

Regulating hate speech and freedom of expression on the Internet: Promoting tolerance and diversity

F CASSIM*

ABSTRACT

The development of new accessible technologies and the expansion of the Internet have changed the informational and communicational realities across the globe. The Internet has become a vital and accessible forum for free speech and a marketplace for the exchange of ideas. Online users rely on the safety, security and anonymity of the Internet and social media in their daily lives. However, the Internet has a potential for misuse and abuse. Hate speech involves the use of abusive, racist and disparaging comments, words or phrases directed against particular groups of people. The anonymity of the Internet has facilitated the dissemination of hate speech with such speech becoming more prevalent. Online hate speech has a significant impact on the lives of the people it seeks to denigrate. This article examines the tension between hate speech and freedom of expression on the Internet. This study reveals that online hate speech has become rife, and many countries have introduced laws placing restrictions on freedom of expression in order to curb online hate speech. It is submitted that limits on freedom of expression can affect the free exchange of ideas and information, and erode confidence in the network technologies that are used. It is recommended that appropriate mechanisms should be put in place to preserve the use of the Internet as a marketing, communication and educational tool, and at the same time teach online users to embrace 'pluralism, tolerance and broadmindedness'. A collaborative effort by all stake holders (such as governments, non-governmental organisations, Internet Service Providers, international organisations) is also necessary to curb hate speech on the Internet and to promote tolerance and respect for diversity. The Internet should be preserved as an accessible forum for free speech and the exercise of freedom of expression should be exercised in a responsible manner.

* BA (Law) (University Durban-Westville) LLB (University of Natal – Durban) LLM LL.D (University of South Africa). Associate Professor in the Department of Criminal and Procedural Law, University of South Africa (UNISA), admitted attorney and conveyancer of the High Court of South Africa.

1 Introduction

The Internet has revolutionised information and communication realities worldwide, introduced an accessible market for the sale of products and the exchange of ideas and introduced new platforms for social interaction, such as Twitter and Facebook.¹ However, the development of new accessible technologies and the expansion of the Internet have created new opportunities for criminal and extremist activities.² The rise of the Internet has resulted in the dissemination of hate speech and cyber hate-related activities becoming widespread, with an increase in the number of websites offering racist content and hate speech being reported recently.³ The publication of such materials has aroused much public concern from international bodies.⁴ There has also been a reported increase in the number of South Africans who make racist comments on social networking sites such as Twitter and

¹ See D Moeckli, S Shah, S Sivakumaran and D Harris (ed) *International Human Rights* 2ed (2014) 431-432; T McGonagle 'Minorities and online 'hate speech': A parsing of selected complexities' (2010) 9 *Euro'n Yearbk Minority Issues* 419 at 431-432; YA Timofeeva 'Hate speech online: restricted or protected? Comparison of regulations in the United States and Germany' (2002-2003) 12 *J Transnat'l Law & Pol'y* 253 at 255; EM Chen 'Global Internet freedom: Can censorship and freedom co-exist?' (2003) 13 *DePaul J Art & Ent Law* 229 at 237-238; B Perry and P Olsson 'Cyberhate: The globalisation of hate' (2009) 18 *Info & Comms Technology L* 185 at 187; A Tsesis 'Hate in cyberspace: Regulating hate speech on the Internet' (2001) 38 *San Diego L R* 817 at 818 and R Delgado and J Stefancic 'Hate speech in cyberspace' (2014) 49 *Wake Forest L R* 319 at 323-324.

² New technologies have also simplified the task of criminals to commit existing crimes (such as *inter alia*, hate speech) which can be generated and spread by a computer in a matter of seconds to millions of computer users around the globe. See MS Nuth 'Taking advantage of new technologies: For and against crime' (2008) 24 *Comp Law & Sec Rev* 437 at 437-438. It should be emphasised at the outset that hate speech would incorporate racist speech. See further, the definition of online hate speech in section 2 below.

³ M Gercke *Understanding Cybercrime: A Guide for Developing Countries* 2ed (2011) 68. Also see DP van der Merwe et al *Information and Communications Technology Law* (2008) 444-445; D Thomas and BD Loader (eds) *Cybercrime: Law Enforcement, Security and Surveillance in the Information Age* (2000) 243, 246, 249; Delgado and Stefancic op cit (n1) 324-325; JS Henry 'Beyond free speech: novel approaches to hate on the Internet in the United States' (2009) 18 *Info & Comms Technology L* 235 at 235; R Cohen-Almagor 'Countering hate on the Internet – A rejoinder' (2010) 2 *Amsterdam Law Forum* 125 at 127; McGonagle op cit (n1) 419-440; Timofeeva op cit (n1) 256. It is noteworthy that there was a surge in online hate speech in Kenya ahead of the contested March 2013 elections. This demonstrates the use of the Internet as a platform to fuel hostilities. See 'Internet: Online hate speech surged ahead of elections' 17 April 2013, available at <http://legalbrief.co.za/story/online-hate-speech-surged-ahead-of-elections-report/>, accessed on 17 April 2013.

⁴ Thomas and Loader op cit (n3) 13.

Facebook.⁵ Thus, there has been an increase in racist and hate attacks on the Internet where the anonymity, accessibility and immediate publication of digital technology make it easier for individuals to target their victims.⁶ Publication of material on the Internet also offers many advantages to the offender such as lower distribution costs, non-specialist equipment and a global audience.⁷ The Internet provides every user with the potential to become a publisher.⁸ The Internet provides a source of 'harmful' information that is difficult to regulate due to the anonymity of cyberspace.⁹ The global reach of social media also compounds the problem of seeking adequate redress for victims.¹⁰ The lack of direct governance by an international body to curb illegal content and activity on the Internet has also compounded the problem of regulation.¹¹ This has led to the perception of the Internet as a platform that requires a delicate balance, with companies and governments requiring sufficient freedom to leverage its power for their respective goals while consumers and online users expect the protection of their legitimate privacy rights.¹² The dissemination of hate speech on social media can also have serious repercussions for

⁵ S Naik 'Jail risk with racist comments on Twitter' *Saturday Star*, 31 March 2012 at 9; See 'Internet: Online hate speech' op cit (n3). Also see 'Preacher apologises for racial tweet' available at <http://www.elaw@legalbrief.co.za>, accessed on 11 January 2012, where a self-proclaimed preacher made a racial tweet about killing white people to eradicate racism and singled out DA leader Helen Zille in the tweet.

⁶ Gercke op cit (n3) 58, 68. Also see C Kim et al 'Computer crimes' (2012) 49 *Am Crim LR* 485; Henry op cit (n3) 251; Thomas and Loader op cit (n3) 12, 220; Perry and Olsson op cit (n1) 196; L Shaw 'Hate speech in cyberspace: Bitterness without boundaries' (2011) 25 *Notre Dame J Law, Ethics & Pub Pol'y* 279 at 280; S Kierkegaard 'Cybering, online grooming and ageplay' (2008) 24 *Computer Law & Sec'y Rep* 41 at 41; C Wolf 'Needed: Diagnostic tools to gauge the full effect of online anti-semitism and hate' (2004) 8 *J Internet Lawyer* 1; 11 at 12-14.

⁷ Gercke op cit (n3) 68.

⁸ Some online users may regard anonymity to be essential to their creativity. See E Burch 'Censoring hate speech in cyberspace: A new debate in a new America' (2001) 3 *N Carolina J Law & Technology* 175 at 175; R Marsden 'Daggers of online anonymity could kill off all restraint' *Saturday Star*, 29 March 2014 at 11.

⁹ It should be noted that the term 'cyberspace' refers to a unique medium or space that has no specific geographical location but it can be available to anyone, anywhere in the world who has access to the Internet (as defined in *Reno v ACLU* 521 US 844, 851 (1997)). Also see Kim et al op cit (n6) 485-486.

¹⁰ It is problematic if the victim and the offender live in different locations or jurisdictions but meet via the Internet. The anonymity and potential permanency of the online environment can contribute to an exacerbation of the emotional or psychological harm inflicted on victims of hate speech. See McGonagle op cit (n1) 433-434.

¹¹ Kierkegaard op cit (n6) 41.

¹² J Nguyen 'Internet privacy class actions: How to manage risks from increasing attacks against online and social media' (2011) 28 *Computer & Internet Lawyer* 8 at 11.

the utterer as employers have been known to consider social media screening.¹³

However, online hate speech may not necessarily be regarded as a criminal activity depending on the jurisdiction, with some countries limiting freedom of expression on the Internet by prohibiting hate speech. Since World War II, a number of European countries have enacted laws to restrict hate speech to promote respect and equality.¹⁴ What constitutes 'acceptable speech' in certain countries differs greatly as not all countries criminalise hate speech offences.¹⁵ The dissemination of Nazi propaganda and paraphernalia is considered illegal in Germany and France; however such material is easily available on the Internet.¹⁶ Thus, while the dissemination of xenophobic material is considered illegal in many European countries,¹⁷ it can be protected by the principle of freedom of speech enshrined in the First Amendment in the United States of America.¹⁸ To illustrate this, a French court ordered Yahoo! (a company based in the United States) to prevent access to Nazi-related material by French users; however as the sale of such material was considered legal in terms of the First Amendment of the United States Constitution, an American court decided that the French order was unenforceable against Yahoo! in the

¹³ Justine Socco, a public relations director at an Internet media company IAC made some hateful racist statements on Twitter before boarding a plane for South Africa during December 2013. She tweeted: "Going to Africa. Hope I don't get AIDS. Just kidding. I'm white!" This created a storm on the Internet and when her employers noticed the tweet, they distanced themselves from the offensive comments and subsequently fired her. See S Al-Dosari 'Watch out! Internet can make you or break you' 24 December 2013 at 23, available at www.arabnews.com/news/497381, accessed on 17 April 2015. This demonstrates that one should display caution when posting messages on social media.

¹⁴ R Cohen 'Regulating hate speech: Nothing customary about it' (2014) 15 *Chicago J International Law* 229 at 231; Thomas and Loader (n3) 246.

¹⁵ To illustrate this, the e-mail addresses of a group of Jewish students in Germany were targeted with threats to repeat the Holocaust. The messages were found to come from an adolfo@hitler.com website. However, the German police were found to be powerless to investigate such threats, as the e-mails emanated from a server in the United States where such material fell outside the jurisdiction of German laws that make neo-Nazi propaganda a crime in Germany. Gercke op cit (n3) 68. Also see Kim et al op cit (n6) 486; Thomas and Loader op cit (n3) 220.

¹⁶ See s 130 of the German Penal Code; French Penal code; Timodeeva op cit (n1) 260-261; Kim et al op cit (n6) 486 and BC Lewis 'Prevention of computer crime amidst international anarchy' (2004) 41 *Am Crim LR* 1353 at 1360.

¹⁷ See s 86a of the German Penal Code.

¹⁸ Gercke op cit (n3) 57; Kierkegaard op cit (n6) 41 and Lewis op cit (n16) 1360; Timofeeva op cit (n1) 257, 259; Cohen op cit (n14) 231.

United States of America.¹⁹ Similarly, the use of derogatory remarks regarding the Holy Prophet ('peace be upon him') may be regarded as criminal in many Arabic countries, but it is not regarded as such in some European countries.²⁰ This also creates the potential for conflict in some countries which may find the content of some online material to be objectionable.²¹ The challenges regarding jurisdiction also make the prosecution of offenders more difficult.²²

¹⁹ Gercke op cit (n3) 69; Van der Merwe et al op cit (n3) 445; Timofeeva op cit (n1) 270. Also see J Spector 'Spreading angst or promoting free expression? Regulating hate speech on the Internet' (2002) *Univ Miami Internat'l & Compar L R* 155 at 173-176.

²⁰ Gercke op cit (n3) 58. It is noteworthy that during 2008, the 'Joint Declaration of the UN Special Rapporteur on Freedom of Opinion and Expression' highlighted that international organisations, including the United Nations General Assembly and Human Rights Council, should desist from the further adoption of statements supporting the idea of defamation of religions. Gercke op cit (n3) 58. Also see the Charlie Hebdo terror incident in France which highlighted, *inter alia*, the issue of religious tolerance and the right to freedom of expression. For further discussion regarding the Charlie Hebdo incident, see *inter alia*, P McGeough 'Charlie Hebdo: total freedom of expression has little chance of survival in an imperfect world' *Sydney Morning Herald* 12 January 2015; P McGeough 'Balancing act of fighting terrorism while retaining right to freedom of expression' *The Age* 13 January 2015 at 6; A Harber 'Editors must edit without fear of being attacked' *Business Day BDLive* 15 January 2015, available at <http://www.bdlive.co.za/opinion/columnists/2015/01/15/editors-must-edit-without-fear-of-being-attacked>, accessed on 17 April 2015; D Chazan "'We're not Charlie': France divided over Charlie Hebdo Prophet Mohammed cartoons' *The Telegraph* 18 January 2015, available at <http://www.telegraph.co.uk/news/worldnews/europe/france/11353689/Were-not-Charlie-France-divided-over-Charlie-Hebdo-Prophet-Mohammed-cartoons.html>, accessed on 17 April 2015. These articles also explore the exercise of freedom of expression in a responsible manner. Please note that a detailed discussion about the Charlie Hebdo terror incident is beyond the scope of this article.

²¹ See 'Internet: Outrage over "blasphemous" Facebook post' 10 December 2014 available at <http://legalbrief.co.za/search/?q=Internet%3A+Outrage+over+%E2%80%98blasphemous%E2%80%99+Facebook+post>, accessed on 10 December 2014. Similarly, an anti-Islamic film made in the United States caused controversy and riots in many Muslim countries during 2012. Google settled with the Jordanian authorities to block all links connecting Internet users in Jordan, whilst it restricted the film in countries where it was considered illegal by local authorities such as India, Indonesia, Malaysia and Saudi Arabia. See 'Google extends blockade of anti-Islamic film' available at <http://legalbrief.co.za/story/google-extends-blockade-of-anti-islam-film>, accessed on 26 September 2012. It should be noted that countries such as Saudi Arabia, the UAE, Egypt, Indonesia and Malaysia have entered into agreements with the United States called the Mutual Legal Assistance Treaty ('MLAT'), whereby they can request American companies such as Google to follow local laws and remove objectionable content from its website. See 'Stalemate over YouTube clip continues' 30 January 2013, available at <http://legalbrief.co.za/search/?q=Governance%3A+Stalemate+over+YouTube+clip+continues>, accessed on 30 January 2013.

²² Gercke op cit (n3) 58.

The question also arises how regulation of online hate speech may impact on the right to freedom of expression. Countries have responded differently in regulating freedom of expression issues. European countries such as France and Germany tried to target Internet Service Providers ('ISPs'), but Germany has now decided not to hold ISPs liable for content that they transmit; China has introduced regulations that criminalise the distribution or consumption of harmful information on the Internet while the United States protects ISPs from liability provided that they are not the publishers of the information posted on their sites.²³ As stated earlier, the lack of agreement regarding the content of online material and to what degree specific acts should be criminalised exacerbates the problem of regulation of the Internet. States can maintain strict laws, employ education and awareness campaigns, block 'unacceptable' websites and filter the contents to regulate the dissemination of such material.²⁴ Approaches to filter systems may include the installation of programmes to analyse such websites and to install filter software on users' computers.²⁵ However, some countries may use filter technology to practise censorship.²⁶ The practice of websites relocating to a more favourable jurisdiction will mean that the blocking or banning of a website will not necessarily be an effective remedy.²⁷ It is submitted that legal approaches to criminalise illegal and hateful content on the Internet should not interfere with the right to freedom of expression. However, the right to freedom of expression may be subject to restrictions.²⁸

The dearth of adequate regulation of the Internet to address online hate speech has raised concerns for many policymakers. This article will examine the controversy between the restrictions on online hate speech and the interests of freedom of expression in the United States

²³ Kim et al op cit (n6) 486; see Spector op cit (n19) 174; Cohen op cit (n14) 239-240, 243-244. Also see s 230(c)(1) of the Communications Decency Act of 1996 which protects ISPs from liability in the United States.

²⁴ Gercke op cit (n3) 59; McGonagle op cit (n1) 434.

²⁵ Gercke op cit (n3) 59; Also see Timofeeva op cit (n1) 278-280, regarding the use of filtering and censorship measures.

²⁶ Gercke op cit (n3) 59. It should be noted that the power of censorship is based on the concept of state sovereignty with states being free to control and regulate the media according to their internal policy. See A Shytov 'Indecency on the Internet and international law' (2005) *Internat'l J Law and Info Tech* 13 260 at 267; Chen op cit (n1) 229-267.

²⁷ See McGonagle op cit (n1) 432, 434.

²⁸ See for example, the limitation clause in s 36 of the South African Constitution and article 10(2) of the European Convention for the Protection of Human Rights and Fundamental Freedoms.

of America, the United Kingdom and South Africa.²⁹ The article will also examine international instruments addressing online hate speech and recent legal developments concerning such speech in the selected jurisdictions. The study reveals that the interconnection between technologies and social media and the rise in socio-pathologies like racism and other forms of discrimination on the Internet, have the potential for cyber racists to abuse the cyber space afforded to them and inflict harm on their victims. It is advocated that countries should consider the diverse interests of online users and victims when addressing the rise in online hate speech. It is also necessary to ensure that the right to freedom of expression is exercised in a responsible manner and to recognise the societal impact of new technologies on free speech.³⁰ It is submitted that countries should employ education and public awareness campaigns to promote intercultural dialogue and understanding among its citizens, employ the media to eliminate negative stereotyping of minorities and vulnerable groups and outlaw hate speech through whichever medium it is disseminated. ISPs should also be held more accountable for the content that they transmit and they should use filtering technology in appropriate circumstances.

2 Defining online hate speech

Hate speech refers to disparaging and abusive comments, words and phrases directed at individuals or groups representing a specific race, religion, ethnic background, gender or sexual preference.³¹

²⁹ It is noteworthy that hate speech does not enjoy any constitutional protection in South Africa, unlike the United States where the First Amendment affords public discourse extraordinary protection and limits hate speech in exceptional circumstances (where the speech amounts to 'true threats'). The United Kingdom does not have a Bill of Rights. However, it has a Human Rights Act, 1998 which codifies protections in the European Convention for the Protection of Human Rights and Fundamental Freedoms, 1950. Both these countries have taken concerted steps to address the tension between online hate speech and the right to freedom of expression; hence these countries were chosen for the comparative study. The discussion in sections 6.1 and 6.2 below will elaborate on the steps taken in these countries. It should also be emphasised that the terms 'freedom of expression', 'free speech' and 'freedom of speech' may be used interchangeably in this article.

³⁰ McGonagle op cit (n1) 440.

³¹ Van der Merwe op cit (n3) 444. Moeckli et al op cit (n1) 230; Timofeeva op cit (n1) 257, McGonagle op cit (n1) 420. It should be noted that vilification is a category of hate speech whose intention is to wound, insult or intimidate a particular group. Vilification is motivated by hostility or contempt for the group's racial or religious identity. See C Young 'Does freedom of speech include hate speech?' (2011) 17 *Res Publica* 385 at 394. However, Cohen maintains that there is no universally accepted definition of 'hate speech'. See Cohen op cit (n14) 251.

Such speech has a propensity to injure its victims.³² Hate speech may be distinguished from hate crime laws that target crime motivated by prejudice.³³ According to the Special Rapporteur to the United Nations, due to population movements, increased immigration levels and concern about terrorism, there is a growing tendency to stigmatise specific groups and communities.³⁴ The Internet provides an avenue for individuals to resort to racist and xenophobic propaganda due to its flexibility and anonymity, with the result that it is often complex and difficult to monitor extremist sites online.³⁵ It has added a global dimension to the availability of propaganda advocating violence and hatred of others.³⁶ However, one should bear in mind that hate speech on the Internet has its roots in the biases and prejudices that existed before the advent of the Internet.³⁷ The Internet is said to be rife with hate speech, chat groups, blogs and websites touting white supremacy and asserting the inferiority of gays, blacks, Jews, Latinos, Muslims, foreigners and women.³⁸ Online hate speech may take the form of posting abusive, hateful or racist messages or comments on social networking sites such as Twitter or Facebook.³⁹ A number of

³² Hate speech may also cause psychological harm and damage to self-esteem. See Shaw op cit (n6) 282 and F Brennan 'Legislating against Internet race hate' (2009) 18 *Info & Comms Technology L* 123-153 at 127 and McGonagle op cit (n1) 421, 433-434 regarding the impact of hate speech on victims. Regarding remedies for victims of social media, see J Holschuh 'Utilising torts to combat hate speech in online social media' (2013-2014) 82 *Univ Cincinnati L R* 953-977.

³³ G Mason 'Prejudice and paedophilia in hate crime laws: *Dunn v R*' (2009) 34 *Alternative L J* 253 at 253.

³⁴ K Boreham 'The right to truth: The freedom to speak, to know and to be protected from lies' (2014) 26 *LegalDate* 2 at 3.

³⁵ Wolf op cit (n6) 13; Thomas and Loader op cit (n3) 249.

³⁶ Moeckli et al op cit (n1) 230; F Aumueller 'Hate propaganda law and Internet-based hate' (2000-2001) 44 *Crim L Q* 92 at 95. However, it is difficult to ascertain where online hate speech is uttered and who is responsible for its utterance due to the communication medium of the Internet. See C Reed 'The challenge of hate speech online' (2009) 18(2) *Info & Comms Technology L* 79 at 79 and Burch op cit (n8) 185.

³⁷ Timofeeva op cit (n1) 256.

³⁸ See Delgado and Stefancic op cit (n1) 324-325.

³⁹ The posting of racial and offensive comments on social media is called cyber-smearing. See Naik op cit (n5) 9 and Delgado and Stefancic op cit (n1) 328-332. Other examples of Internet hate speech may include computer users being verbally attacked online, terrorists exchanging plans for mass destruction, racist comments by white supremacists intimidating non-whites and Jews, neo-Nazism and Holocaust denial. See further, Wolf op cit (n6) 12; A Guichard 'Hate crime in cyberspace: The challenges of substantive criminal law' (2009) 18 *Info & Comms Technology L* 201 at 210. It is noteworthy that YouTube guidelines define hate speech as content that promotes hatred against members of a protected group. See 'Clip violates all constitutions and laws in the world' 19 September 2012, available at <http://legalbrief.co.za/search/?q=Clip+violates+all+constitutions+and+laws+in+the+world>, accessed on 19 September 2012.

websites also contain material that may be regarded as anti-religious or illegal in some jurisdictions such as the defamation of religions or the publication of cartoons.⁴⁰ Although most Western countries have rejected the need to protect religious institutions or doctrine from robust criticism, satire or ridicule, the Human Rights Council has endorsed the idea of protecting groups from intolerance and violence.⁴¹ The European Court of Human Rights has also accepted the legitimacy of a prosecution for blasphemy where 'believers' may legitimately have felt themselves to be the object of unwarranted and offensive attacks and the punishment imposed was an 'insignificant fine'.⁴²

3 The impact of online hate speech on human rights

The dissemination of online hate speech may infringe some basic human rights, such as the right to privacy, right to human dignity and the right to freedom of religion, belief and opinion.⁴³ Section 16 of the South African Constitution ('the Constitution') provides that everyone has the right to freedom of expression, which includes *inter alia* freedom to receive or impart information and freedom of artistic creativity.⁴⁴ However, freedom of expression is not an absolute right and its protection does not extend to incitement of imminent violence or advocacy of hatred based on race, ethnicity, gender, religion and any action that constitutes incitement to cause harm.⁴⁵ It

⁴⁰ See Gercke op cit (n3) 70. It is noteworthy that the posting of the controversial film 'Innocence of Muslims' on YouTube was regarded as hate speech by the *Huffington Post* analyst Philip Seib. See 'Clip violates all constitutions...' op cit (n39).

⁴¹ See Human Rights Resolution 16/18 (12 April 2011) (Combating Intolerance, Negative Stereotyping and Stigmatization of, and Discrimination, Incitement to Violence Against, Persons Based on Religion or Belief). Also see Moeckli et al op cit (n1) 231. For further discussion about Islamophobia (prejudice and hostility towards Muslims in the West) and anti-Semitism and Holocaust-denial (minimises the Nazi genocide during World War II), see Moeckli et al op cit (n1) 230 and Brennan op cit (n32) 128,132.

⁴² See *IA v Turkey* (2007) 45 ECHR 703, 30, paras 29 and 32.

⁴³ See McGonagle op cit (n1) 421, 433-434.

⁴⁴ See s 16(1)(a)-(d) of the Constitution. Free speech is thus guaranteed in s 16.

⁴⁵ See s 16(2)(a)-(c) of the Constitution. Messages which threaten other users with violence fall into the category of "incitement to violence" and they constitute unprotected expression in terms of the Constitution. The effect is that an individual's right to exercise his artistic creativity on the Internet or a social networking site will have to be weighed against the victim's rights to privacy and dignity. See further, ss 14 and 10 of the Constitution which protect the right to privacy and human dignity respectively. It is submitted that hate speech on the Internet can be hurtful and insulting, and such speech may infringe the privacy and dignity of a person. As hate speech is expressly excluded from the ambit and scope of the right to freedom of expression in s 16 of the Constitution, the rights to dignity, privacy, bodily and psychological integrity will always outweigh the use of hate speech.

is submitted that a person's dignity is impaired when one is subjected to treatment which constitutes degrading or humiliating conduct such as the posting of abusive or racist texts, messages or images on social networking sites such as Twitter and Facebook. Similarly, the ability of individuals to invade the privacy and sanctuary of one's home by electronic media espousing hate speech infringes one's right to communication, access to information, and one's right to privacy. However, all rights in the Constitution are not absolute as they may be limited by s 36.⁴⁶ According to Thomas and Loader, hate speech on the Internet may lead to an examination of issues of balance in terms of freedom of expression and responsibility.⁴⁷ It is submitted that countries should use educational and public awareness campaigns to raise awareness about the scope and impact of online hate speech and to foster tolerance and respect for diversity.

Some South African courts have generally required a clear incitement of harm or violence for speech to be labelled hate speech.⁴⁸ However, in *AfriForum v Malema*,⁴⁹ public utterances by Julius Malema (the then President of the African National Congress Youth League) in the infamous 'shoot the Boer/farmer' song were found to constitute hate speech by the Equality Court. Although the Equality Court acknowledged that the song is part of South Africa's history, it found

⁴⁶ See s 36(1) of the Constitution, where the limitation has to be reasonable and justifiable in an open and democratic society.

⁴⁷ Thomas and Loader op cit (n3) 230. It is noteworthy that this position does not exist in South Africa where hate speech is expressly excluded from the scope and ambit of the right to freedom of expression in the Constitution.

⁴⁸ See *inter alia Human Rights Commission of South Africa v SABC* 2003 (1) BCLR 92 (BCCSA) (where the Broadcasting Complaints Commission found that derogatory and inflammatory statements about the Indian population in a Zulu song amounted to advocacy of hatred based on race) and *Ramesh Dharashee Jethlalal v Mbongeni Ngema and Universal Music* (unreported, 3524/2002, 28 June 2002) (court declined to extend interim interdict prohibiting a Zulu song found to denigrate the Indian population because no violent actions had been perpetrated against Indians). See further, 'Hate speech and freedom of expression in South Africa' (2013) *Freedom of Expression Institute* 21-25, available at https://www.fxj.org.za/fxi_downloads/hate_speech_and_freedom_of_expression_in_SA.pdf, accessed on 13 April 2015. However, see the discussion in 6.3 below where the Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000 (PEPUDA) is discussed. Section 10 of PEPUDA regulates hate speech in South Africa. It is submitted that the PEPUDA test requires the interpretation of words to demonstrate a clear intention to be hurtful rather than an intention to incite harm or violence. See section 6.3 below further.

⁴⁹ 2011 (6) SA 240 (EqC). The facts were that Afri-Forum brought a case against Malema on the grounds that the 'objectionable utterances' by Malema ('shoot the Boer/farmer song') undermined the human dignity and adversely affected the rights and freedoms of the white Afrikaans speaking community (including Afrikaner farmers). Malema stated that he had a right to use these words which were taken from a liberation song.

that the song had the ability to incite and promote violence against the white Afrikaans speaking community, and it was unacceptable to sing it in the present peaceful climate prevailing in the country.⁵⁰ The Equality Court concluded that the singing of the song constituted hate speech because the words referred to a recognisable group in society (Afrikaner population) and the words undermined their dignity. The words were found to be discriminatory and harmful to that group. As such, no justification existed for the words to be sung. The Equality Court thus prohibited Malema from singing the song at public or private functions.

Hate speech is said to be on the rise on social media in South Africa, with the South African Human Rights Commission reporting an increase in the number of complaints relating to racial posts on social networking sites.⁵¹ The racial insults are said to emanate from all racial groups and the insults include use of the K-word and violations and threats against minority groups.⁵² This demonstrates the irresponsible use of social media which is a global trend. This has led to calls by some South African politicians for publishers to regulate racist online comments more actively.⁵³

American courts have also faced the challenge of balancing the First Amendment right to free speech against restrictions imposed on free speech through laws governing the Internet. The First Amendment provides that 'Congress shall make no law abridging the freedom of speech or of the press, or the right of the people peaceably to assemble, and to petition the Government for a redress of grievances'.⁵⁴ Hate speech and other forms of racist speech receive the

⁵⁰ *AfriForum* supra (n49) at paras [108], [109].

⁵¹ See 'Internet' op cit (n3).

⁵² *Ibid.* It is noteworthy that the 'K-word' refers to the word 'kaffir'. This word is used to describe the non-white (black) population in South Africa. It is submitted that the use of this word is insulting, derogatory and offensive. See <http://www.oxforddictionaries.com/definition/english/kaffir>, accessed on 9 September 2015.

⁵³ Comments were made by the South African Communist Party general secretary, Blade Nzimande. Also see B Nzimande 'It's time to confront cyber racists and hate speech' (2015) 14 *Umsebenzi online*, available at <http://www.sacp.org.za/main.php/ID=4599>, accessed on 13 April 2015.

⁵⁴ See H Rubin et al 'US international law aspects of the Internet: Fitting square pegs into round holes' (1995) (3) 2 *Int J L & Info Technology* 117 at 118. It should be noted that article 10(1) of the European Convention on Human Rights and article 19 of the International Covenant on Civil and Political Rights guarantee the right to freedom of expression. For further discussion about these international instruments, see section 5 below. It is noteworthy that the First Amendment ranks the freedom of expression as pre-eminent over all other rights. This is unlike the South African Constitution which does not grant the right to freedom of expression superior status over other human rights. See E de Waal, R Mestry and CJ Russo 'Religious and cultural dress at school: A comparative perspective' (2011) 14 *PER* 62 at 80, available at <http://dx.doi.org/10.4314/pelj.v14i6.3>, accessed on 5 March 2013.

same protection on the Internet as they have always received under the First Amendment.⁵⁵ It was held in *Simon & Schuster, Inc v Members of New York State Crime Victims Board*,⁵⁶ that 'the fact that society may find speech offensive is not a sufficient reason for suppressing it'. This case demonstrates that offensive speech may be acceptable in certain circumstances. It is submitted that offensive speech may be distinguished from hate speech in that offensive speech may be found to be annoying, irritating or it may cause displeasure or anger, whilst hate speech incorporates disparaging and abusive comments, words and phrases directed at individuals or groups representing a specific race, religion, ethnic background, gender or sexual preference, with the propensity to injure its victims.⁵⁷ Racist speech may also be protected on the Internet provided that it does not cause a reasonable person to fear for his or her safety.⁵⁸

4 Respecting freedom of expression on the Internet

The Internet offers many advantages for those users who wish to engage in lively debate or discourse as people may post material and write articles without having to disclose their identity.⁵⁹ The Internet is said to endorse the principle of freedom of speech; however its application is subject to conditions and laws in countries around the world.⁶⁰ In the United States, the First Amendment protects speech including speech that is hateful and offensive.⁶¹ As a result of growing international concern, governments have been prompted to police the Internet and to take legal action against those who post offensive and illegal hate material on the Internet.⁶² Although there have been

⁵⁵ See Kim et al op cit (n6) 450; 485. Also see Henry op cit (n3) 235; Burch op cit (n8) 179. This makes the United States a favourable jurisdiction for propagators of online hate speech. See McGonagle op cit (n1) 437.

⁵⁶ 502 US 105 (1991) at 118.

⁵⁷ See *inter alia* Van der Merwe et al op cit (n3) 444, Moeckli et al op cit (n1) 230; Timofeeva op cit (n1) 257, McGonagle op cit (n1) 420, 421, 433-434; Cohen op cit (n14) 251 and Shaw op cit (n6) 282. Thus, it is submitted that hate speech has a wider impact on its victims.

⁵⁸ Thus, the exception to the general free speech principle exists for 'true threats' which include sending a threatening e-mail message or expressing an intention to commit an act that is racially motivated by way of a public announcement on the Internet. See Henry op cit (n3) 235.

⁵⁹ Gercke op cit (n3) 70-71.

⁶⁰ Gercke op cit (n3) 71.

⁶¹ Henry op cit (n3) 235. However, there is criticism that much offensive and illegal material originates in America where the First Amendment protects all free expression of opinion except that which directly incites violence. See Thomas and Loader op cit (n3) 248.

⁶² Thomas and Loader op cit (n3) 242.

some successful prosecutions of individuals who disseminate hate speech over the Internet, web-based hate continues to receive broad First Amendment protection.⁶³ The online writing revolution may also result in liability issues. In the United Kingdom, a British student, Liam Stacey, was sentenced to 56 days in prison for posting racist comments on Twitter concerning the cardiac arrest of footballer Fabrice Muamba during 2012.⁶⁴ According to South African law firm Webber Wenzel, South Africans who post racist and offensive comments on social networking sites may be charged with *crimen injuria* or hate speech.⁶⁵

Internet freedom is regarded as a controversial issue around the world. Google regards the United States government's attempt to tackle terrorist threats in cyberspace as constituting snooping on the Internet and a threat to democracy; the Turkish's government's recent attempt to block access to Twitter was regarded as an attack on personal freedoms and an attack on its international rights obligations, whilst China has been criticised for limiting Internet freedom through its adoption of stricter rules on Internet Service Providers ('ISPs') and users.⁶⁶ It has been suggested that the foreign policy of governments should consider the realities of diplomacy and endeavour to maintain a steady stream of information that is suitably presented for the global audience.⁶⁷ However, a Canadian human rights group has reported that a number of nations are using American-made Internet surveillance technology to censor online content and track their citizens.⁶⁸ This has stimulated debate over whether intelligence agencies should have access to online communications and it has raised concerns over citizens' digital privacy rights.⁶⁹ During July 2012, the United Nations Human Rights Council passed its first resolution to protect free speech of individuals online.⁷⁰ This was hailed as a step in the right direction by privacy groups.

⁶³ Henry op cit (n3) 236.

⁶⁴ See 'Investigation into racial tweets' available at <http://legalbrief.co.za/search/?q=Cybercrime%3A+Investigation+into+racial+tweets>, accessed on 27 June 2012.

⁶⁵ Naik op cit (n5) 9. See further, the discussion in 6.3 below regarding the position in South Africa.

⁶⁶ See 'Turkey's Twitter ban backfires' 26 March 2014, available at <http://legalbrief.co.za/story/turkeys-twitter-ban-backfires> accessed on 26 March 2014. Also see 'Beijing in fresh Internet crackdown' 14 January 2015, available at <http://legalbrief.co.za/search/?q=%27+Beijing+in+fresh+Internet+crackdown>, accessed on 14 January 2015.

⁶⁷ See 'Clip violates all constitutions...' op cit (n39).

⁶⁸ 'Turkey's Twitter ban...' op cit (n66).

⁶⁹ Ibid.

⁷⁰ It was approved by all 47 members including China and Cuba op cit (n66).

5 The role of international instruments in regulating online hate speech

The International Covenant on Civil and Political Rights ('ICCPR') is an international human rights treaty that protects the right to freedom of expression.⁷¹ Article 20(1) of the ICCPR prohibits the advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence. The United Nations Universal Declaration of Human Rights ('UNDHR') is widely regarded as a milestone document in the history of human rights.⁷² Article 2 of the UNDHR provides that everyone is entitled to all rights and freedoms contained in the Declaration, and it prohibits *inter alia* any 'distinction' on the basis of race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. The right to freedom of opinion and expression is protected in article 19 of the UNDHR, and this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media. Article 4 of the United Nations ('UN') Convention on the Elimination of All Forms of Racial Discrimination requires all states to criminalise speech amounting to war, propaganda or advocating racial hatred.⁷³ States ratifying the Convention are obliged to adopt legal measures to eliminate racial discrimination in all its forms and manifestations and to prevent and combat racist doctrines and practices.⁷⁴

The European Convention for the Protection of Human Rights and Fundamental Freedoms ('ECHR') regulates *inter alia* the European response to hate speech. The right to freedom of expression is entrenched in article 10(1), whilst article 10(2) allows certain legal limitations that are necessary in a democratic society such as those in the interests of national security, for the prevention of disorder

⁷¹ See article 19. The ICCPR came into effect on 23 March 1966 and thus far 167 countries, including the United States, United Kingdom and South Africa have ratified the treaty. See N Ghana 'Expression and hate speech in the ICCPR: Compatible or clashing?' (2010) 5 *Religion & HR* 171-190.

⁷² It was drafted by representatives with different legal and cultural backgrounds from different countries. It was proclaimed by the United Nations General Assembly in Paris on 10 December 1948 by General Assembly resolution 217A. See further, <http://www.un.org/en/universal-declaration-human-rights/>, accessed on 24 November 2015.

⁷³ However, these limitations must comply with article 19(3), be provided by law and be the least restrictive means of achieving the relevant aim. The Convention was adopted in 1965. See Cohen op cit (n14) 248-249 and Thomas and Loader op cit (n3) 242-243 for a discussion about the Convention. The Convention is said to be the oldest and most ratified UN human rights convention.

⁷⁴ Thomas and Loader op cit (n3) 242. South Africa is a signatory to the UN Convention on the Elimination of All Forms of Racial Discrimination, which prohibits hate speech.

or crime and for the protection of the rights of others.⁷⁵ In *Jersild v Denmark*⁷⁶ the European Court of Human Rights ('EC') held that an intention to spread racist ideas and views is required to attract liability for hate speech. The EC has also held that the right to freedom of expression extends to information or ideas that offend, shock or disturb because the values of pluralism, tolerance and broadmindedness are essential to a democratic society, and every limitation placed on the freedom of expression must be proportionate to the legitimate aims of governments.⁷⁷

The Convention on Cybercrime criminalises certain computer actions and it is regarded as the first international treaty on crimes on the Internet.⁷⁸ The Convention strives to, *inter alia*, achieve unity between member states, foster co-operation between states by adopting a common criminal policy aimed at the protection of cybercrime, and by the adoption of sufficient measures to combat such criminal offences.⁷⁹ Due to the inability of countries to agree on a common position regarding the criminalisation of the dissemination of xenophobic material, the entire topic was initially excluded from the Convention on Cybercrime and instead addressed in a separate First Protocol.⁸⁰ However, the Additional Protocol to the Convention on Cybercrime which came into effect during March 2006, addresses the criminalisation of acts of a racist and xenophobic nature committed

⁷⁵ See McGonagle op cit (n1) 424-425.

⁷⁶ 23 September 1994, Series A298, reference 36/1993/432/510. This case involved the conviction of a Danish journalist for aiding and abetting in the dissemination of racist statements in a televised interview. The journalist alleged an infringement of article 10 (freedom of expression) and the EC agreed. The EC held that the Danish court had failed to establish that the interference with the reporter's freedom of expression was necessary in a democratic society. See Cohen op cit (n14) 250 and McGonagle op cit (n1) 475 for a discussion about the case. The EC is said to maintain a high level of protection for freedom of expression. See M Oetheimer 'Protecting freedom of expression: The challenge of hate speech in the European Court of Human Rights case law' (2009) 17 *Cardozo J Internat'l & Comp L* 427-443.

⁷⁷ *Handyside v United Kingdom* Series A24;1 EHRR 737 (1979-1980). Also see McGonagle op cit (n1) 425 and Brennan op cit (n32) for a discussion about the case.

⁷⁸ It was signed in Hungary on 23 November 2001. Its aim is to combat cybercrime. See F Cassim 'Addressing the growing spectre of cyber crime in Africa: Evaluating measures adopted by South Africa and other regional role players' (2011) 44 *CILSA* 123 at 126; Kim et al op cit (n6) above at 487; Timofeeva op cit (n1) 265-266.

⁷⁹ See Preamble. Also see Gercke op cit (n3) 69.

⁸⁰ Gercke op cit (n3) 70.

through computer systems.⁸¹ The Protocol requires each signatory state to adopt and implement legislative measures to criminalise intentional conduct committed by means of computers, such as *inter alia*, the distribution of racist and xenophobic material and the utterance of threats and insults motivated by racism and xenophobic material.⁸² The Protocol recognises that freedom of expression is one of the essential foundations of a democratic society; it is concerned about the risk of misuse or abuse of computer systems to disseminate racist and xenophobic propaganda and it is mindful of the need to ensure a proper balance between freedom of expression and an effective fight against acts of a racist and xenophobic nature.⁸³ However, the Additional Protocol has been subjected to criticism.⁸⁴ A commendable initiative has been introduced by the Council of Europe to promote youth awareness of online hate speech called the 'No Hate Speech Movement'.⁸⁵ The 'No Hate Speech Movement' is based on human rights education, youth participation and media literacy and it offers project funding through its programmes.⁸⁶

The African Charter on Human and People's Rights ('ACHPR') is a regional human rights instrument which strives to reflect the history, values, traditions and development of the African people.⁸⁷ Article 2 of the ACHPR contains the non-discrimination provision, and it guarantees rights and freedoms in the Charter without distinction of

⁸¹ See article 2(1). The Protocol defines 'racist and xenophobic material' as any written material, any image or any other representation of ideas or theories which advocates, promotes or incites hatred, discrimination or violence, against any individual or group of individuals, based on race, colour, descent or national or ethnic origins and religion. South Africa signed this Protocol on 4 April 2008. Also see 'Additional Protocol to the Convention on Cybercrime (Concerning the Criminalisation of Acts of a Racist and Xenophobic Nature Committed Through Computer Systems (adopted 7/11/2002),' available at <http://conventions.coe.int/Treaty/en/Treaties/Htm1/189.htm>, accessed on 24 February 2015; McGonagle op cit (n1) 427-428 and Van der Merwe et al op cit (n3) 449.

⁸² See articles 3(1) and 4-7 in Chapter 1.

⁸³ Ibid.

⁸⁴ It has been criticised by Brennan as not standing up to the test of effectiveness as it prioritises freedom of speech over freedom from racial discrimination. See Brennan op cit (n32) 123-153 at 141, 143.

⁸⁵ See 'Online hate speech surged ahead of elections – report,' available at <http://legalbrief.co.za/story/online-hate-speech-surged-ahead-of-elections-report/>, accessed on 17 April 2013. Also see 'Preventing online hate speech through social media and democratic participation' available at <https://www.theewc.org/statement/preventing-online-hate-speech-through-social-media-and-democratic-participation>, accessed on 13 April 2015.

⁸⁶ Ibid.

⁸⁷ It is also called the Banjul Charter and was adopted by the Assembly of Heads of States and governments of the Organisation of African Unity in 1981. It came into force in 1986. See 'African Charter on Human and People's Rights,' available at <http://www.achpr.org/instruments/achpr/>, accessed on 24 November 2015.

any kind on grounds such as race, ethnic group, colour, sex, language, religion, political or any other opinion, national and social origin, fortune, birth or other status. It is submitted that the contents of this article are similar to article 2 of the UNDHR. Article 9 of the ACHPR guarantees *inter alia* the right of every individual to express and disseminate opinions within the law.

The above discussion demonstrates that international instruments are playing an important role in promoting diversity and in addressing hate speech in all media platforms in society.

6 An overview of legislation and policies – a comparative perspective

The use of speech to incite violence is regarded as a criminal offence in most countries, but it is difficult to strike an acceptable balance between the right to freedom of expression and restrictions on forms of objectionable speech.⁸⁸ The following discussion will address legislative frameworks and policies introduced in the United States of America, the United Kingdom and South Africa to regulate online hate speech.

6.1 United States of America

In the United States ('US'), there are legal frameworks protecting freedom of expression.⁸⁹ The issue of online hate speech enjoys the same protection as any other forms of speech under the First Amendment.⁹⁰ The First Amendment does not protect speech that constitutes a threat of imminent harm to an identifiable victim.⁹¹ The cornerstone of free speech in the United States is referred to as the 'free marketplace of ideas.'⁹² This means that citizens can sort out beliefs and ideals which best resonate with them and discard those that do not.⁹³ It should be noted that the government may not suppress lawful speech as a means to suppress unlawful speech.⁹⁴ This has

⁸⁸ Moeckli et al op cit (n1) 230.

⁸⁹ However, hate speech may provide evidence in a hate crime. See Van der Merwe et al op cit (n3) 445.

⁹⁰ US Constitution, First Amendment. US courts are wary of infringing the right to free speech in the First Amendment. See Henry op cit (n3) 235. However, it has been suggested that hate speech is inconsistent with the principle of equality in the Fourteenth Amendment. See Burch op cit (n8) 179.

⁹¹ See *Watts v United States* 394 US 705 (1969).

⁹² Spector op cit (n19) 159; Burch op cit (n8) 181.

⁹³ Henry op cit (n3) 236.

⁹⁴ Kierkegaard (n6) 45. Also see *Ashcroft v Free Speech Coalition* 535 US 234 (2002) at 254-55.

resulted in hate organisations finding a safe haven in the United States from which to launch their hateful messages.⁹⁵ The government may thus only prosecute individuals who express a 'true threat' on the Internet. Foreign pressure has been placed on the United States to respect the laws of other countries that restrict hate speech by assisting in preventing the distribution of hate speech on the Internet.⁹⁶ The export of Neo-Nazi propaganda from the United States to countries such as France and Germany which prohibit such speech has come into conflict with First Amendment liberties.⁹⁷ However, there is no *per se* absolute protection of speech in the United States.⁹⁸

Freedom of speech in the United States is said to depend on the political atmosphere at the time. The Aliens and Seditious Acts which were passed after the adoption of the Bill of Rights curtailed freedom of speech; convictions were upheld for freedom of speech during World War I, and during World War II the fear of communism reduced the protection of the First Amendment to a test based on an assessment of the gravity of evil, discounted by its probability.⁹⁹ Protection afforded to racist speech has also changed. In *Beaubarnais v Illinois*,¹⁰⁰ the US Supreme Court upheld a conviction for 'group libel' against the writer or publisher of a white supremacist periodical for distributing a petition that contained many inflammatory statements about African-Americans. According to the court, the statements were regarded to be 'fighting words', a category of speech which is unprotected by the First Amendment.¹⁰¹ However, in *Brandenburg v Ohio*,¹⁰² the speech of a Ku Klux Klan leader calling for vengeance at a rally in the event

⁹⁵ Henry op cit (n3) 236.

⁹⁶ Spector op cit (n19) 155.

⁹⁷ Ibid. Also see Timofeeva op cit (n1) 275. It should be noted that the First Amendment plays a critical role in a US Court's decision whether or not to enforce the judgment of a foreign court. See *Yahoo! Inc v LICRA (TGI de Paris)* 169 F. Supp. 2d 1181, 22 May 2000, interim court order No 00/05308, 00/05309). A French court had ordered Yahoo! to take necessary steps to remove access to Nazi material on its auction site. Inc was a US corporation with no assets in France. Inc brought an action in the Californian courts that the French judgment could not be enforced in the US. The US court held that it could not enforce a foreign order that violates the protections of the US Constitution by chilling protected speech that occurs in the US.

⁹⁸ Spector op cit (n19) 162.

⁹⁹ Spector op cit (n19) 160.

¹⁰⁰ 343 US 250 (1952). For a discussion about the case see Timofeeva op cit (n1) 271-273 and Chen op cit (n1) 252-259.

¹⁰¹ Spector op cit (n19) 160. It should be noted that the *Beaubarnais* decision is now widely regarded as being no longer binding, although it has never been overturned explicitly and the US Supreme Court still mentions it in its decisions. Also see Timofeeva op cit (n1) 271 and Spector op cit (n19) 161.

¹⁰² 395 US 444 (1969). The Klu Klux Klan leader had made disparaging comments about blacks and Jews. Also see Timofeeva op cit (n1) 271-273 and Cohen op cit (n14) 245-246 for discussion about the case.

of the President suppressing the white, Caucasian race was deemed to 'be protected speech' as it did not amount to an incitement or production of an imminent threat of violence.¹⁰³ US courts have also held that regulations may not favour one form of protected speech over another.¹⁰⁴

The issue of free speech and the Internet has been addressed in *Reno v ACLU*¹⁰⁵ where the US Supreme Court struck down Title V of the Telecommunications Act of 1996, also known as Communications Decency Act of 1996 ('CDA'). The above Act tried to reduce the exposure of minor children to pornographic materials on the Internet.¹⁰⁶ The court found no basis for subjecting the Internet medium to First Amendment scrutiny, and concluded that the interest in encouraging freedom of expression in a democratic society outweighs any unproven benefit of censorship.¹⁰⁷ This demonstrates the court's belief in non-interference and respect for democracy and respect for the exchange of ideas in the marketplace. The *Reno* case also demonstrates that the government cannot restrict online speech unless that speech falls within an unprotected category such as obscene speech or speech that represents a 'true threat'.¹⁰⁸

In *Planned Parenthood of the Columbia/Williamette Inc v American Coalition of Life Activists*,¹⁰⁹ the American Coalition of Life Activists ('ACLA') had compiled a series of dossiers on doctors who performed

¹⁰³ This created a new standard for 'fighting words.' See Spector op cit (n19) 161.

¹⁰⁴ See *RAV v City of St Paul* 505 US 377 (1992). In this case, an African-American family became the target of racially offensive incidents. A white teenager was convicted of misdemeanour by the Minnesota Supreme Court when he placed a cross on the family's lawn. See Spector op cit (n19) 163; Cohen op cit (n14) 246; Tsesis op cit (n1) 849-853.

¹⁰⁵ Supra (n9). In this case, the United States Supreme Court ruled that the First Amendment applies in full measure to speech on the Internet. Section 230 of the CDA, which was struck down as unconstitutional by the court, protects ISPs from liability for the contents of its user websites.

¹⁰⁶ The CDA prohibited *inter alia* the 'knowing' transmission of obscene or indecent messages to any recipient under 18. The American Civil Liberties Union challenged this, finding the anti-indecency provisions of the CDA to be overbroad. The US Supreme Court agreed.

¹⁰⁷ *Reno* supra (n9) at 870, 885. The US Supreme Court elected not to institute a new test for unprotected speech on the Internet in this decision. The decision demonstrates that American legal jurisprudence tends to favour public discourse rather than speech regulation. According to Spector, this case can be distinguished from the *Yahoo!* case in that the concern in the *Reno v ACLU* case was to prevent minors from accessing the pornographic material when others had a right to speak on the issue; whereas in the *Yahoo!* cases, authorities in France (and Germany) were seeking to enforce laws that strove for blanket enforcement. Spector op cit (n19) 171.

¹⁰⁸ Moreover, the government can only regulate speech based on content where it can be shown that the regulation is necessary to serve a compelling state interest. Henry op cit (n3) 236.

¹⁰⁹ 244 F. 3d 1007 (9th Cir. 2001).

abortions and an activist posted personal information about the doctors and their families on an anti-abortion web site. A group of doctors sued the ACLA and twelve other activists alleging, *inter alia*, that the defendant's speech had harmed them. The Ninth Circuit Court of Appeals acknowledged that the language of the website may have frightened the doctors. However, the court regarded the test to be whether the speech in question constituted a 'true threat' in order for the speech to constitute unprotected speech. The defendant's speech on the Internet was held to be protected as the speech did not constitute a 'true threat'. The Ninth Circuit Court had considered the case of *NAACP v Claiborne Hardware Co*¹¹⁰ in reaching its decision. In *NAACP v Claiborne Hardware Co*, the NAACP had organised a boycott of white businesses and one boycott organiser, Charles Evers, had threatened black patrons who went to such stores. In this case, the US Supreme Court found that the words uttered by Evers (despite expressing a call for violence), amounted to political statements made at a public rally and as such, they did not constitute direct threats to people who disregarded the boycott.¹¹¹ The *ACLA* decision went on appeal. The Ninth Circuit Appeals Court (2002) found that the *ACLA* website constituted a 'true threat' that fell outside the scope of protected speech under the First Amendment, and that while 'advocating violence is protected, threatening a person with violence is not' and the *ACLA* was forced to shut down its site.¹¹²

Whilst it is acknowledged that the *ACLA* decision was a civil case, there have also been successful criminal prosecutions of Internet hate speech based on the 'true threat' doctrine. During 1996, a college student, Richard Machado, was convicted and sentenced to one year in prison for sending hate-based e-mail transmissions to Asian students at the University of California;¹¹³ in a similar incident during 1999, another college student was sentenced to two years in prison for sending anti-Hispanic e-mails to faculty members at the California State University and students at the Massachusetts Institute of Technology.¹¹⁴ An injunction was also brought against a college student for sending

¹¹⁰ 458 US 886 (1982).

¹¹¹ *NAACP v Claiborne* supra (n110) at 928-929.

¹¹² *Planned Parenthood v ACLA* 290 F. 3d (9th Circuit, 2002) 1058 at 1072. It is submitted that the web site had used 'threatening' information to identify the abortion providers (doctors) by their status, such as using the words 'working', 'wounded' and 'fatality' which were found to constitute 'true threats'. According to Henry, numerous attempts by the *ACLA* to bring further appeals have been declined by the US Supreme Court. As such, the definition of 'true threats' set out in the 2002 decision continues to be the prevailing law. Also see Henry op cit (n3) 238; 248.

¹¹³ Henry op cit (n3) 238. Also see Burch op cit (n8) 186.

¹¹⁴ Henry op cit (n3) 238.

e-mail messages threatening gay and lesbian students.¹¹⁵ The Federal government also brought charges against a website operator, Ryan Wilson, the leader of a white supremacist group for using his website to intimidate, threaten and harass a housing specialist, one Bonnie Jouhari and her teenage daughter (of mixed race).¹¹⁶ The above discussion demonstrates the US government's application of criminal law to regulate online threats targeting non-whites, Asians, Hispanics, gays and lesbians. The successful criminal prosecutions demonstrate that online hate speech directed at minorities and vulnerable groups will not be tolerated. It is submitted that these prosecutions send out a strong message that discriminatory practices should be eliminated and that a culture of tolerance and respect for diversity should be fostered by all American citizens.

The US government's attempts to regulate online material have met with limited success. As stated earlier, the anti-indecency provisions of the CDA were struck down by the US Supreme Court in *Reno*. The Child Online Prevention Act of 1998 ('COPA') was then passed to protect children from harmful content on the Internet, but this was challenged on First Amendment grounds.¹¹⁷ Attempts by Congress to regulate child pornography on the Internet through the Child Pornography Prevention Act of 1996 ('CPPA') were also struck down on First Amendment grounds in *Ashcroft v Free Speech Coalition*.¹¹⁸ The US Supreme Court found that the CPPA banned material that did not harm actual children and as such had scientific, literary, artistic or political value.¹¹⁹

Some states have drafted legislation criminalising certain types of webpages and other Internet communications.¹²⁰ To illustrate this, Pennsylvania has drafted a bill which makes it a first degree misdemeanour to threaten anyone electronically, including threats

¹¹⁵ Henry op cit (n3) 239.

¹¹⁶ *Department of Housing and Urban Development (HUD) v Wilson* July 19 2000. It should be noted that the Department of Justice had declined to pursue the matter because of First Amendment concerns; so the Housing Department (HUD) took up the matter because of Jouhari's employment in the housing sector. Also see Henry op cit (n3) 239-240.

¹¹⁷ Henry op cit (n3) 237.

¹¹⁸ 535 US 234 (2002). The Supreme Court found that the provisions of the CPPA were too broad and that it violated the freedom of speech. Also see Shytov op cit (n26) 271-272 for a discussion about the case.

¹¹⁹ Ibid. However, the passing into law of the Prosecutorial Remedies and Other Tools to End the Exploitation of Children Today Act ('PROTECT') to regulate child pornography has passed constitutional muster. See *United States v Williams* 128 S.Ct 1830 (2008).

¹²⁰ RJ Yen 'Free speech on the Internet: Regulating web authorship by students' (2000) 62 *Computer Law Rev & Technology J* 61 at 67.

through electronic mail or other Internet communications.¹²¹ This bill was passed in response to a student's use of a webpage to obtain funds to hire a hit man to murder his teacher.¹²² California has passed legislation that regards electronic words in the same way as written or spoken words, and speech that 'seriously threatens' is regarded as a crime.¹²³ States such as California and Pennsylvania have made valiant attempts to criminalise the use of threatening Internet communications. It is submitted that the above statutes protect the rights of online users to freely communicate and access information on the Internet without threats or harassment. Those individuals who have been specifically targeted online are also afforded protection from abuse and threats online by these statutes.

Two non-governmental organisations ('NGOs') such as the Southern Poverty Law Centre ('SPLC') and the Anti-Defamation League ('ADL') are using novel approaches to address online hate crime in the United States. The SPLC is using the Internet as weapon in the fight against hate groups, whilst the ADL works with ISPs to identify hate speech and to encourage enforcement of terms of service contracts in order to remove hateful content from the web.¹²⁴ The SPLC provides training to law enforcement agencies about hate crimes and domestic terrorism, provides information to law enforcement agencies in their investigation and prosecution of hate crimes, investigates and reports on extremist groups and disseminates its findings by way of Internet-based publications.¹²⁵ Its Teaching Tolerance Programme teaches children to embrace differences and turn away from hate.¹²⁶ Thus, the SPLC uses the Internet as an arsenal against hate groups. The ADL has partnered with YouTube to control bigoted online speech and to this end, YouTube launched its Abuse and Safety Centre during 2008 to identify and control hateful content online.¹²⁷ The ADL also uses the Internet to collaborate with law enforcement agencies and it launched the World Difference Institute (an education site) to provide anti-bias education and diversity training programmes.¹²⁸ According to the ADL,

¹²¹ A threatening webpage may alarm its target audience, with the result that the creator of that webpage may face prosecution in terms of the law. *Ibid.*

¹²² See MF Cattabiani 'Senate oks net threat legislation; Case of Bethlehem teen's anti-teacher web page inspired Dent's Bill' *Allentown Morning Call*, 24 March 1999, available at http://www.articles.mcall.com/1999-03-24/news/3233388_1_harass-threats-legislation, accessed on 10 September 2015.

¹²³ See Thomas and Loader *op cit* (n3) 248 for further discussion about US state legislation regulating hate speech.

¹²⁴ Henry *op cit* (n3) 236.

¹²⁵ Henry *op cit* (n3) 242.

¹²⁶ Henry *op cit* (n3) 243.

¹²⁷ Henry *op cit* (n3) 245.

¹²⁸ Henry *op cit* (n3) 245-246.

evidence demonstrates that hate online inspires hate crimes, with transgressors reporting that they were introduced to racist ideology and teachings of hate and intolerance on anti-government sites and extremist sites online.¹²⁹

The United States is a party to international agreements such as the Universal Declaration of Human Rights,¹³⁰ the International Covenant on Civil and Political Rights¹³¹ and the UN Convention on the Elimination of All forms of Racial Discrimination,¹³² which requires the US to enact legislation or take positive measures to prohibit hate speech or eradicate discrimination. The United States government's commitment to the protection of the First Amendment has resulted in it signing such agreements with reservation.¹³³ The United States is also a signatory to the Cybercrime Treaty, which includes an Additional Protocol to the Convention on Cybercrime addressing the criminalisation of acts of a racist and xenophobic nature committed through computer systems. However, the United States did not sign the protocol relating to online hate speech because of First Amendment concerns.¹³⁴ This demonstrates that the United States will not accede to any international treaty that calls for the restriction of hate speech on the Internet. Similarly, the United States government will not extradite a US citizen wanted for dissemination of hate literature in another country if such behaviour is considered to be legal in the United States.¹³⁵

The above discussion demonstrates that the First Amendment presents challenges to any hate crime prosecution on the Internet. The United States government has also reserved the right to protect free speech above its obligations to international treaties and agreements. The absence of an international treaty will create jurisdictional problems for litigation lawyers in the United States and abroad. It has been mooted that the right to free Internet speech must continue to enjoy the protection of the First Amendment in the United States,

¹²⁹ Wolf op cit (n6) 12-13; Cohen-Almagor op cit (n3) 128-129.

¹³⁰ The Universal Declaration contains the following articles relating to freedom of speech: article 18 regulates free expression; article 19 addresses freedom to express; article 29(2) addresses the exercise of rights limited by law for 'morality, public order and the general welfare in a democratic society', whilst article 30 states that one cannot engage in an activity to destroy any rights or freedoms.

¹³¹ Article 20(2) prohibits any advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence.

¹³² Article 4 prohibits all propaganda and all organisations based on ideas or theories of superiority of one race group of persons or one colour or ethnic origin or the justification or promotion of racial hatred and discrimination.

¹³³ See Spector op cit (n19) 166 and Timofeeva op cit (n1) 254.

¹³⁴ Henry op cit (n3) 240.

¹³⁵ See for example, the case of Gary Lauck, an American citizen wanted in Germany for Nazi-related offences. See Henry op cit (n3) 241 and Cohen op cit (n14) 247.

whilst other countries must look for solutions within their own countries.¹³⁶ The *Reno* case protects ISPs from liability for the contents of its user websites; however ISPs can voluntarily choose to regulate such speech.¹³⁷ Due to the increased power of the Internet to reach both public and private places, courts and legislative bodies will face the difficult challenge of creating legal standards that will protect individual safety and privacy, while preserving free-speech rights protected in the First Amendment.¹³⁸ NGOs such as SPLC and ADL can provide positive online messages to counter online bigoted speech.¹³⁹ The roles of SPLC and ADL in the United States in addressing online hate speech are commended.

6.2 United Kingdom

It is an offence in the United Kingdom ('UK') to stir up hatred on the grounds of race, religion and sexual orientation.¹⁴⁰ Many statutes prohibit expressions of hatred towards someone on the basis of colour, race, nationality, ethnic or national origin, religion or sexual orientation.¹⁴¹ The Criminal Justice and Public Order Act of 1994 prohibits any communication which is threatening or abusive and which is intended to harm, alarm or distress someone.¹⁴² The Protection from Harassment Act of 1997 prohibits conduct including speech which amounts to harassment of another person.¹⁴³ Although the above Act does not specifically relate to hate speech, it can be regarded as another tool in combating hate speech.¹⁴⁴ It is submitted that the above statutes prohibit discrimination on the basis of *inter alia*, race, religion and sexual orientation. This demonstrates that the UK government is commendably striving to tackle discrimination and

¹³⁶ Spector op cit (n 19) 155.

¹³⁷ An example is America Online (AOL) which adopted a term of service contract. Users could violate this contract if they disseminated content that was found to be racially or ethnically offensive. Henry op cit (n3) 247.

¹³⁸ Yen op cit (n120) 77.

¹³⁹ Henry op cit (n3) 248.

¹⁴⁰ 'True vision' available at http://www.report.it.org.uk/reporting_internet_hate_crime, accessed on 2 March 2015.

¹⁴¹ See *inter alia* the Public Order Act of 1986, Race Relations Act of 1965, Racial and Religious Hatred Act of 2006 and Criminal Justice and Immigration Act of 2008. The Race Relations Act of 1965 criminalises public speech including 'threatening, abusive, insulting' publications and prohibits incitement to hatred. The Racial and Religious Hatred Act of 2006 criminalises speech stirring up racial and religious hatred. See Cohen op cit (n14) 241-242.

¹⁴² See s 5. The penalties for hate speech include fines, imprisonment or both. See Cohen op cit (n14) 241 and Thomas and Loader op cit (n3) 246.

¹⁴³ See s 1.

¹⁴⁴ Cohen op cit (n14) 241.

at the same time, promote respect for tolerance and diversity among its citizens.

The United Kingdom is a party to the European Convention for the Protection of Human Rights and Fundamental Freedoms 1950 ('ECHR').¹⁴⁵ Article 10 of the ECHR protects the right to freedom of expression which includes the right to hold opinions and to receive and impart information and ideas without interference by a public authority.¹⁴⁶ However, this right is subject to conditions, restrictions and penalties prescribed by law and which are necessary in a democratic society.¹⁴⁷ Indeed, freedom of expression is said to be one of the cardinal rights guaranteed under the ECHR, and the interference with this right must be weighed against whether there is a less intrusive way to protect the public interest.¹⁴⁸ The Human Rights Act of 1998 codifies the protections in the ECHR into UK law.¹⁴⁹

It is accepted that the Internet has become a central component to the exercise of free speech.¹⁵⁰ The United Kingdom has introduced a self-regulatory organisation called the Internet Watch Foundation ('IWF') to reduce the availability of child sexual abuse and criminally obscene content on the Internet.¹⁵¹ The IWF creates a blacklist for unacceptable content, and filtering technologies are used to disable access to such contents by members.¹⁵² The blacklist is available to national and international law enforcement agencies and INHOPE hotlines (International Association of Internet Hotlines).¹⁵³ There are no structures in place to guard against the misuse of power by IWF employees. Therefore, the activities of the IWF are considered to be a

¹⁴⁵ It should be noted that the UK has a statutory Bill of Rights, the Human Rights Act, 1998 which incorporates the ECHR.

¹⁴⁶ See article 10(1) of the ECHR.

¹⁴⁷ See article 10(2) of the ECHR.

¹⁴⁸ EB Laidlaw 'The responsibilities of free speech regulators: An analysis of the Internet Watch Foundation' (2012) 20 *Internat'l J Law & Info Technology* 213 at 329.

¹⁴⁹ The Human Rights Act was introduced in the UK during October 2000. Article 10 refers to the right to freedom of expression. The right to freedom of expression can be restricted in circumstances where the views expressed encourage racial and religious hatred. However, such restrictions must be necessary, appropriate and not be excessive in the circumstances. See 'Equality and Human Rights Commission,' available at <http://www.equalityhumanrights.com>, accessed on 9 September 2015.

¹⁵⁰ Laidlaw op cit (n148) 331.

¹⁵¹ Its members are Internet Service Regulators. Key policy decisions concerning content that is filtered in the UK are made by the IWF. Positive obligations are imposed on the state regarding the governance of the IWF in terms of article 10 of the ECHR. See Laidlaw op cit (n148) 312, 326.

¹⁵² Laidlaw op cit (n148) 312.

¹⁵³ Ibid. It should be noted that INHOPE enables members of the public to anonymously report online content they suspect to be illegal such as child sexual abuse material and illegal types of hate speech. See McGonagle op cit (n1) 433.

disproportionate interference with the right to freedom of expression and thus a breach of article 10(2) of the ECHR.¹⁵⁴ The Electronic Commerce Directive (E-Commerce Directive) is a piece of legislation that regulates some of the responsibilities of the IWF members. It was introduced into UK law through the Electronic Commerce ('EC Directive') Regulations 2002 No 2013, which prescribes the circumstances for liability for unlawful content communicated by a third party in terms of the Information Society Service ('ISS').¹⁵⁵ Services that would qualify as an ISS would include activities of IWF members.¹⁵⁶ Unlawful content includes obscene content, terrorism-related content and content that stirs up religious or racial hatred.¹⁵⁷ A call for a better understanding of the business and human rights dynamic online is required with a clearer awareness of the cross-over of regulatory and rights-based systems of governance.¹⁵⁸ Until April 2011, the IWF's power also extended to incitement to racial hatred content hosted in the UK, but such content has now been re-directed to a new police body called True Vision.¹⁵⁹ The True Vision website is a website that collates information on hate crimes and reports on such crimes in the United Kingdom.¹⁶⁰ Thus, there is clear intent in the UK to prosecute offenders for the publication of hate material irrespective of the medium used for publication.¹⁶¹

There have been cases where individuals have been prosecuted for violations.¹⁶² A Neo-Nazi was recently jailed for attacking an Asian

¹⁵⁴ IWF employees are only questioned about their activities if a user or website owner questions the blocking of a site. See Laidlaw *op cit* (n148) 335.

¹⁵⁵ Laidlaw *op cit* (n148) 319-320. It should be noted that the National Criminal Intelligence Service is also combating racism on the Internet and acting against threatening, abusive and racist material. The information is then passed to relevant national authorities for further investigation. See Thomas and Loader *op cit* (n3) 246.

¹⁵⁶ Laidlaw *op cit* (n148) 320.

¹⁵⁷ *Ibid.* If an IWF member learns about any unlawful content being communicated by a third party, it must act quickly to remove or disable access to such content.

¹⁵⁸ Laidlaw *op cit* (n148) 3-5.

¹⁵⁹ Laidlaw *op cit* (n148) 316.

¹⁶⁰ Examples of hate material reported on may include illegal material; words, pictures, videos or messages of racial or religious violence; web pages, pictures or videos that glorify violence against anyone on the grounds of race, religion, disability, sexual orientation and chat forums where people ask others to commit online crime. See further, 'True Vision' *op cit* (n140).

¹⁶¹ Thomas and Loader *op cit* (n3) 13.

¹⁶² According to Cohen, during 1994-2004, there were about 37 prosecutions for incitement to racial hatred in Britain. See Cohen *op cit* (n14) 252.

man online.¹⁶³ The case of *Regina v Sheppard & Whittle*¹⁶⁴ involved the prosecution and conviction for possession, publication and distribution of racially inflammatory material online. During March 2012, a student, Liam Stacey, was jailed for 56 days after racially abusing two Twitter users and mocking midfielder Fabrice Muamba after his collapse during a soccer match.¹⁶⁵ Similarly, the racial tweets about two black English footballers (Ashley Young and Ashley Cole) who missed penalties in England's quarterfinal loss to Italy during Euro 2012, were regarded as 'appalling and unacceptable' by the British Football Association and the British police launched an investigation into the tweets.¹⁶⁶ However, the above prosecutions have been criticised as being selective and displaying 'politically correct' attitudes to hate speech.¹⁶⁷ Nevertheless, the above discussion demonstrates that the UK government has taken concerted steps to address hate speech in different media platforms, to protect online users from threats and harassment, and to promote respect for diversity.

6.3 The position in South Africa

There has been an increase in the posting of racist comments on social networking sites in South Africa recently.¹⁶⁸ A top-level employee of the Market Theatre in Johannesburg was fired after a disciplinary committee found him guilty of hate speech on Facebook, when he posted derogatory comments about Jewish directors in the South African theatre industry on his Facebook profile.¹⁶⁹ There is no specific legislation addressing online hate speech in South Africa or regulation of online content in South Africa *per se*. However, the dissemination of hate speech can be addressed by remedies found in criminal law or the Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000 ('PEPUDA').¹⁷⁰ It has been mooted that South Africans who post racist and offensive comments on social networking sites may be

¹⁶³ 'True Vision' op cit (n140).

¹⁶⁴ [2010] EWCA Crim 65. The offender had distributed a pamphlet which was published online denigrating Jewish people and stirring racial hatred. The court held that the offence of displaying, distributing or publishing racially inflammatory written material does not require proof that anybody actually read or heard about the offensive material. See Cohen op cit (n14) 242.

¹⁶⁵ See 'Investigation into racial tweets' op cit (n64).

¹⁶⁶ Ibid.

¹⁶⁷ See Cohen op cit (n14) 253-254.

¹⁶⁸ See Naik op cit (n5) 9 and 'Internet' op cit (n3).

¹⁶⁹ See 'Hate speech on Facebook gets employee fired' 10 March 2011, available at <http://legalbrief.co.za/story/hate-speech-on-facebook-gets-employee-fired/>, accessed on 16 March 2011.

¹⁷⁰ See s 16 of the Constitution and the discussion on PEPUDA below.

charged with *crimen injuria*.¹⁷¹ The term *crimen injuria* refers to the unlawful, intentional and serious violation of the dignity or privacy of another person.¹⁷² Facebook users may also be sued for damages as a result of defamatory remarks made on Facebook.¹⁷³

The Promotion of Equality and Prevention of Unfair Discrimination Act 4 of 2000 ('PEPUDA') which was introduced to promote the equal treatment of persons, may also be used to address online hate speech. Section 10 of PEPUDA prohibits the publication or communication of hate speech which is intended to cause hurt, harm or incite hatred.¹⁷⁴ The prohibition is said to apply to hate speech based on the prohibited grounds set out in s 1(1) of PEPUDA.¹⁷⁵ It is submitted that the PEPUDA test for hate speech could be generously interpreted to mean that the hateful words should demonstrate a clear intention to be hurtful rather than an intention to incite harm or violence. The Equality Court has been established in terms of s 16(1) of PEPUDA to hold inquiries to

¹⁷¹ A person found guilty of *crimen injuria* may also face imprisonment. See Naik op cit (n5) 9. See s 16(2)(a)-(c) of the Constitution which excludes hate speech from constitutional protection.

¹⁷² The criminal offence of *crimen injuria* may be used to apply to acts of online hate speech or racist speech. See CR Snyman *Criminal Law* 6ed (2014) 461 regarding the definition of *crimen injuria*. Also see C Badenhorst *Legal Responses to Cyber Bullying and Sexting in South Africa* (2011) (Centre for Justice and Crime Prevention Issue paper 10) 1-20 at 8.

¹⁷³ See *H v W* 2013 (2) SA 530 (GSJ), 2013 (5) BCLR 554 (GSJ), [2013] 2 All SA 218 (GSJ). The offender was ordered to remove all postings relating to the victim or applicant on all social media sites and to pay the costs of the application. Also see K Ajam 'SA man wins interdict for slander on Facebook' *Saturday Star*, 2 February 2013 1; and E Sadleir 'No need to go to court over offensive Facebook posts' (2013) *Saturday Star*, 2 February 2013 5. The judgment also stressed the need for the law to develop to keep abreast with technological innovations. Also see F Cassim 'Formulating adequate legislation to address cyber-bullying: Has the law kept pace with advancing technology?' (2013) 26 *SACJ* 1-20, for further discussion about the case and the impact of new technologies.

¹⁷⁴ See s 10(1) of Act 4 of 2000. Also see *Hate Speech and Freedom of Expression in South Africa* (2013), available at http://www.fxj.org.za/home/fxi_downloads/Hate_Speech_and_Freedom_of_Expression_in_SA.pdf, accessed on 13 April 2015.

¹⁷⁵ The prohibited grounds encompass race, gender, sex, pregnancy, marital status, ethnic or social origin, colour or sexual orientation, age, disability, religion, conscience, belief, culture, language and birth or any other ground which causes systematic disadvantage, undermines human dignity or adversely effects the equal enjoyment of a person's rights and freedom in a serious manner that amounts to discrimination. See further, s 1(1)(a)-(b) of PEPUDA.

examine allegations of hate speech.¹⁷⁶ It is noteworthy that s 10 does not preclude the *bona fide* engagement of artistic creativity, academic and scientific inquiry, fair and accurate reporting in the public interest or publication of any information, advertisement or notice in terms of s 16 of the Constitution.¹⁷⁷

The Films and Publication Act 65 of 1996 ('FPA') also prohibits the distribution of hate speech. According to s 29(1)(c) of the FPA, any person who knowingly distributes a publication which advocates hatred that is based on race, ethnicity, gender or religion and which constitutes incitement to cause harm is guilty of an offence. A publication includes any message or communication including a visual presentation that is placed on any distributed network including but not limited to the Internet.¹⁷⁸ The word 'knowingly' has important ramifications: A Internet Service Provider ('ISP') which has no knowledge of the distribution of the content of the material cannot be found guilty of distributing hate speech; similarly a user who downloads hate speech also does not commit this offence provided that he or she does not distribute the hate speech any further.¹⁷⁹ However, there are exceptions to the provisions of s 29(1)–(3) of the FPA, which prohibits the broadcast, distribution, exhibition and presentation of material advocating *inter alia*, war, violence and hatred. These exceptions encapsulate the following: a *bona fide* scientific, documentary, literary or religious publication, film, entertainment or play; or a publication, film, entertainment or play which results in a

¹⁷⁶ See s 21(1) of PEPUDA. The South African Human Rights Commission ('SAHRC') recently accused Afrikaans singer Sunette Bridges of hosting racial commentary on her Facebook page which constitutes hate speech in terms of the Equality Act (PEPUDA). The Equality Court held that the comments posted on Ms Bridge's Facebook page may constitute hate speech and harassment in terms of s 10(1) of PEPUDA. Bridges reached a settlement with the SAHRC which was made an order of the Equality Court in the Western Cape High Court on 31st March 2015. According to her lawyer William Booth, websites allowing publication of contents will now have to direct more attention to the issue of defamation and hate speech to avoid possible court action. See 'Hate speech test case against Afrikaans singer,' available at <http://legalbrief.co.za/story/hate-speech-test-case-against-afrikaans-singer/>, accessed on 21 January 2015; J Etheridge 'Law expert: Sunette Bridges order could set precedent' 1 April 2015, available at <http://www.news24.com/Law-expert-Sunette-Bridges-order-could-set-precedent-20150401>, accessed on 2nd April 2015; and I Mangena 'Equality court orders Sunette Bridges to ensure she does not promote hate speech, harassment and violence on her Facebook pages' 31 March 2015, available at www.sabrc.org.za/home/index.php?ipkArticleID=330 accessed on 23 April 2015.

¹⁷⁷ See s 12 of PEPUDA. Also see Van der Merwe et al op cit (n3) 447.

¹⁷⁸ See s 1 of the FPA.

¹⁷⁹ Van der Merwe et al op cit (n3) 448.

bona fide discussion, argument or opinion on a matter pertaining to religion, belief or conscience; or a publication, film, entertainment or play which results in a *bona fide* discussion, argument or opinion on a matter of public interest.¹⁸⁰ A contravention of s 29 will result in a fine or a term of imprisonment not exceeding five years or both if aggravating factors exist.¹⁸¹

The Protection from Harassment Act 17 of 2011 came into effect on 27 April 2013.¹⁸² The definition of harassment in the Act may encompass online hate speech since online hate speech may involve the threat to harm or harass a particular racial group or individual.¹⁸³ The Act refers to the use of 'electronic communications' and the sending of messages by 'electronic mail' to 'cause harm or inspire the reasonable belief that harm may be caused' in the definition of harassment.¹⁸⁴ The enactment of this legislation is regarded as a step in the right direction because it can be used as another tool in combating online hate speech.¹⁸⁵ However, criticism has been levelled at the Act because protection for victims of harassment is dependent on the victims knowing the identity and location of the perpetrators.¹⁸⁶

The Electronic Communications and Transactions Act 25 of 2002 ('the ECT') was introduced to address *inter alia*, the facilitation and regulation of electronic communications and to prevent abuse of information systems. Cybercrime is addressed comprehensively in chapter 13. Section 85 defines 'access' as the action of a person

¹⁸⁰ See s 29(1)-(3) of the FPA.

¹⁸¹ See s 30(1) and (2) of the FPA.

¹⁸² 'Department of Justice and Constitutional Development champions anti-harassment Act' *South African Government Information*, 28 April 2013, available at <http://www.gov.za/departments-justice-and-constitutional-development-champions-anti-harassment-act>, accessed on 21 January 2015; also see 'Protection from harassment bill approved', 16 August 2011, available at <http://www.news24.com/SouthAfrica/Politics/Protection-from-harassment-bill-approved-20110816>, accessed on 15 March 2013; L Bick 'Freedom day fitting for launch of protection act' *Saturday Star*, 14 April 2013, 14.

¹⁸³ See s 1 of the Act. This legislation emanates from the South African Reform Commission's Discussion Paper 108 (Project 130) *Stalking* (2004). Also see Badenhorst op cit (n172) 14-15 for a discussion about harassment law in South Africa.

¹⁸⁴ See s 1. It is submitted that online hate speech may involve a threat to harm the victim and the perpetrator may be unknown to the victim. Harm may be direct or indirect and it may include mental, psychological or economic harm.

¹⁸⁵ Anti-harassment laws are regarded as valid regulators of hate speech because these laws are directed at intent and effect. See Burch op cit (n8) 186.

¹⁸⁶ This requirement is problematic since the anonymity and accessibility of the Internet makes it difficult for victims and enforcement agencies to track down offenders. See Badenhorst op cit (n172) 15.

who after taking note of any data, becomes aware of the fact that he or she is not authorised to access that data and still continues to access that data. Online hate speech may involve the potential abuse of electronic communications or information systems to degrade, humiliate or threaten another person or a group of persons due to *inter alia*, race, ethnicity and religious grounds. Similarly, a person may 'hack' into a person's computer to intimidate, harass or threaten him or her. Therefore, it can be argued that actions amounting to online hate speech may conceivably fall under the scope of cybercrime, and online hate speech can therefore be addressed by chapter 13 of the ECT. Although the ECT does not specifically address online hate speech, it is submitted that the ambit of s 86 can be widened to address the dissemination of such speech.¹⁸⁷ The liability of ISPs is regulated by chapter 11 of the ECT. The question arises whether ISPs can face both civil and criminal liability in instances where they act as mere conduits. However, such limitations do not affect a court order to remove, block or deny access to any data message that amounts to hate or racist speech.¹⁸⁸

The above discussion demonstrates that there is legislation in place in South Africa to curb online hate speech; such as, *inter alia*, PEPUDA and the FPA. It is noteworthy that the preamble in PEPUDA alludes to *inter alia*, principles of 'diversity, fairness, justice and human dignity' and 'human relations that are caring and compassionate'. Section 2 of PEPUDA promotes the values of non-racialism and non-sexism. The FPA prohibits the distribution of any publication espousing hate speech, except where the distribution, broadcast, exhibition and presentation is for *bona fide* purposes. Similarly, s 12 of PEPUDA also protects *inter alia*, the distribution or use of publications or communications for *bona fide* purposes. Thus, it is submitted that both PEPUDA and the FPA promote a culture of respect for tolerance and diversity and protect online users who use the Internet for *bona fide* purposes. Both the Protection from Harassment Act and the ECT regulate the use of electronic communications and prevent the abuse of information systems. In this way, these statutes protect online users from threats, harm and harassment committed online.

¹⁸⁷ Section 86 prevents unauthorised access to or interception of or interference with data.

¹⁸⁸ This may well mean that the Equality Court has the authority to order the ISP to remove, block or deny access to hate or racist speech in terms of PEPUDA. See Van der Merwe et al op cit (n3) 447. Also see F Marx 'Hate speech on social network sites: perpetrator and service provider's liability' (2011) 32 *Obiter* 322-340 regarding liability of ISPs for hate speech posted by third parties on their websites.

7 Conclusion

The popularity of the Internet and the introduction of new social media have opened up new channels of communication. However, the global nature of the Internet, and the ease with which illegal materials can be produced, distributed and accessed in many countries in the world, have created a major problem for computer users and governments worldwide.¹⁸⁹ The Internet has also brought new challenges and introduced new opportunities for criminal behaviours such as online hate speech. The dissemination of online hate speech may result in psychological and emotional harm and its effects can be devastating.¹⁹⁰ Measures are required to adequately regulate the Internet to curb online hate speech. Indeed, introducing legislation to adequately regulate online content sends a strong message to potential transgressors to take note that their actions will not go unpunished. A need thus arises to address the emergence of new challenges that new technologies pose. However, one should not lose sight of constitutional freedoms such as freedom of expression. After all, free-speech guarantees are foundational to a democratic society and should not be infringed except in extraordinary circumstances.¹⁹¹ The above discussion has demonstrated that while the US views freedom of expression as worthy of protection in almost absolute terms (except where the speech amounts to 'true threats'), European countries and the international community seek to promote values of pluralism, tolerance and broadmindedness as being essential to a democratic society.¹⁹² It is submitted that the introduction of any legislation regulating online content should encourage online users to express their free speech and diverse opinions in a responsible manner without government censorship.¹⁹³ Technology advances faster than the law; hence the need also arises to amend and strengthen the law to address such challenges.¹⁹⁴

The tension between hate speech laws of countries is exacerbated by the character of the Internet, which transcends jurisdictional boundaries. The Internet is said to present numerous problems because of the cultural diversity of the planet; hence a new standard for free

¹⁸⁹ Kierkegaard op cit (n6) 42.

¹⁹⁰ See *inter alia*, Shaw op cit (n6) 282 and McGonagle op cit (n1) 433-434.

¹⁹¹ Yen op cit (n120) 84. However, see Brennan op cit (n32) 124 who is in favour of placing anti-racism as the pinnacle of a mark of a decent society rather than freedom of speech as sacrosanct, because freedom of expression is used as a vehicle to vent racial hatred on the Internet.

¹⁹² Cohen op cit (n14) 238. Also see the discussion in sections 5, 6.1 and 6.2 above.

¹⁹³ However, governments should be wary of censorship since overzealous censorship is as much a threat to human dignity as hateful propaganda. Shaw op cit (n6) 285.

¹⁹⁴ Yen op cit (n120) 55.

speech should not be ushered in; rather, it has been mooted that 'end-user' solutions should be investigated to overcome jurisdictional problems to avoid straining international relations between countries as illustrated by the *Yahoo!* cases.¹⁹⁵ Present attempts by the international community to curb hate speech provide good arsenals against hate speech on the Internet.¹⁹⁶ However, more can be done. It has been recommended that more concrete steps should be taken to address online hate speech such as alerting users about sites espousing hate ideology; indexing search engines so that hate sites can be properly displayed; encouraging the responsible use of filtering technologies by countries; making ISPs more responsible and accountable for hate ideology posted on their websites; the international community should introduce more cohesive and uniform hate speech regulations, employ more effective and consistent prosecution strategies and introduce an educational process that outlaws race hate through whichever medium it is disseminated.¹⁹⁷ It is submitted that governments should also use education and public awareness campaigns to promote intercultural dialogue among citizens and eliminate negative stereotyping of minorities and vulnerable groups. Some governments may censor online content to prevent its citizens from viewing hate content hosted in other countries that allow such hate speech; however, such government censorship of Internet content may in the long run result in the exclusion of some information that is not necessarily hateful or offensive.¹⁹⁸ States should rather regulate the content of the Internet internationally and they should engage in international co-operation to address online content.¹⁹⁹ An international harmonisation of laws will also assist in stamping order on a 'wild and shifty Web'.²⁰⁰ The deeper societal problems of hate speech also need to be investigated

¹⁹⁵ Spector op cit (n19) 177.

¹⁹⁶ See for example, the Additional Protocol to the convention on cybercrime, concerning the criminalisation of acts of a racist and xenophobic nature committed through computer systems (adopted 7/11/2002). Steps taken in the US, UK and South Africa to combat online hate speech are noteworthy. The roles of NGOs such as ADL and SPLC in the US and the policing body True Vision in the UK are also commendable. However, governments should be wary of undermining principles of free academic research. See Cohen op cit (n14) 244.

¹⁹⁷ See Cohen-Almagor op cit (n3) 130; Thomas and Loader op cit (n3) 249; Timofeeva op cit (n1) 285; Cohen op cit (n14) 255; Tsesis op cit (n1) 864-868 and Perry and Olsson op cit (n1) 196.

¹⁹⁸ Henry op cit (n3) 241.

¹⁹⁹ It is submitted that a more cohesive effort by all role players is advocated. Also see Shytov op cit (n26) 268-269; Nuth op cit (n2) 446; Timofeeva op cit (n1) 280. There should also be regular monitoring by states of the use of the Internet for hateful purposes. See Wolf op cit (n6) 14 and Chen op cit (n1) 232, 267.

²⁰⁰ This phrase means that the Web or Internet is always changing due to new technologies and new challenges. Shytov op cit (n26) 268-269.

and studies conducted as to how vulnerable people become exposed to hate sites online.²⁰¹ At the end of the day, computer users should freely embrace the Internet to share ideas, respect differences, turn away from hate speech and hate provocation, and embrace 'pluralism, tolerance and broadmindedness'.²⁰²

²⁰¹ Wolf op cit (n6) 14.

²⁰² These principles are a hallmark of a democratic society. See McGonagle op cit (n1) 436.