



## Podcast Transcript: *A More Social Media: Hate Speech in the Digital Public Sphere*

### **Narrator:**

Welcome to the Irish Human Rights and Equality Commission Podcast. In this podcast, we'll explore the issue of hate speech in the digital public sphere. You will hear contributions from speakers at an event held by the Commission at Dublin's Science Gallery in November 2018. The event brought together experts from the worlds of law, academic, media, journalism and tech to discuss online hate speech in the Irish context.

Eleanor Roosevelt, who was one of those involved in authoring the Universal Declaration of Human Rights, famously spoke about human rights starting in "small places, close to home". In 1948, Roosevelt could not have envisaged the digital revolution that we have witnessed, or the fact that hate speech online can now be automated and targeted at individuals.

Today, there can be no smaller places closer to home than the mobile phone screen you carry, or the laptop sitting on a desk in a closed bedroom. 97% of consumers in Ireland have access to either a smartphone or tablet and 98% of smartphone owners use their devices daily.



**Dr Sindy Joyce (Lecturer at University of Limerick):**

It's a great platform for the young people. In one way, it has built their confidence in different aspects; connecting with people from their own community who are in different locations and bringing those connections and that pride with the community. But then, the other side of it is the negative aspect of it, where you have those comments.

It affects their everyday life because particularly the young Travellers when they're on Twitter and Facebook and other social media platforms - every single day they're seeing negative comments about their community. It's so normalised that they just slide through it. It's rarely reported because if the community had to report everything that was on there, you would want to spend nearly 24 hours on the platforms reporting.

Comments saying that the community should be exterminated, calling the children feral, rodents, parasites. Talking about throwing us all into the sea, talking about burning caravans...

**Narrator:**

When we talk about the spread of racism and intolerance online, it's important to remind ourselves that we are not talking about an abstract idea. Toxic discourse, even when online, can have a real world, oppressive and damaging effect on those who fall victim to it. Hate speech can cause people to withdraw from society, to avoid expressing their identity. Hate speech can also carry a powerful message beyond those directly targeted, impacting an

entire community. When such expression becomes normalised in the public sphere, on the airwaves, and in the wider media, that impact is compounded.

**Dr Sindy Joyce:**

Most of the comments on Facebook are done by adults. Of course, you have a few kids negatively commenting, but mostly it's adults. So you have adults commenting on children in the community. That message of being inferior internalises within the community and brings on a lot of mental health issues because they're thinking about who they are. We're a proud community, but it's very difficult to be proud when you have negative comments continuously at your community, as if your community was some type of inferior group in society.

**Emma Dabiri (Broadcaster and Social Historian at the School of Oriental and African Studies, London):**

I went on 'Newsnight' and spoke about cultural appropriation. The backlash against that was absolutely wild. I've never experienced such an avalanche of hate. Also very extreme stuff, as well. I think because I'd written the piece about being Irish and not being white the week before, all the people who were coming for me for the cultural appropriation thing googled me and saw this piece and started reacting extremely violently to my claiming of Irishness. That's when all these American white supremacists came out of the woodwork and I was like 'Wow! This is a big thing online.' There was so much stuff written about me. There were long blog posts. People really did their research. Any information they could glean about me online was coalesced together; a lot of stuff about my mother and father. My dad is



Nigerian and my mum is white Irish, so there were a lot of those ideas about African men, black men, the types of Irish women who would have children with them and about how those offspring, me, could never be Irish. But this was not phrased in such a polite way...

**Narrator:**

One of the Commission's statutory functions is to encourage intercultural understanding. At its most basic, intercultural understanding involves us moving on step beyond the concept of the multicultural. It means acknowledging that a modern country like Ireland is host to a diversity of cultures; that cultural identity matters, and that disparaging, erasing or ignoring it can cause significant harm at individual and community level.

**Emma Dabiri:**

There was a picture going around Twitter, and people [said to me], "This is the kind of Ireland people like you want to create." It was just a picture of maybe eight black boys. They were standing next to a street sign that said 'Tyrrelstown'. I was like, "I see nothing wrong with this Ireland". The person was like, "These are gang members!" I was introduced to the African gang story which is not just online but is also in the mainstream media and very common amongst people I speak to, people that I would have imagined would know better. I was first introduced to it on Twitter. When you actually look at that picture that was spread far and wide and the assumption that these are gang members - it's not like they have weapons or anything. It's the combination of those things, that narrative of black criminality, of African criminality. That image seemed very potent and people just seemed to respond like "This is the worst thing that Ireland could become", but I was just like "Well



it has some black teenage boys - I don't understand where the terror is coming from from with this image." And that image wouldn't be regarded as representative of hate speech.

**Siobhan Cummiskey (Head of Content Policy at Facebook EMEA):**

My team is the Content Policy team at Facebook. We're the team that creates the Community Standards. We have a huge amount of people using our site - 1.5 billion people come to Facebook every day, and that number increases to 2.5 billion when you look at this figure on a monthly basis. There's a huge amount of content on Facebook and most people go there to share very positive things. What my team is responsible for is when that content is harmful, and we create the Community Standards as a result. My team is based in ten offices around the world. We hire people from a variety of different backgrounds. On my team alone, nobody shares the same nationality, they're all from different places in the world. They also come from many different working backgrounds, so we hire people from politics, we hire people with an NGO background. Our lead of the team globally is a former prosecution lawyer and as you heard earlier, I'm a former human rights lawyer. We hire a diverse group of people to work on the Community Standards but we don't develop them by ourselves, we develop them with a large number of people inside and outside of Facebook. We do that at a biweekly meeting that is called the Content Standards Forum. That's an internal meeting at Facebook where someone will come with an idea to change the policy. That idea could have come from an NGO that we work with, it could have come from what the content reviewers are seeing in their queues. They will come to this meeting and say "Heads up, I'm going to make an update to our hate speech policy or our suicide prevention policy - is anyone interesting in joining this working group?" So you'll have

people from Engineering and Legal on there, you'll have people from Public Policy and others. They will join this working group and that's where the real work happens and they create a number of options for improving the policy. That then goes to external third party groups, so nothing that goes through this Community Standards forum process goes through it without going to external third party experts. We've worked with a number of NGOs, child safety organisations and a number of academics and think tanks. Once they have approved a particular recommendation, it will go back to the group and that will become a live policy. When it becomes a live policy, we update our public-facing Community Standards and we also push that out to our Content Reviewers. We have thousands of Content Reviewers based all over the world and a large contingent here in Dublin. They will then start to operationalise that policy. To tell you a little bit about how Content Review works, we use a combination of technology and human review when we're looking at content. So automation is very useful to us in sorting the millions of reports that we receive. We receive millions of reports every week and automation will sort them to make sure that they go to the right person. If it's on a child safety issue then it goes to a child safety expert. If it's from a particular language and is a bullying or a hate speech issue, that will go to a native speaker. So we don't hire people who just happen to speak French as a second language to review French content; we hire native French people who can understand the context and nuance of language. Automation has been really useful to us and A.I. has been really useful to us. 99% of the terrorist content we delete before anyone ever reports it. When you look at hate speech, that number is 52%, so significantly lower and that is largely because hate speech can be very subtle and we rely on real people to review that content. We formally had 10,000 people working across Safety and Security at Facebook at the



beginning of this year; we now have 30,000 people working across Safety and Security across the company and about half of those are Content Reviewers who review reports 24/7. They have a follow-the-sun model. We get to the vast majority of reports within about 24 hours.

I'll finish up just talking about remedies. We have an appeals system in place so that if your content is removed, you can appeal that. We've had that in place for a long time around groups, pages and larger objects. We now have that at a content level for posts and videos for things like nudity and hate speech. We're also exploring the idea of having an independent appeals body that you can appeal to that would be separate from Facebook and if you're unhappy with our decision, you can appeal to them. That's at nascent stages and we hope to get to that within about a year. Finally, on the issue of transparency, we have a transparency report where some of those numbers that I've spoken to, you can find them there around how much content we remove. So in the first three months of this year, we removed 2.5 million pieces of hate speech from the platform. All of those figures are available there to view. Our Community Standards, which were made public last April, now contain a combination of our external-facing Community Standards but also what our internal reviewers actually use. You can see in there a very detailed definition of hate speech which includes ten protected characteristics which is quite a lot, including disability and disease, so we're ahead of some countries on that. Ethnicity is protected in there, and Travellers have always been recognised as an ethnicity by Facebook from long before I started working there in 2012. Our Content Standards Forum - we've had a number of external people come and view that meeting where I said we review the policies, including the UN Special Rapporteur on Freedom of Expression. We also get audited under the Global



Network Initiative which are the UN Guiding Principles translated for internet companies' use. And finally, just to mention the European Code of Conduct on hate speech, we are signed up to that and we removed 79% of reported hate speech content. 89% of those were removed within 24 hours. Those are just some of the initiatives that we are doing.

**Narrator:**

Recent international developments represent an increasingly urgent challenge for our society and democracy. We have witnessed in the past few years a worrying trend which has seen many countries in Europe and further afield succumb to the incipient rise of populism, unilateralism, racism and an increasingly narrow and inward-looking vision of statehood, sovereignty and national belonging. This trend and the threat it poses was described in stark terms by the outgoing High Commissioner for Human Rights, Zeid Ra'ad al-Hussein.

**Zeid Ra'ad al-Hussein:**

This is the way that wars are made: with the snarl of belligerence and the smirk of dehumanization; the lash of injustice and the incremental erosion of old and seemingly wearisome checks. The path of violence is made up of the unreckoned consequences of banal, incidental brutality seeping into the political landscape.

Here is one lesson: Intolerance is an insatiable machine. Its wheels, once they begin to function at a certain amplitude, become uncontrollable — grinding deeper, more cruelly and widely. First one group of people is singled out for hatred; next it will be more, and then more, as the machine for exclusion accelerates into violence, and into civil or international





warfare — feeding always on its own rage, a growing frenzy of grievance and blaming. As that tension begins to peak, no obvious mechanism exists that is capable of decompressing and controlling its intensity, because the machine functions on an emotional level that has very little contact with reason. Release may come only after tremendous violence. This is something those of us who work for human rights have witnessed time and again.

**Narrator:**

The role of the digital public sphere in this trend has been unmistakable. The potential for intolerance online to shape the public debate and the resulting political debate offline is becoming one of the hallmarks of the digital age.

**Dr Eugenia Siapera (Associate Professor, School of Communications at Dublin City University):**

The ‘trigger events’ that we identified, high profile events, were the case of Ibrahim Halawa, that triggered all these discussions about who qualifies to describe themselves as Irish. Also, events or posts or anything that has to do with housing and welfare - and here we can see the effects of austerity - would trigger racially toxic speech. Any news articles on Travellers and the Roma communities, or Muslims, refugees and immigration, these would be typical triggers, and terrorism as well.

**Dr Sindy Joyce:**

Not only was the hate speech on social media throughout those two to three weeks - as a community, we were bombarded by mainstream media as well. Newspaper articles, radio, television - we were bombarded with negative comments, as well, in our everyday life. So I



think that it was a triple - we had the social media, the mainstream media and the physical world itself coming down on top of us as a community, bombarding us with these negative comments and ideas. And as I said earlier, because anti-Traveller racism is so embedded in society, it's seen that it's not hate speech, it's seen that it's not anti-Traveller racism, it's seen that it's okay, that they're just criticising the community for not being settled.

**Narrator:**

This dynamic represents a significant area of threat to human rights and equality. As Ireland's national human rights and equality institution, the Commission believes that it is crucial to better understand online hate speech, and to see what can be done to challenge it.

There has been much discussion in academic studies about the shortcomings of Ireland's legislative infrastructure where hate speech is concerned. These criticisms are shared by international human rights treaty monitoring bodies. There is no doubt that our existing legislation, the *Prohibition of Incitement to Hatred Act* from 1989 is not fit for purpose and should be modernised. Strong law acts as an important symbol, sending a message of intolerance of intolerance.

**Dr Tarlach McGonagle (Senior Researcher at the Institute for Information Law, IViR, University of Amsterdam):**

In calibrating appropriate responses to counter hate speech it's important to look at the proportionality of those responses in terms of freedom of expression and other rights. But



also their effectiveness, that they actually serve their intended purposes for the victims and targets of hate speech.

**Dr Sindy Joyce:**

If you look at the physical world, we're regulated. You can just go out on the street there with your racist speech without something happening. Without you being called out on it and reported to the police station in the physical world, but it doesn't happen online. So it's like, what's the difference in freedom of expression if we're on the digital platform or if we're in the physical world itself.

**Siobhan Cumiskey:**

We recently announced we're working with the French government a cooperative effort around regulation. We know that regulation is coming, we know that it's something that people want and that is potentially very good for the community of people who use Facebook. So what we want to do is make sure that that's smart regulation, so what we're going to do is find out what the regulators want, what is it that they're interested in, what are their needs, and also to show them a little bit about what we currently do to find out if that's sufficient or what they think is lacking. I think hate speech is the most interesting area in which to do that and that's where we're starting because there is no commonly agreed definition of hate speech. If you take the European Code of Conduct, we're complying with that, but even within Europe, there's no agreed definition of hate speech. So it falls to us then to define it, because we don't want to allow hate speech on our platform. So we do



that carefully with experts and academics to ensure that we have a rule that works for our site. But regulation definitely has a role to play here as well.

**Narrator:**

While it's true that we can legislate away intolerance, a focus on absolutes can close down a more fruitful and nuanced debate about culture and norms. Cultural change is possible, and new new norms can be established, particularly by those with power and influence.

**Dr Tarlach McGonagle:**

I think it's very important that besides the regulatory approaches to also look at the power of information, education and awareness-raising. Very often the results achieved by those strategies are not instantly tangible, they're more long term solutions but they're absolutely crucial if you're trying to create a shared space in which people can express opposing viewpoints but in a way that is respectful of each other. I think that's a key for a pluralistic society.

**Dr Sindy Joyce:**

I think that the social media platforms' employees need to be trained, because the majority of comments that I report on platforms, I get messages back saying that it doesn't go against their community standards. It seems that the people working on these platforms just don't understand what hate speech is. They don't understand anything about the community, about minorities and particularly my community. When I talk to international people, the majority of them wouldn't even hear of my community. We've kind of been neglected not



only in history but in policy as well and that feeds from the top down the the bottom to the everyday person that perhaps has grown up with this embedded anti-Traveller racism, that it's okay to speak like that about a community. So I think the people working in those areas need to be trained in anti-Traveller racism particularly in Ireland. They also need to be aware of the different types of racial slurs that there is, that need to be taken down, and to be aware not only of direct hate speech comments but that a lot of people on social media platforms have gotten around direct hate speech and have comments that are insinuating different types of hostility, but it can't be directly seen as hate speech. So, for example, you might have a newspaper article about my community about accommodation or education. For that, then, to feed into social media, what's happening on the ground, like we had a lot of our young people going into schools, listening to how the community should be forced into settlement. I think that if people who worked on these platforms knew more about the community they would be able to pick it up and they would be able to see these indirect comments as well as the direct comments.

**Narrator:**

The traditional media has a crucial role to play. The way in which the media reports, engages with or denies racism and other prejudicial attitudes, matters a great deal. The media can serve to raise awareness of abhorrent views, to challenge them, and to engage with them critically. However, the media can also serve to amplify such views, providing them with a platform that is disproportionate and which serves to lend them legitimacy.



**Dr Gavan Titley (Senior Lecturer, Media Studies at Maynooth University):**

This entire responsive media ecosystem, which is desperate for content, which is competing for attention - what that results in is not just hate speech on Twitter about Travellers, but it results in two or three days where Travellers become the population which is to be debated, adjudicated about, commented on, and central to the discursive economy. And I don't know how we're going to tackle those kind of dynamics, but I know that what this report does, by asking us to think about the production of hate speech, not just as an act in and of itself, but as something that has a corporate logic, as something that has a platform logic, and as something that has a media system logic to it, that if we can start to put those pieces together, then we can at least map out the challenges that we undoubtedly face.

**Narrator:**

To meet these challenges will require a society-wide focus; a collective investment in challenging the underlying ideologies which underpin prejudice. A fostering of a more civil public sphere - both on and offline.

Thank you for listening to the Irish Human Rights and Equality Commission podcast. For more information on human rights and equality, follow the Commission on Twitter, @\_IHREC, or visit [www.ihrec.ie](http://www.ihrec.ie).

**[ENDS]**