DEMOCRACY AND SOCIETY 16.03.2017 | Jennifer Baker

Fakebook

A refugee's selfie made him the target of racists. His failed court case shows the justice system needs to catch up.

Last week a German court rejected efforts by teenage refugee, Anas Modamani, to force Facebook to protect him from trolls, harassment and hate speech.

The 19-year-old who fled Syria in 2015, became briefly famous for taking a selfie with German Chancellor, Angela Merkel, as she visited a refugee shelter in September of that year.

However last year events took a darker turn as Modamani's (often doctored) image was circulated on Facebook alongside accusations that he was responsible for the Brussels terrorist bombings, the Berlin Christmas market lorry attack, and even the murder of a homeless man.

Despite repeated reports to Facebook that these "news stories" were fake, Modamani says that many of the images are still circulating on the social network. As a poster boy for Germany's refugee policy, Modamani became a target. Whenever there were allegations that refugees were involved in criminal activity, his image was wheeled out and altered to suit an anti-immigration agenda.

There are two distinct problems with this activity: the harassment of the individual, and the manipulation of "news" to set an agenda.

To cope with his personal situation, Modamani sought legal representation and filed an injunction against Facebook in the southern German city of Würzburg.

On 7th March, the court rejected that injunction, agreeing with Facebook's claim that it was nearly impossible to track and shut down every single iteration of a post that has gone viral. The judge said Facebook was not obliged to "proactively seek out and delete defamatory posts" since it was "neither a perpetrator or participant in the smears".

Privatised policing versus legislation

While many will have some sympathy with Facebook's argument that there is no "miracle software" that can protect individuals like Modamani, the tech giant has shown itself to be more than capable of taking down thousands of images that breach other parts of its "Community Standards" and has successfully campaigned against legislation forcing it to act responsibly.

Facebook said it would "continue to respond quickly" to reports from Modamani or his legal team, and that it continually updates its Community Standards in an effort to eliminate inappropriate posts and comments. However, Facebook's Community Standards are not law and the company is wilfully opaque in how it handles takedown requests.

Facebook users may "report" posts by clicking on a button and selecting a reason the content is inappropriate from a multiple choice menu. The options include:

- It's annoying or distasteful
- It's pornography
- It goes against my views
- It advocates violence or harm towards a person or animal
- It's a fake news story
- It shows someone using drugs
- It harms or humiliates based on race, sex, orientation or ability
- It describes buying or selling drugs, guns or adult products
- It shows someone harming themselves or planning to harm themselves
- I think it's an unauthorised use of my intellectual property

The "fake news" option is, unsurprisingly, a recent addition to the list and does give credence to Facebook's claim that it is constantly updating in efforts to cope with content. What the list doesn't show, is how often Facebook gets it wrong.

In January, Facebook took down a photo of a statue of the sea god Neptune because it violated its policy on nudity, and last year the company removed the iconic, 1972 Pulitzer-winner photograph of the Vietnam War, which showed naked children running from a napalm attack. Globally, so many women complained that images of breastfeeding were removed, that the company was forced to draw up almost ridiculously-specific rules on what it would and wouldn't allow.

In its defence of these "false-positives" Facebook says although it has

thousands of people moderating content, sometimes they have to make split-second decisions. To someone like Modamani, who wants harassing content removed, split-second decisions aren't good enough.

Facebook had not responded to *International Politics and Society*'s requests for more details at time of publishing, so we cannot say how many takedown requests are managed by how many employees per hour. What we do know is that the company relies on "our community to report this content to us." According to Facebook, once a post has been reported, "one of the moderators then assesses the post according to *the social network's guidelines*, and decides whether it indeed needs taking down."

New laws?

The German Justice Ministry is currently considering new laws that would take that decision out of Facebook's hands. Some MPs have proposed that social media companies should hire legally qualified ombudsmen to carry out deletions, and be subject to fines of up to €500,000 if they fail to remove hate speech posts within 24 hours.

In October, following reports in *Die Zeit* that 100,000 Facebook posts had been deleted, the Ministry said that it had "no information with regard to the extent to which the deleted content was illegal."

"The taskforce set up by Minister Maas on illegal hate messages on the internet does not check whether individual instances of hate speech are illegal or, in particular, are criminal. Since Autumn 2015 it has developed standards for how the participating internet companies can effectively take action against hate messages. The companies check and delete, on their own responsibility, the content that is reported to them," it said.

Last month, German Economy Minister, Brigitte Zypries called on the European Commission to draw up laws against hate speech and fake news arguing that the current "soft approach" with social media firms "privatizes justice."

She may face an uphill battle as the Commission unveiled its "Code of Conduct" on illegal online hate speech to much fanfare last May. The code amounts to little more than a "gentleman's agreement" between the Commission and a group of IT companies including Facebook, Twitter, YouTube and Microsoft.

The companies promised to combat the spread of illegal hate speech

online in Europe, and the Commission in turn promised not to legislate the platforms.

"While the effective application of provisions criminalising hate speech is dependent on a robust system of enforcement of criminal law sanctions against the individual perpetrators of hate speech, this work must be complemented with actions geared at ensuring that illegal hate speech online is expeditiously reviewed by online intermediaries and social media platforms, upon receipt of a valid notification, in an appropriate timeframe," said the Commission.

By signing this code of conduct, the IT companies committed to developing internal procedures and staff training "to guarantee that they review the majority of valid notifications for removal of illegal hate speech in less than 24 hours and remove or disable access to such content, if necessary."

"With a global community of 1.6 billion people we work hard to balance giving people the power to express themselves whilst ensuring we provide a respectful environment. As we make clear in our Community Standards, there's no place for hate speech on Facebook. We urge people to use our reporting tools if they find content that they believe violates our standards so we can investigate. Our teams around the world review these reports around the clock and take swift action," said Facebook Head of Global Policy Management, Monika Bickert, at the time.

However a review conducted for the Commission in the last quarter of 2016 found that only 40% of flagged content was reviewed in less than 24h—a further 43% of cases were examined the following day. "Facebook assessed the notified content in less than 24 hours in 50% of the cases and in 41.9% of the cases in less than 48 hours. The corresponding figures for YouTube are 60.8% and 9.8% and for Twitter 23.5% and 56.1%, respectively," according to the report.

The amount of content removed is relatively small: Facebook removed the content in 28.3% of cases, Twitter in 19.1% and YouTube in 48.5%. A second monitoring cycle will be carried out during 2017 to observe trends. Meanwhile users, like Modamani, remain in the dark about how their cases are assessed.

Similar questions were raised following a Right to be Forgotten ruling against Google, but the search giant now publishes a "Transparency Report", which while certainly far from perfect, provides at least some statistics on how much content is de-indexed. Facebook, Twitter, et al would do well to take note.

There is indeed a fine line between protecting freedom of expression and stamping out online hate speech. The question the EU must answer is whether it wants private companies policing that line or laws enforced by the courts. Right now, most social media have stronger filters for nudity and copyright infringement than for criminal incitement to violence, if that continues Modamani's case will not be an isolated one.

Dutch MEP Marietje Schaake summed up the current paradox: "The rule of law must apply online as well as offline. Yet in practice, a lot remains unclear about what is legal and what is not, and especially about who decides. Is it just for Facebook to take down pictures of women who breastfeed, but not of people who are falsely framed as belonging to a terror network? Pressure on dominant tech companies like Facebook to take down illegal content as fast as possible is mounting. But there are no hard obligations in the e-commerce directive that address the specifics of such takedowns. Companies can implement these at will. If we want to avoid privatized law enforcement we need updates of EU laws with the aim of ensuring the public interest and people´s fundamental rights are better safeguarded," she told *International Politics and Society*.



Jennifer Baker
Brussels

Jennifer Baker has been a journalist in print, radio, and television for nearly 20 years, the last eight specialising in EU policy and legislation especially in the tech/digital sector. Her aim is to translate EU policy-speak and tech jargon into understandable English. She likes good books and bad films.