

SELF-CONTROL ON NETWORKS!

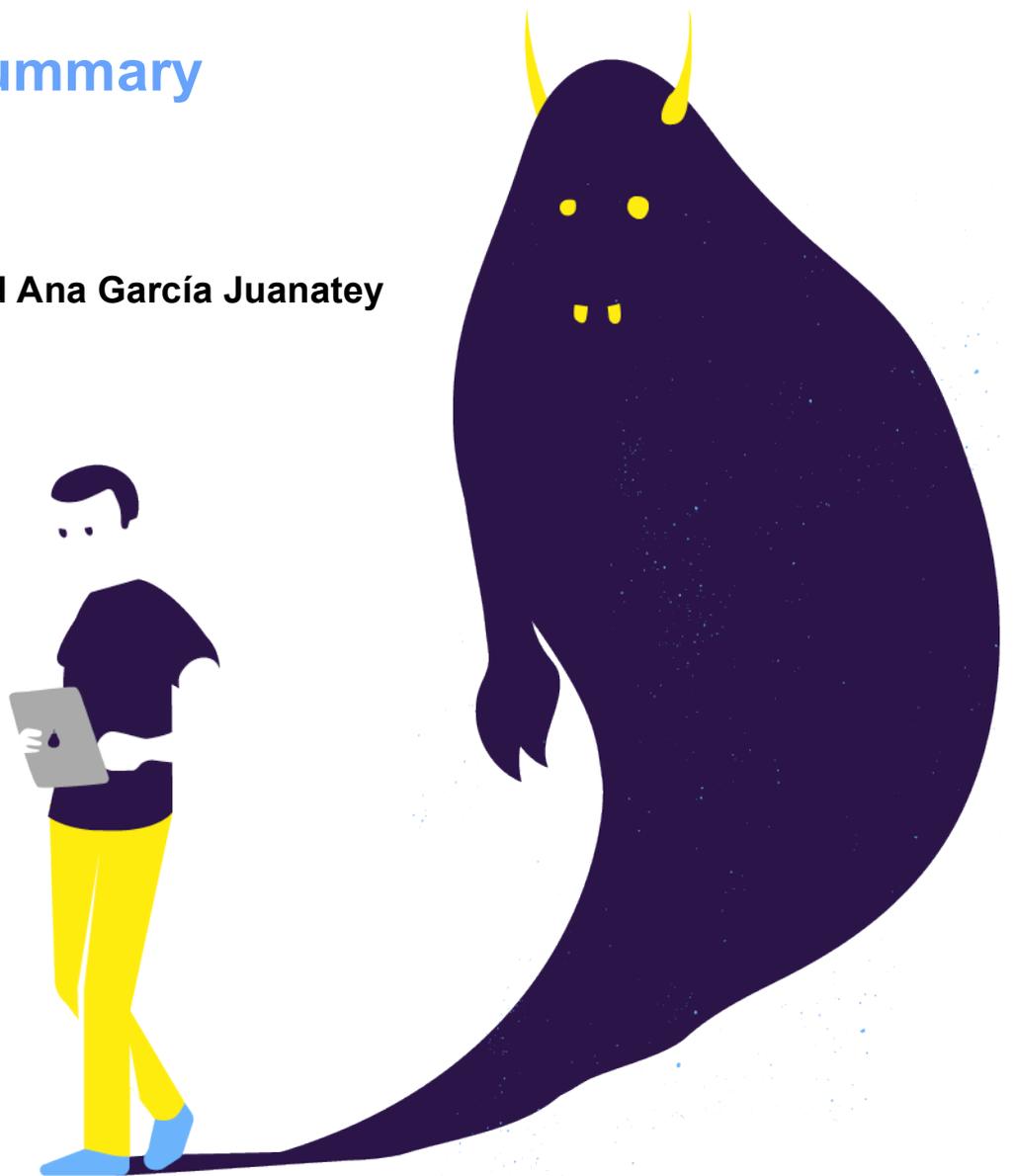
#BCN vs ODI

**Hate speech in social
media: a state-of-the-
art review**

Executive summary

12/12/2016

Alex Cabo Isasi and Ana García Juanatey



SELF-CONTROL ON NETWORKS!

2 and 3 March 2017 at the CCCB
What we do in networks has repercussions.

#BCN vs ODI



Contents

1.	THE PROBLEM OF HATE SPEECH IN SOCIAL MEDIA	1
1.1.	What is understood by hate speech?.....	1
1.2.	The particular features of online hate speech	2
1.3.	The consequences of hate speech	2
1.4.	Who are the perpetrators of hate speech?.....	2
2.	RESTRICTING HATE SPEECH	3
2.1.	The dilemma between freedom of expression and suppression of hate speech 3	
2.2.	Legal aspects of hate on the Internet and in social media	4
2.3.	Spanish legislation	4
3.	THE ROLE OF SOCIAL MEDIA AS INTERMEDIARIES	5
3.1.	Attributing responsibility	5
3.2.	Self-regulation and installing computer filters	6
3.3.	The Code of Conduct with the European Union.....	6
4.	EXTRAJUDICIAL STRATEGIES FOR COMABTING HATE SPEECH IN SOCIAL MEDIA	7
4.1.	Monitoring and research strategies.....	7
4.2.	Strategies for pressurising social media	7
4.3.	Strategies designed to change perceptions and attitudes	7
4.3.1	<i>Campaigns against online hate speech</i>	7
4.3.2	<i>Strategies against online hate speech based on education and training</i>	9
	BIBLIOGRAPHY	9

1. THE PROBLEM OF HATE SPEECH IN SOCIAL MEDIA

There is a general feeling that extreme speech has a disproportionate presence in social media. Sexism, homophobia, xenophobia, Islamophobia, Romaphobia, anti-Semitism and other forms of intolerance use the Internet and social media to insult, humiliate, harass, threaten and incite violence.

In a Europe in economic crisis and suffering a crisis of identity, manifestations of hate, anger and aggressiveness have become common currency on the Net. In the wake of the refugee crisis, and especially in the days following the terrorist attacks in Paris, Brussels and Nice, there were big spikes in hate on social media.

International and European institutions are not immune to this. The problem of hate speech online, and particularly on social media, has become such a big issue that today it can be found on the desks of most European and international bodies.

According to studies carried out by the think tank Demos, Twitter has approximately 10,000 tweets a day with racist insults in English, and every 10 seconds a woman is insulted with the word “whore” or “bitch”. There is no doubt that social media have become a place where people can express their anger and hatred with impunity.

The problem is no less serious in Spain. At a collective level, messages that insult and denigrate women, Gypsies, Muslims, Jews, homosexuals, transsexuals, etc., are worryingly common. The president of SOS Racismo Madrid, Moha Gerehou, the Mayor of Terrassa, Jordi Ballart, a member of the Madrid Assembly, Carla Antonelli, and a youth football referee, Jesús Tomillero, are some of the cases of racist and homophobic threats and harassment that have appeared in the media in 2016, and which are only the tip of the iceberg of a situation that has become normal in social media.

1.1. What is understood by hate speech?

Although there is a degree of consensus over the extent of the problem, the concept of hate speech is still being discussed and is the subject of debate at an international level. It is a complex concept, because it challenges fundamental principles for democracies such as equality, human dignity and freedom of expression, and can therefore be used with more or less legitimate aims.

The term “hate speech” is used to refer to expressions that incite discrimination or violence due to racial hatred, xenophobia, sexual orientation and other types of intolerance, but also to refer in broader terms to those expressions that foster hostility through prejudice and intolerance. In the media, the concept is widely used to refer to a mass of expressions that include manifestations as diverse as mobbing famous personalities, threats to individuals or groups inspired by intolerance, offensive remarks against those in power, advocating terrorism and religious blasphemy, which require different approaches and solutions.

Consequently, in this report, the concept of hate speech is used to refer to expressions which, in a broad sense, foster and incite violence, discrimination and hostility towards individuals or groups for reasons of race, xenophobia, religion, gender, sexual

orientation or disability and other forms of intolerance.

1.2. The particular features of online hate speech

The Internet and social media give hate speech a set of characteristics that help to create a feeling of impunity or turn it into a phenomenon that is out of control and causes more harm:

- Communication is decentralised and on a massive scale.
- Social media have enormous multiplying potential.
- The contents stay out there indefinitely if they are not deleted.
- The contents can jump to another platform.
- The use of pseudonyms and anonymity is widespread.
- They are a medium “without borders”.
- The feeling of “virtualness” helps people lose their inhibitions.

1.3. The consequences of hate speech

Besides turning social media into an uncivilised communication space, hate speech has harmful consequences in the offline world:

- The emotional and psychological damage caused to victims of threats and harassment motivated by hate and intolerance.
- The erosion of the dignity and reputation of individuals and groups.
- The perpetuation of discriminatory stereotypes, dehumanisation and stigmatisation of groups.
- There is less empathy towards dehumanised groups.
- It creates fertile ground for hate crime offences and discriminatory acts.

1.4. Who are the perpetrators of hate speech?

A large part of the hate speech found on social media comes from individuals directly or indirectly linked to fascist groups such as Hogar Social Madrid, Plataforma per Catalunya, Falange Española de las JONS, España 2000, Alianza Nacional, and groups of fanatical football fans known as *ultras*.

However, the phenomenon of hate speech is not exclusive to these far-right groups. In Internet slang it is usual to use the terms “haters” or “trolls” to refer to individuals obsessed with attacking specific groups or who spend their time provoking other users by insulting and being aggressive towards them. These are the types of people responsible for a large part of the hate speech found on social networking sites. But to really understand the scale of the problem and design suitable strategies for combating

it, we should bear in mind that these phenomena, which are a part of Internet culture, are not entirely homogeneous.

2. RESTRICTING HATE SPEECH

2.1. The dilemma between freedom of expression and suppression of hate speech

At an international level, the different approaches to restricting hate speech are usually reduced to the division between a more liberal and “tolerant” American perspective and the European perspective, which is more “intransigent” towards hate speech.

- The USA has traditionally been characterised by its uncompromising defence of freedom of expression.
- Europe, however, has generally taken a harder line against hate speech and rejects the possibility of accommodating it under freedom of expression to protect Holocaust deniers, terrorist apologists and xenophobic or racist messages.

The following table shows, in a simplified form, some of the main arguments that are usually used for and against banning hate speech.

For	Against ¹
Freedom of expression is not an absolute right. It has to be restricted to protect human dignity, equality, peace and social harmony, the right to live without being harassed and intimidated, etc.	Freedom of expression and public debate are essential for a democracy. Restrictions should not be imposed simply on the basis of how offensive certain expressions might be. That runs the risk of restrictions being used as a “criminal law for the enemy” against political opponents.
The proliferation of extreme and offensive comments and expressions could create a social climate that leads to outbreaks of violence.	No clear link has been demonstrated in established democracies between the proliferation of hate speech and an increase in hate crime. The State can already punish hate when it is a motive for criminal acts as an aggravating circumstance.
Hate speech causes direct psychological damage (feeling of being threatened, humiliated, etc.) and contributes towards perpetuating situations of discrimination.	There are already legal means for prohibiting harassment, threats and individual attacks on dignity, as well as more legitimate, proportionate and effective instruments for combating discrimination.

¹ Table based on *Nineteen arguments for hate speech bans - and against them* by Eric Heinze.

<p>The most extreme declarations of hatred incite acts of violence and discrimination, and should therefore be restricted and punished.</p>	<p>The State can already punish acts of provocation and conspiracy, where a clear link can be established with the crime. The notion of incitement is attributed by a punishment mechanism without any need to demonstrate the possibility of damage being caused.</p>
<p>The State cannot remain neutral and bans on hate speech symbolise its commitment to democratic values, equality and human dignity.</p>	<p>The State has more legitimate and more effective ways of positioning itself on a symbolic level. Bans have little effect and may even be counter-productive.</p>

2.2. Legal aspects of hate on the Internet and in social media

As if this political, ethical and legal debate did not already pose enough difficulties, the problem of hate speech is made even more difficult by the Internet and social media. First of all because the Internet has become a kind of freedom symbol, so any attempt to regulate what happens on the Net is automatically branded as reactionary.

But the Internet and social media pose a second problem as well, linked to the “border-free” character of the Net: the limit of jurisdiction.

The absence of borders also represents a challenge in areas such as international judicial cooperation, or communication privacy and data protection.

Furthermore, it poses problems that flow from the technical configuration and functionality of the Internet (server localisation, IP authentication, robot accounts, different encrypting procedures for hiding identity, etc.) that make it more difficult to obtain proof or determine responsibility.

2.3. Spanish legislation

The new regulation brings the various hate speech crimes together in Article 510 of the Criminal Code, and punishes them with the following penalties:

- 1- **1 to 4 years in prison** and a fine of 6 to 12 months,
 - a. Direct or indirect incitement to hatred, hostility, discrimination and violence, against groups or individuals for racist, anti-Semitic and other motives linked to ideology, religion, origin, sex, sexual orientation, illness or disability.
 - b. Production and distribution of material with the same motives as those mentioned in Section a).
 - c. Denial or glorification of the crime of genocide, crimes against humanity or against individuals or property in the case of armed conflict that may have been committed against the aforementioned groups, when that promotes or fosters a climate of violence, hostility and hatred towards them.

- 2- **6 months to 2 years in prison** and a fine of 6 to 12 months (and 1 to 4 years in prison when a climate of violence, hate and discrimination is promoted)
 - a. Holding events, producing and disseminating material that might imply humiliating or disparaging one of the aforementioned groups or its members on discriminatory grounds.
 - b. Glorification or public justification of crimes committed against the aforementioned groups.

Article 510 also establishes a series of aggravating circumstances and accessory consequences, notably the aggravating circumstance of having used the Internet to carry out any of the behaviours described.

Furthermore, the Criminal Code envisages a generic aggravating circumstance for discriminatory motives, which aggravates the criminal responsibility of any offence and, for example, can be applied to typical social media behaviour, such as threats or harassment. Threats against groups may also be considered an offence and this can be used to punish threats against ethnic, cultural and other groups that may be enough to strike fear among them.

The following are some of the main practical problems of prosecuting hate crime in general and hate speech in particular:

- Ignorance of the facts due to the very small number of complaints.
- The people in charge of investigations playing down the facts.
- Not enough police or judicial investigation.
- Difficulties in proving the “hate” or discriminatory motive of the crime.
- Lack of training on equality and non-discrimination for the institutions involved: the judiciary, judicial civil servants, forensic specialists, the police, private security, etc.
- Inherent difficulties of the “social media” scene: problems of jurisdiction, territoriality, identification and obtaining data on the aggressors, etc.

3. THE ROLE OF SOCIAL MEDIA AS INTERMEDIARIES

There is no doubt that the social media themselves, as a communication channel for “hate” messages, play an important role in the fight against hate speech. Their role as intermediaries in digital communication makes them the first arbiter in determining what can and cannot be said.

These intermediary companies propose various alternatives: attributing responsibility to the companies, promoting self-regulation and installing computer filters. Any of these options poses certain dilemmas.

3.1. Attributing responsibility

While service providers in the United States are virtually free of any responsibility for published content, the attribution of responsibility to intermediary service providers in the European Union is based largely on the fact that they have been made aware of the illicit nature of the content published through their service. In theory, they will only be

held responsible when they do not withdraw access to illegal content after having been made aware of its illegality.

3.2. Self-regulation and installing computer filters

As private operators, companies have the right to establish ethical codes and terms of use for their services, which are applicable around the world.

All the main social media make users accept terms of use that, in theory, ban hate speech from their platforms. Furthermore, they have reporting systems that allow users to report content that violates the standards for using social media, so the companies can assess whether to delete it.

Nevertheless, despite this attitude which, theoretically, does not tolerate hate speech, it is clear that the mechanisms for reporting and deleting comments do not work as they should. However sophisticated the algorithms on which the computer filters developed by social networks are based may be, they tend to commit block errors.

Either way, it is only fair to remember the enormous difficulty of managing this problem. Even with the help of computer filtering, the task of moderating and self-regulation is one of giant proportions. Twitter, for example, handles over 300 million tweets a day. Apart from the technical difficulties of checking such quantities of messages, content moderation is a very sensitive task for a company's image; they receive protests for not respecting freedom of expression and accusations of censorship as well as being a space of impunity for hate. The social media themselves are aware of the problem they have with managing online harassment, threats and hate speech, which has become their Achilles heel given the difficulty of striking a balance between freedom of expression and respect for personal dignity.

3.3. The Code of Conduct with the European Union

In the wake of the refugee crisis in 2015, and the obvious growth in xenophobic and racist statements on social media, the European Union pressurised the tech companies to assume a much more active role in the fight against hate speech.

As a result of this pressure, Facebook, Twitter, YouTube and Microsoft signed a Code of Conduct in May 2016 on illegal incitement to hatred on the Internet. The companies made a series of commitments in this agreement, notably to:

- check requests for the withdrawal of illegal content inciting hatred within a period of 24 hours;
- establish clear procedures for examining reported content and to assess the reports they receive on their services in compliance with their self-regulation rules, but also taking national legislation into account, whenever necessary;
- establish cooperation between the signatory companies and exchange good practices with other online social communication services;
- promote initiatives that develop an alternative "counter-speech";
- support educational programmes that foster critical thinking, and collaborate with civil society organisations in training activities.

4. EXTRAJUDICIAL STRATEGIES FOR COMABTING HATE SPEECH IN SOCIAL MEDIA

Below, without any claim to being exhaustive, we outline the main “non-judicial” strategies being pursued in response to hate speech in general and in social media in particular.

4.1. Monitoring and research strategies

These kinds of strategies are important for a comprehensive understanding of the phenomenon of hate speech in social media. This knowledge is valuable in itself but, above all, when it comes to applying it to the search for resources to mitigate the problem, or which might anticipate its most damaging consequences. For example, monitoring can do an important job as an early warning system in unstable situations, to forewarn of violent events.

One of the reference centres in the field of monitoring and research, due to the consistency of its studies, is the Centre for the Analysis of Social Media at the British think tank Demos, which carries out frequent studies, not only on the spread of different types of hate speech (Islamophobia, sexism, etc.), but also on how the counter-narratives challenging hatred in the media work. Facebook has commissioned Demos to carry out a series of studies on content which combats hate speech, to assess which is the most successful and which can be counter-productive.

4.2. Strategies for pressurising social media

These types of strategies, which aim to ensure that social media companies do not allow their platforms to be used to publish extremely intolerant content, can focus pressure directly on the companies themselves, or indirectly by targeting their advertisers.

Civil society organisations pressurise the companies directly by means of online campaigns, collecting complaints, online petitions, etc., with uneven results. Public institutions have more capacity for putting pressure on these companies and have succeeded in getting some commitments from them, such as the Code of Conduct promoted by the European Union, mentioned above.

Pressure campaigns on advertisers have, in turn, achieved considerable success. One example of this kind of initiative, carried out jointly in Great Britain and the United States by WAM! (Women, Action and the Media Group) and the Everyday Sexism Project, to demand the removal of content that is abusive towards women, succeeded in getting 15 large companies to withdraw their advertising from Facebook and getting the company to publicly commit to reviewing its terms of service and its mechanisms for checking content.

4.3. Strategies designed to change perceptions and attitudes

4.3.1 *Campaigns against online hate speech*

By looking at their content and specific orientation, these campaigns can be classified

into three types: awareness, affirmative and restrictive². Before looking at them, and because of its centrality and scope, we will examine *Viviendo juntos online: acción y campaña contra el discurso del odio* (No Hate campaign), launched by the Council of Europe (COE) in March 2013.

a) *The Council of Europe's No Hate campaign*

The No Hate campaign is part of the COE's effort to promote human rights on the Internet. The main goal of this campaign is to mobilise Europe's young people, creating a social movement (No Hate Speech Movement) that can help to reduce the level of acceptance of this type of speech.

This campaign has been reproduced on a national level by means of National Campaign Committees. These committees bring together the main interested parties in each country and follow the directives issued by the Council of Europe to maintain the general coherence of the campaign. So far, more than 40 committees have been set up, both in member states and in other countries, such as the USA, Mexico and Morocco.

This campaign has also been carried out at a local level, through a series of activities held in a number of European cities. In addition, Strasbourg has actively collaborated with the COE, which is based there. However, up to now there has not been a specific movement or network of cities that explicitly supports the campaign.

b) *A type of campaign*

Besides the COE's campaign, other campaigns of a more restricted scope have also been run across Europe. With regard to their content and specific orientation, they can be classified into three types.

- **Awareness campaigns**, which aim to make the general public aware of hate speech and its consequences. One campaign worthy of mention in Spain is "Internet with No Risks" (*Internet sin riesgos*), launched by the Tenerife Island Council Youth Area with the aim of raising awareness of the propagation of online hate and sexual harassment.
- **Affirmative campaigns**, which aim to present minorities to the general public in a positive light in order to prevent discriminatory behaviour. Some examples of this are the British "Islam is Peace campaign", which focuses on breaking down the prejudices against Islam and its practitioners, and the Spanish campaign *#YoNoSoyTrapacero*, which criticises the negative image of Gypsies spread by the Real Academia Española (RAE).
- **Restrictive campaigns**, which aim to gather information on pages and online actions based on intolerant content and to take action to restrict such activity. An example of this is "Get the Trolls Out", an international campaign launched by various organisations to restrict online anti-Semitic hate speech.

² Titley, Gavan, Ellie Keen, and László Földi. 2014. "Starting Points for Combating Hate Speech Online. Three Studies about Online Hate Speech and Ways to Address It."

4.3.2 Strategies against online hate speech based on education and training

Another group of strategies for combating online hate is based on education and training. Without claiming to be exhaustive, these initiatives can be classified into two large groups: on the one hand, those geared towards educating and training the general public, to develop what has been conceptualised as “digital citizenship”; and, on the other hand, initiatives aimed at equipping people who are already aware of the problem and activists with the skills needed to act more effectively.

a) Critical thinking and reflection as key tools against online hate speech

Given the centrality that the Internet and social media have acquired in recent years as a form of communication, it has become clear that educational strategies must include fostering skills such as the critical reception of media messages and empowerment in creating media content. To bring these new skills about, the notion of “digital citizenship” has been put forward, which implies adapting the concept of education for citizenship to incorporate the knowledge and skills required for interacting in a digital medium.

Some initiatives against hate speech have focused on developing these skills by means of training courses and developing specific educational tools. Examples of this are the “Prevention of Online Hate Speech” training project, organised by the European Wergeland Centre in Norway, or “No Hate - Something that concerns you” (*NO HATE - Ce qui vous regarde*), an educational project organised by the Belgian association Loupiote, during which a series of workshops against cyber hate were held with young people that included developing educational tools, available online.

b) Training cyber activists: some specific strategies for training groups already aware of hate speech

Initiatives have also been carried out with the specific aim of empowering people already aware of, or victims of hate speech with specific skills that will enable them to be more effective online. In fact, empowering activists is the main goal of the Young People Combating Hate Speech Online project, which is part of the No Hate campaign and is intended to give young people and youth associations the necessary skills for recognising and acting against hate speech, as well as involving them in launching the campaign. Among initiatives following the same philosophy in Spain, it is worth highlighting the workshop “How to Act Against Online Hate Speech”, carried out as part of the Online Project Against Xenophobia and Intolerance in Digital Media (PROXI), launched by the Human Rights Institute of Catalonia and United Explanations.

BIBLIOGRAPHY

Aguilar García, Miguel Ángel, ed. 2015. *Manual Práctico Para La Investigación Y Enjuiciamiento de Delitos de Odio y Discriminación*. Barcelona: Generalitat de Catalunya Centre d'Estudis Jurídics i Formació Especialitzada.

Article 19. 2015. *'Hate Speech' Explained A Toolkit*. Article 19. London. Available at:

- Bustos Gisbert, Rafael. 2015. "Libertad de expresión y discurso negacionista", in *Libertad de Expresión y discursos del odio*, Miguel Revenga Sánchez (ed.), *Cuadernos de la Cátedra de Democracia y Derechos Humanos* N°12. Universidad de Alcalá.
- Castells, Manuel. 2012. *Redes de Indignación y Esperanza*: Los Movimientos Sociales En La Era de Internet. Alianza Editorial.
- Gagliardone, Iginio, Alisha Patel, and Matti Pohjonen. 2014. "Mapping and Analysing Hate Speech Online: Opportunities and Challenges for Ethiopia". University of Oxford and Addis Ababa University.
<http://pcmlp.socleg.ox.ac.uk/sites/pcmlp.socleg.ox.ac.uk/files/Ethiopia%20hate%20speech.pdf>
- Gagliardone, Iginio, Danit Gal, Thiago Alves, and Gabriela Martinez. 2015. Countering Online Hate Speech. Programme in Comparative Media Law and Policy, University of Oxford.
- Gascón Cuenca, Andrés. 2015. "La Nueva Regulación Del Discurso Del Odio En El Ordenamiento Jurídico Español: La Modificación Del Artículo 510 CP. *Cuadernos Electrónicos de Filosofía Del Derecho*, N° 32.
- Heinze, Eric. 2014. Nineteen arguments for hate speech bans - and against them. Available at: <http://freespeechdebate.com/en/discuss/nineteen-arguments-for-hate-speech-bans-and-against-them/> (Consulted 23/11/2016)
- Hoehsmann, Michael, and Stuart R. Poyntz. 2012. *Media Literacies*: A Critical Introduction. Wiley-Blackwell.
- Jubany, Olga, and Malin Roiha. 2016. "Backgrounds, Experiences and Responses to Online Hate Speech": A Comparative Cross-Country Analysis."
- Jubany, Olga, Roiha, Malin, and Martínez, Arlette. 2016. "Online Hate Speech in Spain. Fieldwork Report. PRISM. Preventing, Redressing & Inhibiting Hate Speech in New Media". Universitat de Barcelona. Fundamental Rights and Citizenship Programme of the European Union
- Kahn, Robert. 2015. Rethinking the Context of Hate Speech, in *First Amendment Law Review*; Univ. of St. Thomas (Minnesota) *Legal Studies Research Paper* N° 15-20.
- Keen, Ellie, and Mara Georgescu. 2016. *Bookmarks - A Manual For Combating Hate Speech Online Through Human Rights Education*. Strasbourg: Council of Europe.
<http://site.ebrary.com/lib/uacj/docDetail.action?docID=10961376>.
- Legault, Lisa, Jennifer N. Gutsell, and Michael Inzlicht. 2011. "Ironic Effects of Antiprejudice Messages: How Motivational Interventions Can Reduce (but Also Increase) Prejudice". *Psychological Science*, 22, 1472-1477. DOI: 10.1177/0956797611427918
- Spanish Ministry of Interior
- 2015. Protocolo de actuación de las fuerzas y cuerpos de seguridad para los delitos de odio y conductas que vulneran las normas legales sobre discriminación. Available at:
 - 2016. "Informe 2015 Sobre Incidentes Relacionados Con Los Delitos de Odio En España". Available at: <http://datos.gob.es/catalogo/informe-2015-sobre-incidentes-relacionados-con-delitos-de-odio-espana>.
- Moretón Toquero, María Aranzazu. 2012. "El «ciberodio», La Nueva Cara Del Mensaje de Odio: Entre La Cibercriminalidad Y La Libertad de Expresión." *Revista Jurídica de Castilla Y León*, N° 27.
- Morozov, Evgeny. 2012. *El Desengaño de Internet: Los Mitos de La Libertad En La Red. Imago Mundi*. Barcelona: Ediciones Destino.
- Phillipson, Gavin. 2015. "Hate Speech Laws: What they should and shouldn't try to do", *Revue générale du droit* (www.revuegeneraldroit.eu), Etudes et reflexions 2015, N° 13.
- Revenga Sánchez, Miguel. 2015. "Los discursos del odio y la democracia adjetivada: tolerante, intransigente, ¿militante?", in *Libertad de Expresión y discursos del odio (op.cit)*.

- Rey Martínez, Fernando. 2015. "Discurso del odio y racismo líquido", in *Libertad de Expresión y discursos del odio (op. cit.)*.
- Rodríguez Izquierdo, Myriam. 2015 "El discurso del odio a través de Internet", en *Libertad de Expresión y discursos del odio (op. cit.)*.
- Ruiz, Carlos, Pere Masip, Josep Lluís Micó, Javier Díaz-Noci, and David Domingo 2010 "Convesación 2.0 y democracia Análisis de los comentarios de los lectores en la prensa digital catalana". *Comunicación y Sociedad*. Vol. XXIII, Núm. 2, 2010, pp.7-39
- Silverman, Tanya, Christopher J Stewart, Jonathan Birdwell, and Zahed Amanullah. 2016. "The Impact of Counter-Narratives. Insights from a Year-Long Cross-Platform Pilot Study of Counter-Narrative Curation, Targeting, Evaluation and Impact."
- Teruel Lozano, Germán M. 2015. "La libertad de expresión frente a los delitos de negacionismo y de provocación al odio y a la violencia: Sombras Sin Luces En La Reforma Del Código Penal." *Indret: Revista Para El Análisis Del Derecho*, N° 4.
- Titely, Gavan, Ellie Keen, and László Földi. 2014. "Starting Points for Combating Hate Speech Online. Three Studies about Online Hate Speech and Ways to Address It."
- Van Spanje, J., and C. de Vreese. 2015. "The Good, the Bad and the Voter: The Impact of Hate Speech Prosecution of a Politician on Electoral Support for His Party." *Party Politics* 21 (1). SAGE Publications: 115 -30 256 DOI:10.1177 1354068812472553
- Vázquez Alonso, Víctor J. 2015. "Libertad de expresión y religión en la cultura liberal: de la moralidad cristiana al miedo postsecular", in *Libertad de Expresión y discursos del odio (op. cit.)*.
- Vives Antón, Tomás. 2015. "Sobre la apología del terrorismo como 'discurso' del odio", en *Libertad de Expresión y discursos del odio (op. cit.)*.