



Selected Papers of #AoIR2017:
The 18th Annual Conference of the
Association of Internet Researchers
Tartu, Estonia / 18-21 October 2017

RACISM AND AFFECT IN THE HYBRID MEDIA SYSTEM: ANALYZING THE 'FINNISH IMMIGRATION DEBATE'

Gavan Titley
Maynooth University/University of Helsinki

Kaarina Nikunen
Tampere University

Mervi Pantti
Helsinki University

The Syrian emergency and wider border crisis has amplified exclusionary practices and political discourses in Europe. Though neither unique to the post-2014 moment nor a 'temporary problem' of European politics, the intensive construction of refugees and people who migrate as a near-existential crisis for the European Union has led to the mainstream legitimation of overt Islamophobia and xenophobia. The extensive media coverage and political instrumentalisation has ensured that in public debates, refugees and migrants have been connected with terrorism - particularly after the Paris attacks in October 2015 - and with sexual violence, most notably after the allegedly organised sexual harassment of women in Cologne on New Year Eve 2015. It is certainly the case that mediated moments of individual tragedy, such as the widely circulated photo of Alan Kurdi, have generated social media events that appeared to shift public opinion towards the plight of refugees rather than the 'problem of migrants' (see Vis et al 2015). However the intersection of these divisive communicative events, and their circulation throughout blogs, social media platforms and mainstream news sites, with the overtness of anti-immigrant movements and 'populist' politics in European politics, have arguably fuelled opposition and racism towards refugees and migrants.

In this broader context, this paper explores the 'immigration debate' in Finland with a particular focus on the *affective force* of racism in social media networks, and on how key communicative events become generative sites for the discursive reproduction – and to resistance to - racism. In the Finnish context, several significant online

Titley, G., Nikunen, K., Pantti, M. (2017, October 18-21). *Racism And Affect In The Hybrid Media System: Analyzing The 'Finnish Immigration Debate'*. Paper presented at AoIR 2017: The 18th Annual Conference of the Association of Internet Researchers. Tartu, Estonia: AoIR. Retrieved from <http://spir.aoir.org>.

communities and blogs played a crucial role in mobilizing racist rhetoric and sentiments during this period, but it is not possible to understand their significance without examining them in terms of what Andrew Chadwick theorises as a hybrid media system (2013), a concept that emphasizes the interdependence of older and newer media logics, and thus invites an exploration of the relations between different communicative and political spaces. Central to Chadwick's analysis is the recognition that a hybrid media system will inevitably be characterized by contingent distributions of power. That is, even within an unequal political economy of communication, the dynamics of intensive communicative events that are simultaneously mediated across platforms allow for a wider range of actors to "create, tap or steer information flows in ways that suit their goals in ways that modify, enable, or disable others' agency"(Chadwick 2013: 218).

The paper argues that these dynamics are critical to understanding how contemporary expressions of racism are shaped and contested in the interactive everyday cultures of digital media. As the sociology of racism broadly recognizes, racism appears to occupy a paradoxical status in contemporary societies (Murji & Solomos 2014). It is held to have been historically overcome, while remaining a constant focus of public debate and political mobilization (Lentin & Titley 2011). The 'refugee and migration' crisis has witnessed a distinct intensification of this dynamic, with endless contestation as to what constitutes racism in public debates that foreground 'the migrant' and 'the refugee' as problems to be deliberated upon, and solved. Thus, we argue, racism are discursively shaped in a context that combines both the political disputability of racism in contemporary societies of migration and multiculturalism, and in an interactive, interdependent and convergent media environment predicated on the 'spreadability' of media content (Jenkins et al 2014) and the 'constant incitement to discourse' (Couldry 2012) of media events and digital participation.

To explore the intensity of these 'events about racism', we turn to a consideration of how affect drives online engagements in different ways (Paasonen 2016). The 'affective dimension' of racism is, of course, not new, having been explored in a range of studies of everyday experiences of being racialized, and investments in racism (Essed 1991, Jackson 2010). We approach affect as practice to be able to explore the ways in which social formations are marked and shaped by affect and emotions (Wetherell 2012, 103). In this way our work involves exploration of the collective dimensions of affective meaning-making. We focus on affective-discursive practices that involve sharing and circulation (Ahmed 2007) of racism in the Finnish immigration debate through three analytical dimensions.

First, affective practice engages emotional discourses shared through personal networks. The topics circulated also draw on intimacy and body – and employ themes such as rape and sexuality - that are used to politicize the intimate sphere.

Second, we analyse the production of racism through an ironic sensibility that provides multiple layers of meanings to images and an increasingly challenging environment for interpretations. The practices of ironic commenting, re-interpretations, and intentional misrepresentations complicate the critique with ever escaping definitions and in this way epitomize the sensibility of the post-factual context (Andrejevic 2013: 10-13).

Third, we look at the structure of these rhetorical practices. The intensity of the interaction and the way publics organize around particular debates produce what has been termed as networked and affective publics (Papacharissi 2013). This emergence of publics around news events, debates and topics is also structured by algorithmic mechanisms that constitute several subcultures or multiple parallel and also racialized 'bubbles' (Gillespie 2014, Pariser, 2011; Sharma 2013).

The data of the research is gathered from Finnish online discussion forums, social media and mainstream news platforms using custom web scrapers and existing application programming interfaces as well as from existing big data sets of Finnish online discussion collected during the Finnish Parliament Elections in 2015 for the project "Cyber-elections 2015" and online discussion forum, "Suomi24", from 2002 to this date publicly available through FIN-Clarín. Acknowledging the limitations related to automated text mining and computational methods (e.g. Grimmer & Stewart 2013), the research has combined quantitative data with qualitative data collection and analysis with focus on specific case studies.

References:

- Ahmed, S. (2004) *The Cultural Politics of Emotion*. Edinburgh University Press.
- Andrejevic, M. 2013. *Infoglut: How Too Much Information is Changing the Way We Think and Know*. London and New York: Routledge.
- Chadwick, A. (2013) *The Hybrid Media System: Politics and Power*. Oxford University Press.
- Couldry, N. (2012) *Media, Society, World. Social Theory and Digital Media Practice*. Polity Press.
- Essed, P. (1991) *Understanding Everyday Racism*. London; Sage.
- Gillespie, T. (2014) *The Relevance of Algorithms*. In T. Gillespie, P.J. Boczkowski, K. A. Foot: *Media Technologies: Essays on Communication, Materiality and Society*. Cambridge: MIT Press.
- Hills, B, Paasonen, S. & Petit, M. (2015) (eds.) *Networked Affect*. Cambridge: MIT Press.
- Jackson, J.L., (2010) *Racial Paranoia: The Unintended Consequences of Political Correctness*. New York: Civitas Press.
- Jenkins, H., Ford, S., and Green, J. (2013) *Spreadable Media*. New York University Press.
- Lentin, A. & G. Titley (2011) *The Crises of Multiculturalism: Racism in a Neoliberal Age*. London: Zed Books.
- Murji, K. & J. Solomos (2014) *Theories of Race and Ethnicity: Contemporary Debates and Perspectives*. Oxford: OUP.
- Paasonen, S. (2015) *A midsummer's bonfire: Affective intensities of online debate*. In K. Hillis, S. Paasonen and M. Petit (eds.) *Networked Affect*. Cambridge: MIT Press, 27-42.
- Papacharissi, Z. (2014) *Affective Publics: Sentiment, Technology, Politics*. Oxford University Press.
- Pariser, E. (2011) *The Filter Bubble: what the internet is hiding from you*. The Penguin Books.

Sharma, S. (2013). Black Twitter? Racial Hashtags, Networks and Contagion. *New Formations* 78(1), 46-64. Shohat, E. and Stam, R. (2012) *Race in Translation: Culture Wars Around the Postcolonial Atlantic*. New York
Wetherell, M. (2012) *Affect and emotion: A new social scientific understanding*. London: Sage.

THEORIZING ONLINE RACISM: THE STREAM, AFFECT AND POWER LAWS.

Sanjay Sharma
Brunel University

The Stream has been the organizing metaphor for the web for the past several years...The Stream represents the triumph of reverse-chronology, where importance...is based exclusively on newness...(Madigral 2013).

The Internet has reached peak hate (Marche 2013).

The Web has supposedly become a dynamic social *stream*. This stream is replete with viral events such as breaking news or political revolts, as well as the frenetic circulation of celebrity gossip, banal chatter, images, videos and status updates. Social media platforms such as Facebook, YouTube, Instagram, Twitter and Reddit are key players in generating the stream. They 'curate' and monetize the enormous amounts of user generated content. Nonetheless, according to editor of Techcrunch, the social stream '...creates a different form of syndication which cannot be licensed and cannot be controlled' (Schonfeld 2009).

Social media commentators lament that the hyper-connectivity of the Web has produced a mass participatory medium that foments over-sharing, produces 'fake news' and creates a fickle online attention-economy. In particular, the stream is cast as a debased flow of communication. Moreover, the mainstream condemnation of 'online hate' flits between blaming malicious users/mob-like crowds for propagating hostility; or singles out the technologies of the medium itself – for example, internet anonymity, ease of sharing – as being responsible for engendering a toxic online culture (Harry 2013).

The characterization of the social stream raging with hate does little to advance an understanding of what are specific forms of online antagonism (Shepard et al 2015). While a burgeoning academic literature is exploring various forms of online hate, there has been a relative lack of research specifically examining Web-based racism. In particular, *theorizing* how racism is manifested in social media remains remarkably under-developed (cf. Daniels 2012; Nakamura 2008). Along with other forms of online antagonism, racism occurs across different social media platforms, with varying degrees

of frequency, magnitude, visibility and affect. Online racism is propagated by individuals, networked crowds, bots and algorithms. And racism can morph as it traverses the Web, gaining traction and visibility across different platforms.

The *multi-modality* of Web racism is not easy to analyse. The lack of meta-theorizations of Web-based racism has thwarted establishing it as a field of study (cf. Kolko et al 2000). A key limitation to our understanding has been a failure to pursue an analysis from a *sociotechnical* standpoint. This paper, organised as three parts, innovates an approach for studying the manifestation of online racism by advancing a 'materialist' account.

Part I interrogates mainstream accounts of online racism. These treat racism as either: (i) an '*extreme*', socially aberrant real world phenomenon which spills over into the virtual world and is amplified by the Web; or conversely; (ii) '*everyday*' online racism is a prosaic phenomenon, primarily caused by the communicative technologies and network effects of the Web. Essentially, both perspectives fail to acknowledge the mutual imbrication and entanglement of racism *and* the Web.

Part II offers an alternative approach that conceives online racism as an *assemblage* (Sharma 2013). Racism is *emergent*, via interactions and encounters on the Web: an entanglement of both human and non-human relations and technologies. Racial differentiations are 'performed not only through human interaction, but also through encounters with all kinds of things...which have the capacity to affect' (Swanton 2010: 8).

Part III identifies a *racialized* 'affective economy' for grasping online racism. There is a wild variation in the attention economy of the Web, because the processes of information propagation and visibility are multiply determined. It is proposed that online racism can be modelled by a power law distribution. In situations when power laws operate, 'decisions' - such as linking to a web page, 'Liking', sharing or retweeting - are influenced by *existing* behaviours of the system. Racism is not an individualized, random or independent phenomenon; it is embedded in techno-socio-material relations.

The power law distribution figuratively describes not just scale, but also variations in frequency, magnitude, visibility and affect of online racialized events. With reference to the flow of affects in the assemblages of online racism, I am concerned with elaborating '... an "affect economy"' (Clough, cited in Fox & Alldred 2014: 403). It is possible to analytically identify at least three modalities of online racism with respect to a power law distribution: *spectacular*, *explicit* and *ambient* racisms.

(i) At the 'head' of the power law distribution is *spectacular racism*. This characterizes racialized events that garner significant attention on the Web, for example, by a trending hashtag, or an image or video intensely shared across social media platforms, or reported by influential online magazine sites and mainstream news media. It can be

formed of an event such as a prominent public figure shamed for a racist diatribe, or the exposure of racist trolling against a non-white celebrity.

(ii) In between the 'head' and 'long tail' of the power law distribution is *explicit* racism. This is less visible on the Web, as it tends to occur in particular sites that foment online racist cultures. It can circulate in a Facebook hate group, or fester in a racist sub-reddit, or take the form of a racist meme created by the notorious trolls of 4Chan.

(iii) Ambient racism languishes in the long tail of the power law distribution. While this mode of racism garners relatively little immediate attention on the Web, it constitutes the majority of online racialized events. Ambient racism appears as everyday antagonistic online chatter, and is composed of seemingly transient racialized infractions, such as vitriolic expressions on online news sites and YouTube's comment space, or tweets harbouring micro-aggressions (Sharma & Brooker 2016).

The three modalities of online racism outlined aim to heuristically identify its differing characteristics. In reality they overlap, and rhizomatically flow across the web.

To conclude, a power law model to grasp online racism is a simplification of what is occurring. Nonetheless, it offers a means of beginning to making sense of the multifarious types of online racism that exist on the Web.

References

- Daniels, J., 2013. Race and racism in Internet Studies: A review and critique. *New Media & Society* 15, 695–719.
- Fox, N.J. & Alldred, P., 2015. New materialist social inquiry: designs, methods and the research-assemblage. *International Journal of Social Research Methodology* 18, 399–414.
- Harry, S., 2014. Attacking the Stream, *Dissent Magazine*.
https://www.dissentmagazine.org/online_articles/attacking-the-stream.
- Kolko, B.E., Nakamura, L., Rodman, G.B., 2000. *Race in Cyberspace*. Routledge.
- Madrigal, A.C., 2013. 2013: The Year "the Stream" Crested. *The Atlantic*.
- Marche, S., 2013. There Are No Saints Online. *Esquire*. <http://www.esquire.com/no-saints-online-0513>.
- Nakamura, L., 2008. *Digitizing Race: Visual Cultures of the Internet*. University of Minnesota Press.
- Schonfeld, E., 2013. Jump Into The Stream. *Esquire*. <http://www.esquire.com/news-politics/a22310/no-saints-online-0513/>
- Sharma, S., 2013. Black Twitter? Racial Hashtags, Networks and Contagion. *New Formations*, 78, 46–64.
- Sharma, S., Brooker, P., 2016. #notracist: Exploring racism denial talk on Twitter, in, Daniels, J., Gregory, K., Cottom, T.M. (Eds.), *Digital Sociologies*. Policy Press, pp. 459–481.
- Shepherd, T., Harvey, A., Jordan, T., Srauy, S., Miltner, K., 2015. Histories of Hating. *Social Media + Society* 1.

Swanton, D., 2010. Sorting Bodies: Race, Affect, and Everyday Multiculture in a Mill Town in Northern England. *Environment and Planning A* 42, 2332–2350.

MORE THAN MEETS THE EYE: UNDERSTANDING NETWORKS OF IMAGES IN CONTROVERSIES AROUND RACISM ON SOCIAL MEDIA

Ariadna Matamoros Fernandez
Queensland University of Technology

Tim Highfield
Queensland University of Technology

Felix Muench
Queensland University of Technology

Introduction

Visual objects are central to social media and its practices, acting as key mediators of discussions of matters of socio-cultural relevance, including politics and news events. Iconic images spread virally across social media, becoming symbols of events and issues, and inspiring further visual reactions and remixes; these include the myriad versions of Shepard Fairey's 2008 Obama 'Hope' poster (Gries, 2015), or of the photograph of three year-old Syrian refugee Alan Kurdi (Vis & Goriunova, 2015). While most of the research on social media communication has focused on text-based analysis, the visual as "widespread social media form" requires critical investigation ([author]). This paper examines images as an opportunity to understand conversations around race and racism online (Nakamura, 2008), and develops new methods to understand processes of meaning-making and sharing patterns across platforms.

Previous research on racism online has examined racist text-based discourse (Hughey & Daniels, 2013), hashtag tropes and their networks (Sharma & Brooker, 2016) and how the visual can highlight antagonistic practices (Everett, 2013; Milner, 2013). Most of the work on how the visual mediates racist practices, though, has used semiotic analyses of small-scale samples of visual content to describe the surrounding practices and cultures. Our empirical approach builds upon these methods by adding a second element of analysis: the examination of networks of visual content across platforms. Do certain genres of images tend to stay in their original communities? Can we identify polarising pictures by their communication patterns? Through a combination of qualitative approaches and community detection algorithms, we will map the actors, objects and arguments involved around the sharing of key visual objects to discuss current controversies around racism.

Context

The techno-social construction of racism online, which is nation- and medium-specific, requires nuanced investigation and the development of new methods to understand its complexity ([author]). We propose an empirical approach to examine how visual objects mediate conversations around race, focusing on the Australian context. Similar to Klugman and Osmond's (2009) study of how a 1993 photograph of Indigenous footballer Nicky Winmar gesturing to his skin triggered a national discussion about racism, we centre our analysis on two contemporary iconic images involving Adam Goodes, another Indigenous footballer and one of the game's most fêted players.

During a game in 2013, Goodes pointed out to the umpire a girl in the crowd who had called him an "ape", and who was subsequently removed. While there was substantial support for Goodes for calling out racism in Australian sporting culture, rival supporters would also boo him every time he played. In 2015, during the annual Indigenous Round, a celebration of the role of Aboriginal and Torres Strait Islander people in Australian Rules football, Goodes celebrated a goal by performing a war dance, which included him mimicking the action of throwing a spear in the general direction of the crowd. This event served to reignite debate about race and racism in Australia and increased the booing campaign towards Goodes at matches and online.

These two events involving Adam Goodes invoked different tropes on social media, from the 'aggressive black man' to 'playing the racist card'. Like Winmar's gesture 20 years earlier, images of Goodes' actions (see Figure 1) became iconic: widely shared, spread, and referenced in responses to these events.



Figure 1. Adam Goodes' gestures in 2013 (left) and 2015 (right).

Method and discussion

Our mixed-methods approach uses Twitter as our seed data source since it is a rich repository of images that users post from other platforms (Thelwall et al., 2015). We used the Tracking Infrastructure for Social Media Analysis (TrISMA) for the data collection, which utilises the Twitter API to continually capture tweets of 2.8m Australian users (Bruns et al., 2016). We chose the TrISMA dataset to guarantee that the images examined were posted by Australian users, which is important to understand the national specificity of racism. We queried for all tweets that matched the keyword "Goodes" between the war dance (29 May 2015) and Goodes' retirement on 16 September 2015. Although this period was focused on the war dance controversy, a

preliminary exploration of the images posted showed that images of Goodes' 2013 gesture were also shared. These were used to deny the racial motives of the booing and to justify opponents' dislike of him because of his involvement in past controversial events.

We qualitatively analysed all the tweets with images (2,174 tweets) to only keep the ones containing Goodes' iconic gestures from 2013 and 2015, ending with 200 tweets. Following this, we coded the images, using the tweet text to help determine whether they are 'pro' or 'against' Goodes¹. This categorization is crucial to study the publics and sharing dynamics around the same images based on different cultures of use: prosocial and antisocial practices.

The retweet, @mention, and co-hashtag networks around those images further develop this analysis. Network visualisations, supporting an explorative analysis of the dissemination of the pictures, as well as key network metrics, and community detection algorithms give us the possibility to find correlations between the origin and features of the content, communication patterns, and dissemination dynamics (e.g. del Val, Rebollo, & Botti, 2015). Furthermore, we can identify different actors and arguments involved in the debate, and the communication dynamics of these images, depending on who shared them and with what purpose. For instance, was the war dance image mostly shared on Twitter to oppose Goodes or to support him? How disconnected are the friends and foes of Adam Goodes?

This paper is a preliminary exploration of our method on Twitter. Building on this, we aim to explore how the visual was used and shared not only on Twitter, but also across other relevant platforms including YouTube and Facebook.

References

- Bruns, A., Burgess, J., Banks, J., Tjondronegoro, D., Dreiling, A., ... Hartley, J. (2016). *TrISMA: Tracking Infrastructure for Social Media Analysis*, <http://trisma.org/>
- del Val, E., Rebollo, M., & Botti, V. (2015). Does the type of event influence how user interactions evolve on Twitter? *PLOS ONE*, 10(5). doi:10.1371/journal.pone.0124049
- Everett, A. (2013). Have we become postracial yet? Race and media technology in the age of President Obama. In L. Nakamura & P. Chow-White (Eds.), *Race After the Internet* (pp. 146–167). New York: Routledge.
- Gries, L. E. (2015). *Still Life with Rhetoric: A new materialist approach for visual rhetorics*. Logan: Utah State University Press.

¹ Some mainstream media shared these images in a purely informative fashion. These were coded as 'neutral' and removed from the subsequent analysis.

Hughey, M. W., & Daniels, J. (2013). Racist comments at online news sites: a methodological dilemma for discourse analysis. *Media, Culture & Society*, 35(3), 332–347.

Klugman, M., & Osmond, G. (2009). That picture-Nicky Winmar and the history of an image. *Australian Aboriginal Studies*, 2009(2), 78.

Milner, R. M. (2013). FCJ-156 Hacking the Social: Internet Memes, Identity Antagonism, and the Logic of Lulz. *Fibreculture*, (22). Retrieved from <http://twentytwo.fibreculturejournal.org/fcj-156-hacking-the-social-internet-memes-identity-antagonism-and-the-logic-of-lulz/>

Nakamura, L. (2008). *Digitizing Race: Visual Cultures of the Internet*. Minneapolis: University of Minnesota Press.

Sharma, S., & Brooker, P. (2016). #notracist: Exploring racism denial talk on Twitter. Retrieved from http://opus.bath.ac.uk/53206/1/Sharma_Brooker_notracist_digital_sociology_preprint.pdf

Thelwall, M., Goriunova, O., Vis, F., Faulkner, S., Burns, A., Aulich, J., ... D'Orazio, F. (2015). Chatting through pictures? A classification of images tweeted in one week in the UK and USA. *Journal of the Association for Information Science and Technology*. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/asi.23620/full>

Vis, F., & Goriunova, O. (Eds.). (2015). *The Iconic Image on Social Media: A Rapid Research Response to the Death of Aylan Kurdi**. Retrieved from <http://visualsecialmedialab.org/projects/the-iconic-image-on-social-media>

ONLINE HATE SPEECH: GENEALOGIES, TENSIONS AND CONTENTIONS

Eugenia Siapera
Dublin City University

Paloma Viejo Otero
Dublin City University

Elena Moreo
Dublin City University

The explosion of discussions of online hate speech suggests raised public concerns. Controversial figures such as Milo Yiannopoulos have generated multiple discussions and commentaries in journalistic articles but also in people's timelines and news feeds. Concerns are raised over the proliferation of online hate and its pernicious effects especially on vulnerable individuals. At the same time concerns are raised over tensions between controlling hate and freedom of speech. Additionally, concerns are raised over the perceived duties and responsibilities of social media platforms. As more discussions, comments and opinions accumulate, there is added confusion and more and more layers of complexity.

On the one hand such discussions are typical of the domain of social media and of what Jodi Dean has described as communicative capitalism. On the other hand, the accumulation of discourse may end up obscuring the very contours of hate speech that it is meant to address and illuminate. Additionally, the very debatability of hate speech, especially racist hate speech, points to the operation of a particular politics in play which denies it during the very act of supposedly seeking to address it – as Gavan Titley (2016) has convincingly shown. Taking a step back, this article is looking to identify and discuss definitions of hate speech encountered in different political actors. In identifying and contextualising the origins, differences, tensions and contentions among these definitions, this article is hoping to contribute to understanding the terms of the discussion with a view to ultimately escaping the limits and impossible dilemmas encountered in much of the public debate. At the same time, this article is concerned with identifying the role played by the various actors in the debate.

The paper focuses on three discrete but occasionally overlapping discourses of online hate speech as an entry point to uncovering the terms of the debate. Firstly, it examines institutional and legal responses, which feed into policy; secondly, social media platforms themselves; using materials from interviews with key informants in Facebook and Twitter, and through an analysis of their online terms of service and reporting mechanisms; and finally, civil society initiatives that have emerged in order to deal with online hate and which act as pressure groups.

More specifically, the article begins by looking at the formal documents of the European Commission, and more specifically the Framework Decision of 2008 and the 2016 Code of Conduct agreed with social media platforms. Tracing the antecedents of these documents and their basis in other decisions referenced in this decision, the analysis

here identifies the roots of current online hate speech in the various international court decisions that took place at the end of WWII and which culminated in the Universal Declaration of Human Rights of 1948. However, this declaration subsumed important tensions between two ideologically opposite understandings, one associated with the Soviet delegation and one associated with US liberalism, revealed in the opposing principles in Article 7 concerning non-discrimination and Article 19 concerning freedom of speech (Viejo Otero, 2017). Addressing this tension in practice means that The EC Framework Decision has focused on what they identify as illegal online hate speech but leave everything else go through, while the Code of Conduct focuses on the efficiency of the mechanisms by which illegal hate speech is removed.

In many ways, existing legislation shapes understandings of hate speech in the policies of social media platforms. Both Facebook and Twitter contain references to the same vulnerable groups identified in the Universal Declaration of Human Rights. Facebook is influenced by the European Convention of Human Rights and has adopted its main principles. The two main sources of tension for these platforms concern firstly the liberal ideology and freedom of speech 'absolutism' and the requirement that they offer a pleasant and enjoyable experience to their customers; and secondly, between the social origins and uses of hate and the individualistic approach to users. The former tension is discursively resolved through references to safety and security for users; this is further accompanied by a realisation of the monumental difficulties of tracing online hate on their platforms a shift towards encouraging users to produce counter-narratives. The latter tension is addressed through the slide between hate speech and cyberbullying and through individualised reporting mechanisms (Siapera and Viejo Otero, 2016). In a third step, the paper will consider understandings of hate speech as they emerge from anti-racist groups in the context of Ireland. This part of the research hasn't been conducted yet, and it will rely on a set of focus group interviews with anti-racist groups discussing their understandings of hate speech online, their experiences and the varieties of hate encountered, as well as their views on how to address this. This analysis is looking to identify the tensions in the discourses and understandings of online hate encountered in civil society groups as well as tensions that emerge between online and offline experiences and because of the difficulties associated with remedial actions.

Taken together these three distinct but related discourses will reveal some of the contours of the debate on online hate speech and point to overlaps but also important differences. The discussion will point to a kind of role reversal in which social media platforms operate as states or state-like actors, and political institutions act as lobbyists. The role of civil society actors is diminished as social media platforms understand their constituencies primarily as individuals and not as members of social groups. At the same time, social media corporations collaborate with some civil society groups and NGOs but under conditions that remain unclear and using criteria that are opaque. The relationship remains one between a service provider and a client. In these terms, the formulation of distinct approaches to online hate speech points to the existence of a parallel politics that sees the various actors jostling for position and influence. In substantive terms, the prioritisation of legal discourses or illegal versus legal hate speech, of hate speech as an issue of safety and security, the individualisation of responses and emphasis on counter-narratives or the production of more speech, as

well as the lack of any open channels for social groups to communicate with social media platforms point to a foreclosure of political responses to online hate speech.

References

Dean, J., 2005. Communicative capitalism: Circulation and the foreclosure of politics. *Cultural Politics*, 1(1), pp.51-74.

Titley, G. ,2016, The debatability of racism. Networked participative media and postracialism. *Rasismista ja rajoista*, available at: <https://raster.fi/category/in-english/>
Siapera E. and Viejo Otero, P. 2016, Hate speech and social media platforms, paper presented at the 17th AioR conference, Berlin, 5-8 October.

Viejo Otero P., 2017, From harmful speech to hate speech, work in progress.