

LAITech2022

**TAL
TECH**
DEPARTMENT OF
LAW

AI AND CYBERBULLYING

Özden Serinay ÖZ

Law

Yasar University, Salzburg Paris Lodron University

28.02.2022

WHAT IS CYBERBULLYING ?

EXAMPLES OF CYBERBULLYING

- Cyberbullying is carried out by using information and communication technologies such as mobile phones, social media and instant messaging applications.
- Online harmful behaviours are discussed under five sub-headings; online harassment, cyberbullying and trolling, cyberstalking, revenge porn and image abuse, hate crime, hate incidents and hate speech.

EXAMPLES OF CYBERBULLYING

- *""Being I have been bullied many times before it makes me not want to live in this world anymore. I've been called probably every name out there.it really hurts. I sometimes contemplate suicide but i think where will that get me? i would only end up hurting the ones that i love."»*

- 14 year-old girl from WA

""Well i always get bullied and i hate it i feel like killing myself sometimes."" - **13 year-old girl from Victoria, Australia**

EXAMPLES OF CYBERBULLYING

Amanda Michelle Todd was a 15-year-old Canadian student and victim of cyberbullying who hanged herself at her home in British Columbia, Canada.





Someone in his school made a website about him saying » Welcome to the website that makes fun of Dave Knight» So many people in his school posted thoughtless and inappropriate comments about him.

He hid at home for weeks away from everyone outside, and he was frightened to death.



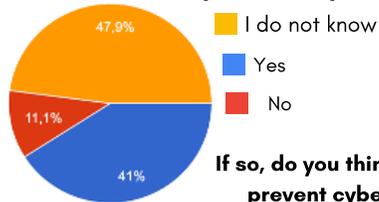
Lena Dunham, Tom Daley, Leslie Jones, Blac Chyna, and many more have been victims of cyberbullying. In the case of Jones, the actress was forced to quit Twitter for a time in response to ongoing harassment.



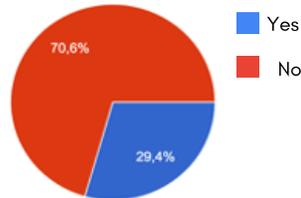
QUESTIONNAIRE RESULTS

- The **questionnaire** was shared with groups and connections through social media applications such as LinkedIn, Instagram and WhatsApp. By 28-30 January of 2022, **144 participants** from **28 countries** have answered the questionnaire.
- According to the questionnaire conducted by the author, the participants chose the application in which cyberbullying was most common as Instagram, Twitter, WhatsApp and Youtube.

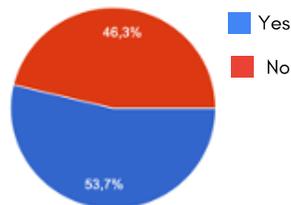
Are there legal sanctions against cyberbullying in your country ?



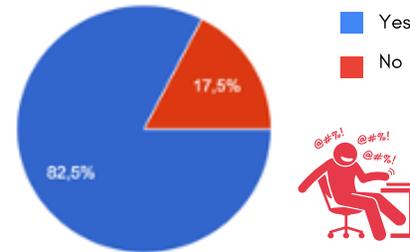
If so, do you think sanctions prevent cyberbullies ?



If so, do you think these sanctions raise awareness about cyberbullying ?



Can cyberbullying be prevented if the bullying is detected by AI technologies, punished according to the legal sanctions, and actions taken against the bully by a related platform?



FURTHER THIRD-PARTY RESEARCH RESULTS

- According to another research during a January 2021 it was found that 41 percent of internet users had personally experienced any kind of online harassment with 27 percent of respondents reporting to having experienced severe forms of online harassment such as physical threats, sexual harassment, stalking and sustained harassment. Furthermore, 77 percent of responding online harassment victims reported that they had been harassed on Facebook

Nearly 1 in 6 (15%) of online teens said they had experienced unwanted forwarding of private communication.

The top 3 countries where cyberbullying is the most prevalent are India, Brazil, and the United States, but it is a common occurrence everywhere.

WHAT IS THE DIFFERENCES FROM CYBERBULLYING AND HATE SPEECH ?

- Hate speech is defined as abuse that directly targets a unique factor beyond the control of the recipient.
- Hate speech is differentiated from cyberbullying in being defined as abuse directed specifically toward a unique, non-controllable attribute of a group of people.



KEY TERMS OF CYBERBULLYING AND HATE SPEECH

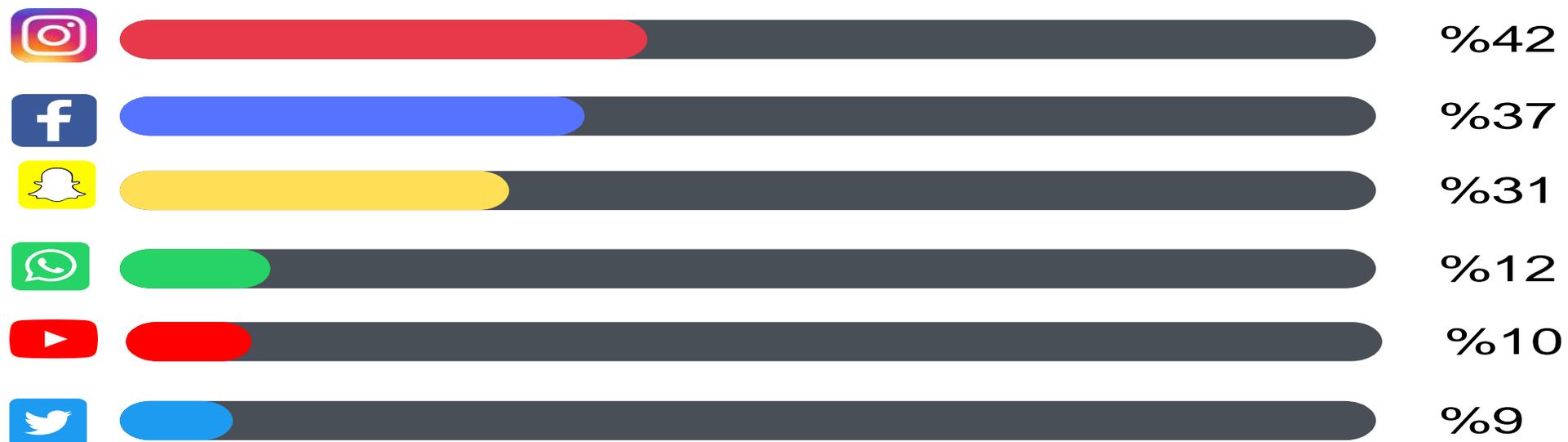
Hate speech – key terms

- | | |
|------------|------------------|
| 1 Faggot | 2 Bitch |
| 3 Nigga | 4 Dirty |
| 5 Fucking | 6 Hoe |
| 7 White | 8 Pussy |
| 9 Black | 10 LMAO |
| 11 Fuck | 12 Tranny |
| 13 Dyke | 14 Lesbian |
| 15 Gay | 16 Dumb |
| 17 Die | 18 Lady boy |
| 19 Shemale | 20 Gender bender |
| 21 He-she | 22 She-he |
-

Cyberbullying – key terms

- | | |
|-----------------|------------|
| 1 Bitch | 2 Dumb |
| 3 Fucking | 4 Fat/Ugly |
| 5 Idiot | 6 Stupid |
| 7 LMAO | 8 Moron |
| 9 Piece of shit | 10 Punk |
| 11 Hate | 12 Nigga |
-

CyberBullying In Social Media



WHAT ARE THE POLICIES AND REGULATIONS IMPLEMENTED IN THE APPLICATIONS ?



In 2016, Identity Guard partnered with Megan Meier Foundation to use IBM Watson technologies that enable natural language process (NLP) and natural language classifiers (NLC) to identify instances of cyberbullying or self-harm.

Parents receive alerts that are generated in instances of recognition. Then, parents can take actions to protect their children's experiences online.



In June 2016, Facebook introduced DeepText as "a deep learning-based text understanding engine that can understand with near-human accuracy the textual content of several thousand posts per second.

Facebook has a content moderating team that employs human in conjunction with using AI technologies to identify harassment online, monitor and remove fake accounts.

WHAT ARE THE POLICIES AND REGULATIONS IMPLEMENTED IN THE APPLICATIONS ?



Instagram started to use DeepText to eliminate spam on its platform. Instagram started to use technology to combat online trolls and harassment.



Twitter uses AI technology to spot spam and recognize negative interactions. Twitter helps users to curate their own experiences by presenting the best, positive experiences on the user's feeds. If you block someone on Twitter, chances are similar accounts from the same person will stop showing up on your Twitter feed.



Google and Jigsaw developed the AI moderating tool, Perspective, to score comments based on similarities with other comments categorized as "toxic" by human reviewers.



Youtube uses this tool to combat online harassment on its platform.

LAW AND CYBERBULLYING

AUSTRIA



Who, using telecommunication or computer systems, harm a person's dignity in public for a long time, or disclose information and pictures related to the private area of a person's life to a large number of people without consent, are sentenced to imprisonment of up to one year or a fine of up to 720 days.

GERMANY



There is no special regulation regarding cyberbullying in the German Penal Code. The German legislator, in order to respond to cyberbullying acts, has also passed Article 201a of the Law. .

SWISS



The Swiss Penal Code does not contain specific regulation on cyberbullying. It has been stated that there is no need for such a regulation in the doctrine, and that the classical forms of cyberbullying, which especially threaten the youth, are regulated as crimes in various provisions of the Law.

LAW AND CYBERBULLYING

The number of US states with state cyberbullying laws, by policy implementation: Cyberbullying laws in the US vary by state and policy implementation.

48 states include cyberbullying and online harassment in their laws while 44 states include criminal sanctions for cyberbullying in their laws.

In Australia, things are a little more clear. According to the Australian Cybercrime Online Reporting Network (ACORN):

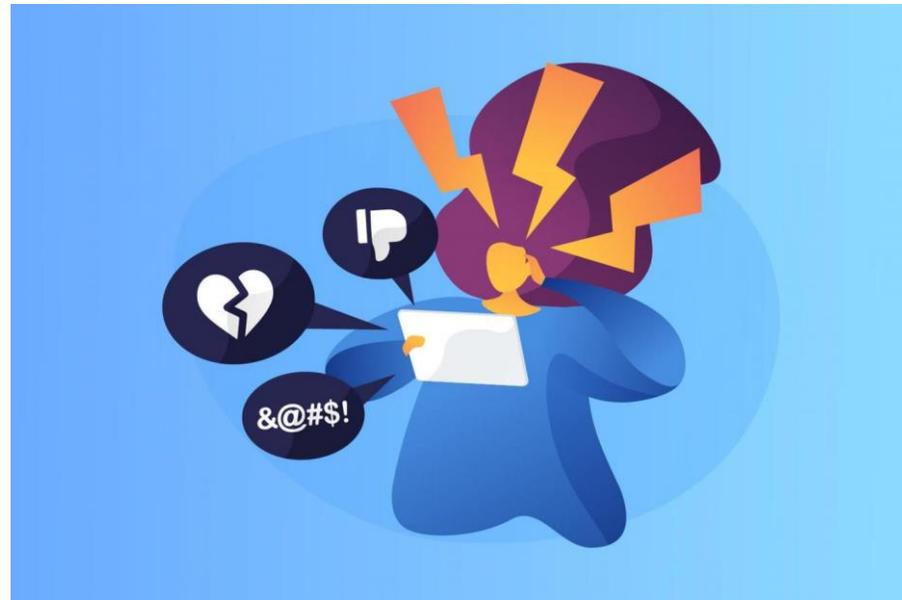
“Under the Criminal Code Act 1995 (Cth) it is an offence to use the internet, social media or a telephone to menace, harass or cause offence.”

This offence carries a **maximum penalty of up to three years in prison** or a fine in excess of \$30,000 AUD.

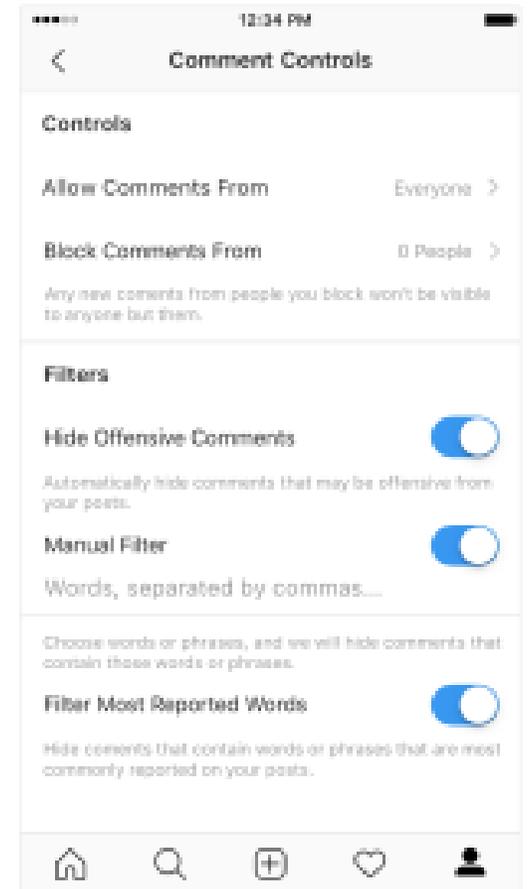
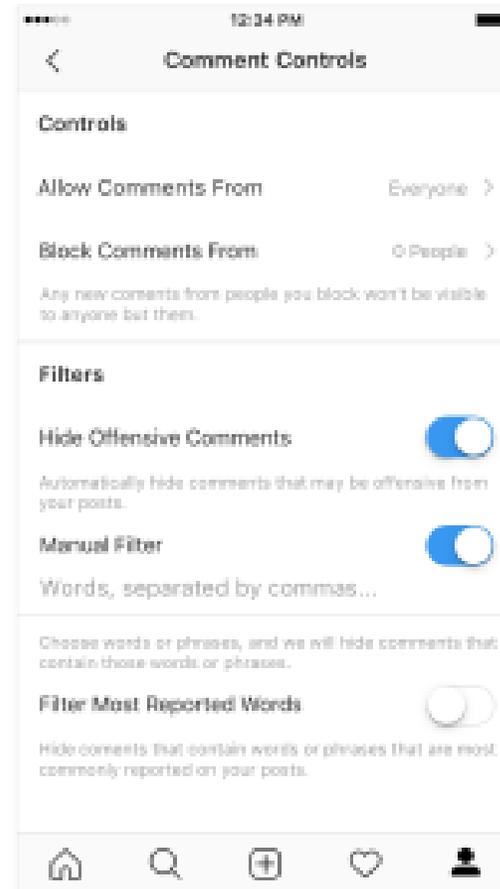
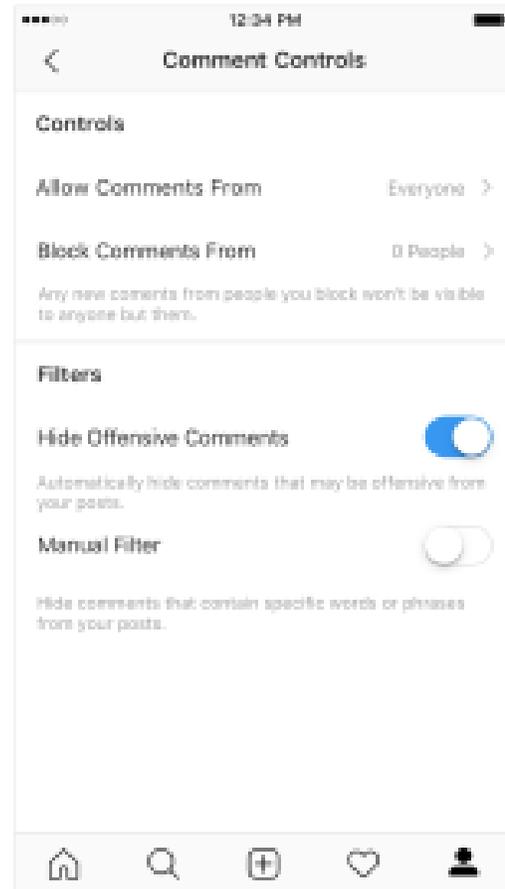
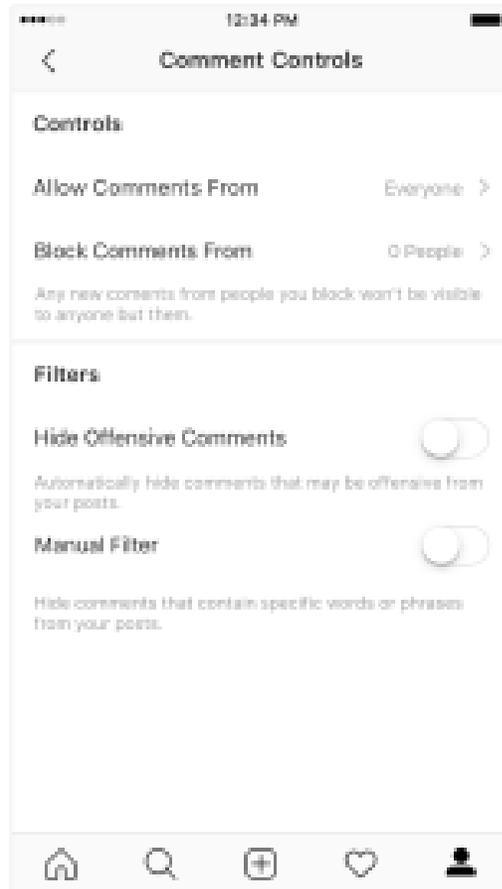
It's also noted that stalking laws — which vary between state and territory — may be applied to online behavior. These tend to carry large maximum penalties.

CAN AI PREVENT CYBERBULLYING ?

According to a study by L1GHT, a company that specializes in AI that is designed to detect and filter toxic content to protect children, online toxicity and cyberbullying on social media sites and video conferencing apps increased by up to 70% due to the pandemic. This included a sharp rise in toxicity and bullying directed at Asians.



WHAT ARE THE CYBERBULLYING PREVENTION ACTIVITIES OF APPLICATIONS?



WHAT ARE THE CYBERBULLYING PREVENTION ACTIVITIES OF APPLICATIONS?



Instagram is expanding its proactive filter for bullying in comment from the feed, Explore, and profile to also protect Live broadcasts. It's launching a "Kindness" camera effect in partnership with Maddie Ziegler, best known as the child dancer version of Sia from her music video "Chandelier".

WHAT ARE THE CYBERBULLYING PREVENTION ACTIVITIES OF APPLICATIONS?



Put a Stop to Bullying

Developed in partnership with the Yale Center for Emotional Intelligence, the Bullying Prevention Hub is a resource for teens, parents and educators seeking support and help for issues related to bullying and other conflicts. It offers step-by-step plans, including guidance on how to start some important conversations for people being bullied, parents who have had a child being bullied or accused of bullying, and educators who have had students involved with bullying.

Yale Center for Emotional Intelligence

Bullying Prevention Hub

In 2013, Facebook launched its dedicated Bullying Prevention Hub that was developed in partnership with the Yale Center for Emotional Intelligence. This is described as “a resource for teens, parents and educators seeking support and help for issues related to bullying and other conflicts.”

WHAT ARE THE CYBERBULLYING PREVENTION ACTIVITIES OF APPLICATIONS?

Parents Portal

The Facebook Parents Portal was rolled out in 2016 and offers parents some basic tips about helping their child navigate social media. While there isn't a large focus on bullying, there are some helpful sections about online privacy and links to guides for blocking and reporting other users within the platform.



In 2017, it implemented a couple of new features that could help prevent bullying and harassment.

For example, if you block an account and the user creates a new one, then Facebook will automatically block the new account on your behalf. Of course, this likely isn't foolproof because it's believed that Facebook uses IP addresses to identify users. As such, the blocked party creating the new account could simply use a VPN to bypass the restriction. Nonetheless, it's a step in the right direction.

WHAT ARE THE CYBERBULLYING PREVENTION ACTIVITIES OF APPLICATIONS?

October 2018, Facebook announced that it would implement tools that give users more control over what is posted about them on social media. For example, multiple comments on posts can be deleted or hidden at once, which can help when people are being bombarded with multiple hateful comments.

Facebook will also give people the option to report bullying or harassment on behalf of another user. Another feature Facebook is testing is the option for users to block certain words from appearing in their comments. Facebook-owned Instagram already has an anti-bullying feature that was released in May 2018.

HOW DEMOCRATIC IS TO INTERVENE NATURAL HUMAN INTERACTIONS WITH TECHNOLOGY ?

Artificial Intelligence has substantial social and cultural implications. As many information technologies do, AI raises issues of freedom of expression, privacy and surveillance, ownership of data, bias and discrimination, manipulation of information and trust, power relations, and environmental impact in relation to its energy consumption. Moreover, AI brings specifically new challenges that are related to its interaction with human cognitive capacities. AI-based systems have implications for human understanding and expertise. Algorithms of social media and news sites can help to spread disinformation and have implications for the perceived meaning of 'facts' and 'truth', as well as for political interaction and engagement. Machine learning can embed and exacerbate bias, potentially resulting in inequality, exclusion and a threat to cultural diversity. The scale and the power generated by AI technology accentuates the asymmetry between individuals, groups and nations, including the so-called 'digital divide' within and between nations. This divide may be exacerbated due to lack of access to fundamental elements such as algorithms for learning and classification, data to train and to evaluate the algorithms, human resources to code, set up the software, and prepare the data, as well as computational resources for storage and processing of data.

AI AND CYBERBULLYING AS A RESULT

Artificial intelligence is not at the desired level at the point of detecting cyberbullying, but people's awareness levels should be high. The awareness rate in countries with cyberbullying sanctions is higher than in other countries.

Therefore, the increase in legal sanctions will be a big step towards stopping cyberbullying in the development process of artificial intelligence.

All these researches and surveys show that cyberbullying is a growing situation that harms people. In this case, in order to combat cyberbullying, it is necessary to think about artificial intelligence, as understood from previous studies. With the help of law, technology and awareness awareness, cyberbullying can be enlightened.



LAITech2022 STUDENT CONFERENCE

[TALTECH.EE/EN/DEPARTMENT-LAW/LAITECH2022](https://taltech.ee/en/department-law/laitech2022)