian jurisdictions.¹⁷ An effective criminal prosecution approach can play a strong educative role. One state to have criminal provisions is Victoria.¹⁸

The Victorian Position

The *Racial and Religious Tolerance Act* 2001 (Vic) proscribes two forms of racial vilification. Racial vilification (s7) is "conduct that incites hatred against, serious contempt

¹¹ Luke McNamara, *Regulating Racism: Racial Vilification Laws in Australia* (2002) 308.

¹² *Ibid.*

¹³ Commonwealth, *Parliamentary Debates*, House of Representatives, 9 April 1975, 1408. Seen in Luke McNamara, *Regulating Racism: Racial Vilification Laws in Australia* (2002) 35.

¹⁴ Luke McNamara, *Regulating Racism: Racial Vilification Laws in Australia* (2002) 35.

¹⁵ *Racial Hatred Act* 1995 (Cth) s18C(1).

¹⁶ *Racial Hatred Act* 1995 (Cth) s18D.

¹⁷ Acting Race Discrimination Commissioner, "I want Respect and Equality" *A Summary of Consultations with Civil Society on Racism in Australia* (2001) 47.

¹⁸ Criminal sanctions exist in Victoria, New South Wales, South Australia, Queensland, and Western Australia,

for, or revulsion or severe ridicule of” a person or class of person on the basis of their race, and it is deemed unlawful.¹⁹ Serious racial vilification (s24) is said to occur when a person “intentionally engage[s] in conduct that the offender knows is likely” to incite hatred and “threaten, or incite others to threaten, physical harm” against the victim or their property.²⁰ Serious racial vilification can result in six months imprisonment. The Act expressly states that conduct includes use of the Internet.²¹

The Victorian Act entirely excludes private conduct from unlawful racial vilification. Private conduct occurs in “circumstances that may reasonably be taken to indicate that the parties to the conduct desire it to be heard or seen only by themselves” provided the circumstances are such that it is reasonable to expect it would not be “heard or seen by someone else”.²² Public conduct is excluded in limited circumstances and along similar lines to those in the Commonwealth legislation.²³ No exception exists for serious racial vilification.

The Victorian Act uses similar language to s 7 to deem religious vilification unlawful (S8) and similar language to s 24 to designate serious religious vilification an offence (S25).

THE SIGNIFICANCE OF THE LAW IN THE WEB 2.0 WORLD

The fact that something is unlawful provides sufficient moral grounds to require corporations to take action to discourage it, particularly in circumstances where they would otherwise be profiting from the activity. These moral grounds can be converted into a legal imperative by enacting new laws that specify which types of unlawful speech must be reasonably discouraged by providers and removed within reasonable time. Social media providers already act of their own volition to remove material such as nudity and potential copyright infringing materials. The lack of will when it comes to hate speech largely derives from the protection afforded in the United States. The extreme protection of the first amendment, which extends to protect hate speech, is a US anomaly and one that puts the United States out of step with both the legal position and international expectations.²⁴ There is a clash of values. This was best seen in the response of Facebook to accusations it was facilitating Holocaust denial.²⁵

The clash of values has resulted in US companies acting globally, facilitating wrongful acts locally in a third country, yet seeking to avoid the implications of the effected countries national laws. This imperialist approach has seen a strong reaction in Europe. The idea that the internet is a zone outside the law has been clearly rejected in the Australian legislation. This is not unusual, and national anti-hate laws in other countries have led to law suits and

¹⁹ *Racial and Religious Tolerance Act 2001 (Vic)*, S7

²⁰ *Racial and Religious Tolerance Act 2001 (Vic)*, S24

²¹ The note appears in S7, S8(1), S24(1), S24(2), S25(1), and S25(2)

²² *Racial and Religious Tolerance Act 2001 (Vic)*, S12

²³ *Racial and Religious Tolerance Act 2001 (Vic)*, S11

²⁴ International Covenant on Civil and Political Rights art 19. Available online <www.unhchr.ch/html/menu3/b/a_ccpr.htm>

²⁵ Andre Oboler, 'Facebook, Holocaust Denial, and Anti-Semitism 2.0', 15 September 2009, Post-Holocaust and Anti-Semitism Series, No. 86, The Jerusalem Center for Public Affairs. Available online <<http://tinyurl.com/facebookholocaustdenial>>

changes in the way internet companies operate. Laws need to be developed that ensure social media companies do not wilfully facilitate the unlawful spread of racist hate within Australia.

YAHOO! AND GOOGLE MEET THE FRENCH AND ITALIAN JUSTICE SYSTEMS

In 2000 two French NGOs filed suit against Yahoo! in the French courts.²⁶ The suit was brought against the sale of Nazi memorabilia to people in France through the Yahoo! website, in contravention of French law. The French courts ruled Yahoo! was at fault and demanded remedial action. The response, played out in the French and American courts provided the test case for internet regulation.

Ronald S. Katz, a lawyer for the NGOs explained, "There is this naive idea that the Internet changes everything... It doesn't change everything. It doesn't change the laws in France."²⁷ Cyber libertarians such as Nicholas Negroponte countered, "It's not that the laws aren't relevant; it's that the nation-state's not relevant. The internet cannot be regulated."²⁸ The issue was finally settled before the Ninth Circuit of United States Court of Appeals.²⁹

The appeal by the NGOs was against a US District Court ruling which held the French Court's ruling could not be enforced in the US due to the First Amendment. The appeal overturned this ruling holding that the district court did not have personal jurisdiction over the French parties. It also noted that "France is within its rights as a sovereign nation to enact hate speech laws" and that the NGOs were "within their rights to bring suit in France against Yahoo! for violation of French speech law." Yahoo! the court found, was being forced to wait for the NGOs to file a US suit for enforcement before the opportunity to raise a First Amendment claim could arise. While noting that this allowed the fine to increase daily, the court held that "it was not wrongful for the French organizations to place Yahoo! in this position".³⁰ The final outcome makes clear the validity of national hate speech laws outside the US, but leaves in question their enforceability within the US.

The French case led to changes in Yahoo!'s approach. When initially refusing to comply, Yahoo!'s CEO stated, "We are not going to change the content of our sites in the United States just because someone in France is asking us to do so."³¹ Later Yahoo! banned the sale of Nazi memorabilia not only in France, but globally. The CEO stated, "we have to comply with local law".³²

In 2008 American Internet lawyer, Christopher Wolf, noted that the US courts have begun to chip away at the immunity Congress gave internet service providers from liability for actions

²⁶ Jessie Daniels, *Cyber Racism: White Supremacy Online and the New Attack on Civil Rights* (2009) 176-77.

²⁷ Lisa Guernsey, 'Welcome to the World Wide Web. Passport, Please?', *The New York Times*, March 15, 2001. Available at < <http://www.nytimes.com/2001/03/15/technology/15BORD.html> >

²⁸ Jack Goldsmith and Tim Wu. *Who Controls the Internet? Illustrations of a borderless world* (2006) 3.

²⁹ 379 F3d 1120 Yahoo! Inc v. La Ligue Contre Le Racisme et L'Antisemitisme. Available online at <<http://openjurist.org/379/f3d/1120/yahoo-inc-v-la-ligue-contre-le-racisme-et-lantisemitisme>>

³⁰ 379 F3d 1120 Yahoo! Inc v. La Ligue Contre Le Racisme et L'Antisemitisme. Available online at <<http://openjurist.org/379/f3d/1120/yahoo-inc-v-la-ligue-contre-le-racisme-et-lantisemitisme>>

³¹ Jack Goldsmith and Tim Wu. *Who Controls the Internet? Illustrations of a borderless world* (2006) 5.

³² Jack Goldsmith and Timothy Wu, 'Digital Borders', *Legal Affairs*, January / February 2006. Available at < http://www.legalaffairs.org/issues/January-February-2006/feature_goldsmith_janfeb06.msp >

of their users. He warned that "the day may come, if a trend continues, where the potential for legal liability for tortious speech of others may compel ISPs and web sites to more actively monitor what goes out through their service."³³

That day arguably came in February 2010, when three Google Executives were, in absentia, given six-month suspended jail terms by an Italian court for breeches of privacy legislation.³⁴ The ruling followed a failure to remove a Google Video of a boy with autism being bullied. While Google argued it removed the video within hours of being formally notified by police, two months had passed since users first reported the video to Google.³⁵ In the mean time, the video had become one of the most viewed Google Videos in Italy. Prosecutors argued the lack of action before the police notification amounted to negligence.³⁶ The case is subject to appeal.

Facing the threat of new internet legislation in Italy, Facebook, Google and Microsoft recently agreed to implement a shared code of conduct.³⁷ The change in attitudes shows that governments can both regulate the internet giants and encourage improvements in self regulation.

THE NEED FOR A NEW SYSTEMATIC APPROACH

The Human Rights and Equal Opportunity Commission (HREOC) noted in 2001 that the burden placed on individuals to bring complaints was "often insurmountable" and that "additional mechanisms for grievance processes which go beyond the individual complaint system" may need legislative change.³⁸ The Commission recommended that organisations be permitted to initiate representative complaints, and that HREOC commissioners be allowed to self-initiate complaints.³⁹ There are, however, other additional ways to go beyond the individual complaint system.

Luke McNamara argued it is "behavioural changes and attitudinal shift" which offer the best protection against racial hatred in the long term.⁴⁰ A new approach should aim to make racism not only illegal but social unattractive. The platform providers are the key to making this happen online.

³³ Christopher Wolf, 'The Dangers Inherent in Web 2.0' (Remarks to the Westminster eForum Keynote Seminar: Web 2.0 -- Meeting the Policy Challenges of User-Generated Content, Social Networking and Beyond, London, UK, April 2008). Available at <www.adl.org/main_internet/Dangers_Web20.htm>

³⁴ Manuela D'Alessandro, 'Google executives convicted for Italy autism video', *Reuters*, 24 February 2010. Available online <<http://www.reuters.com/article/idUSTRE61N2G520100224>>

³⁵ Rachel Donadio, 'Larger Threat Is Seen in Google Case', *New York Times*, 24 February 2010. Available online <<http://www.nytimes.com/2010/02/25/technology/companies/25google.html>>

³⁶ John Hooper, 'Google executives convicted in Italy over abuse video', *The Guardian*, 24 February 2010. Available online <<http://www.guardian.co.uk/technology/2010/feb/24/google-video-italy-privacy-convictions>>

³⁷ Manuela D'Alessandro, 'Google executives convicted for Italy autism video', *Reuters*, 24 February 2010. Available online <<http://www.reuters.com/article/idUSTRE61N2G520100224>>

³⁸ Acting Race Discrimination Commissioner, "I want Respect and Equality" *A Summary of Consultations with Civil Society on Racism in Australia* (2001) 47.

³⁹ *Ibid.*

⁴⁰ Luke McNamara, *Regulating Racism: Racial Vilification Laws in Australia* (2002) 312.

Reasonable steps that platform owners could take range from removing content, through to disabling the accounts of problem users. Other options include suspending accounts for a fixed period of time and revoking privileges. For example, a user who creates a group that incites hate might have their group shut down and additionally lose the ability to become a group administrator. Repeated problems might see them eventually banned, a step that provides punishment, deterrence, and improved safety for the rest of the community. In some cases, internal sanctions are not sufficient and a matter might be handed over to the relevant authorities in the user's country of residence. Only a legal obligation to take reasonable steps in reasonable time will ensure sufficient effort is invested in responding effectively to hate.

NEW FORMS OF REGULATION IN THE ONLINE ENVIRONMENT

The Online Antisemitism working group of the Global Forum to Combat Antisemitism met in December 2009 and outlined a number of recommendations, many of them generic to combating online hate.⁴¹

One recommendation was that carrier immunity "is too broad and needs to be limited in the case of antisemitism and other forms of hate."⁴² The report explained, "While real time communication may be immune, stored communication e.g. user published content, can be brought to a service provider's 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."⁴³

The question of reasonable time was addressed through a model dividing online platforms into categories and suggesting different rules could apply to each. The categories differentiate between websites without user input, and those with user input that are unmoderated, exception moderated, post-moderated, or pre-moderated. The higher the investment to prevent incitement, the longer companies should have to handle complaints. This would encourage more systematic vigilance by providers.

In Australia new laws would need to: (a) state what is a reasonable time for each category; (b) allow complaints registered with the platform to be separately registered with a third party such as HREOC or an online ombudsman; and (c) allow the third party to take action on complaints that are both beyond reasonable time and considered by the third party to be valid.

Under such a scheme platform providers would act to reduce their liability and create preventative mechanisms. One such mechanism would be attaching obligations to some user privileges. The administrator of a Facebook group might, therefore, become the first port of call for complaints about content in their group. Failure to respond in reasonable time could result in a loss of administration privileges. An administrator who marks content that has been complained about as acceptable, might find their decision reviewed (by staff or elected

⁴¹ Andre Oboler and David Matas, 'Working Group 3 - Antisemitism Online: Cyberspace and the Media' (co-chairs report from Working Group 3 at the Global Forum to Combat Antisemitism, Jerusalem, 16-17 December 2009). Available online <<http://tinyurl.com/onlineantisemitismreport>>

⁴² Ibid.

⁴³ Ibid.

community members) if complaints continue. If they make too many mistakes, their position as an administrator might be revoked. Such a system would dramatically decrease the number of complaints Facebook itself needed to review. In a similar manner, YouTube users could get initial complaints on videos they posted. The technical problems are not insurmountable if the companies find it is in their interest to act.

CONCLUSION

To combat online hate, we need new laws that enable a more proactive response. We need laws that ensure online spaces enjoy the same protection as physical spaces. The current situation is that companies create terms of service, as they see fit, and are then able to apply them as they see fit. In some cases hateful content lasted over six months before finally being removed.⁴⁴ In 2009 Facebook removed the prohibition against posting “racially, ethnically or otherwise objectionable” content from its terms of service.⁴⁵ It kept the prohibition on “hateful content”, but then claimed Holocaust denial, an example provided in the EUMC Working Definition, would not qualify.⁴⁶ As a public wrong, hate speech should be receiving a public response. It is not enough to leave it to the discretion of powerful social media companies. 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.

⁴⁴ See above, n 25.

⁴⁵ Ibid.

⁴⁶ Ibid.