Journal of Education and Learning; Vol. 10, No. 4; 2021
ISSN 1927-5250 E-ISSN 1927-5269
Published by Canadian Center of Science and Education

SafeGuard: A Web-Based Application to Guard Against Cyberbullying

Mudasser F. Wyne[1], Jigyasaa Sood[1], Christopher Kempton[1] & Thuyet Dao[1]

[1] College of Professional Studies, National University, USA

Correspondence: Mudasser F. Wyne, College of Professional Studies, National University, USA. E-mail: mwyne@nu.edu

Received: April 20, 2021 Accepted: May 25, 2021 Online Published: May 31, 2021

doi:10.5539/jel.v10n4p63 URL: https://doi.org/10.5539/jel.v10n4p63

Abstract

In this day and age there is a consistent rise in the usage of social media amongst all age groups thus, more people have established an online presence that makes these users more susceptible to becoming a target of cyberbullying. SafeGuard is a web-based application that monitors social media accounts and detects critical keywords such as "hurt", "die", etc., and phrases to detect cyberbullying threats and emotional distress in a person's social media post. Once a monitor discovers a match, an alert is generated for the end user in which they can initiate human intervention. This application is a new strategy for academic institution that allows them to be proactive rather than reactive against cyberbullying.

Keywords: bullying, cyberbullying, online relationships, social media

1. Introduction

Traditional physical bullying happens in playgrounds, work area and other aspects of daily life. In such cases the victim come in physical contact either verbally or physically with their bully. There is a widespread public concern regarding bullying among youth because of substantial impact on healthy and development. as indicated in published research (Eriksen, 2018; Eslea & Smith, 2012; Fekkes, Pijpers, & Verloove-Vanhorick, 2005; Papanikolaou, Chatzikosma, & Kleio, 2011; Tenenbaum, Varjas, Meyers, & Parris, 2011). Most of the interaction among younger generation these days is online where they express their emotions and ideas, such exchanges can be emotionally stressful for some thus causing negative impact on their personalities and mental development. The advancement of information technology started seeing a new form of bullying called cyber bullying. It is also referred to as electronic bullying, cyberbullying, or online social cruelty (Kowalski & Limber, 2007). The National Crime Prevention Council of Canada (Paullet & Pinchot, 2014) defines cyberbullying as "the process of using the Internet, cell phones or other devices to send or post text or images intended to hurt or embarrass another person." One also needs to understand that both cyberbullying and physical bullying have many things in common however, cyberbullying has distinct difference. Cyberbullying occurs when youth repeatedly use technology to threaten or harass their peers (Paullet & Pinchot, 2014).

In the 21st century as a society we spend a lot of time on one of the social media platforms and wake up to the tweets on Twitter, instead of newspaper. Among these users there are millions of youths, that post sensitive and disturbing content and are susceptible to cyberbullying, etc. Authors (Schneider, O'Donnell, Stueve, & Coulte, 2012; Lenhart, Maddeen, & Hitlin, 2005) report 93% of youth are active users of the Internet and 75% own a cell phone thus there is great potential for cyberbullying among these users. In cyberbullying, the offenders can hide his/her identity (Donegan, 2012) and can remotely contact the victim in numerous ways anonymously using any of the electronic social medium, thus may feel diminished responsibility and accountability as compared with traditional bullying (Juvonen & Gross, 2008; Mishna, Saini, & Solomon, 2009). In addition, the victim can be harassed round the clock by using modern technology which can otherwise be used to improve our quality of life. Kowalski and Limber (2007) conducted a study of 168 undergraduate students to understand student's perception of cyberbullying including their knowledge, opinion, and personal experiences if they have any. The paper reports 66% have witnessed and 50% of the students believe that cyberbullying has become a normal part of life. The paper also reports what participants of the study think needs to be done to avoid cyberbullying in future. In addition, authors in (Stacey, 2009) investigated the issues of cyberbullying and report on how students are coping with this issue through discussion with 74 students. Since students in academic institution have access to cyberworld via internet including computers and mobile phones. The research in the paper states that social networking sites and synchronous chat sites are places where cyberbullying most commonly occur, with email

and texting on mobile phones.

The frequency of suicide among youth victims of cyberbullying at national level (Schneider, O'Donnell, Stueve, & Coulte, 2012; Lenhart & Hitlin, 2005) has raised concerns thus many states in USA now have legislation in place that requires academic institutions to have policies to address cyberbullying in their antibullying polices. However, academic institution may lack the ability or knowhow on how to detect cyberbullying and identify which youths are offenders and which are victims before it causes any physical or mental damage to the individual. Bullying in general is a problem among youth sometimes as a victim or as a bully or even both as well as not limited to academic institutions (Burk, DiRenzo, & Fuller, 2019). It is because of rising violence in academic institutions in USA and cyberbullying amongst youth we have conclude that there is an immediate need to develop an application to address this situation.

2. Proposed Application

Recently, there has been a rise in mass shootings, cyberbullying, and suicide cases in academic institutions, where a good portion of youth usually post on their social media platform before committing to any of these acts. If someone in the school management can identify such individuals as well their post in a timely manner, then there is a possibility that prospective victim can be protected and saved from any physical and/or mental harm. The goal is to tackle the imminent danger of school shootings, suicides, and cyberbullying by carefully monitoring student's social media activities for any mentions of suicide, death, violence, etc. and sending an alert to the school officials in a timely manner. Thus, we are proposing an application that will monitor social media postings, utilizes web technologies, and is hosted in a cloud environment that is accessible from any location. It provides seamless integration with an academic institution's existing security solutions that will allow for better threat and emotional distress detection of the students using Twitter in the academic institution. The intent of this application is to alert users and institution's administration before a tragedy occurs by looking at suspicious signs on students' online posts. Using this web-based application, provides users a way to create and maintain lists of student accounts as well as monitor their social media text exchanged or sent that will be used to detect potential threats or emotional distress by parsing posts from a student's social media account, like Twitter. By monitoring social media posts, SafeGuard will be able to determine if the language used, matches any preset keywords or phrases. This will generate an alert in the system if a condition is met and will notify the end user of the system. It will require an initial setup and minimal maintenance to achieve a high level of monitoring, the content will be filtered down to only alert about specific posts which would otherwise not be possible if done by human monitoring alone. This greatly decreases the amount of time and personnel effort needed to monitor online media posts among student population. Though the primary focus of this product is that detection of a violent act or identifying a person in emotional distress can prevent tragic incidents occurring such as suicide, shooting or cyberbullying.

3. Application Requirement

A major benefit is to catch an issue before it becomes something big and/or tragic. The institution and law enforcement will anticipate the incident through our alerts and will be more prepared and might even be able to prevent that tragic incident from occurring. It will not only benefit other innocent people involved in the tragedy, but also the youth who might be suffering from pain and anger from being bullied by getting them the help they need with the help of more informed officials. The academic institute management need to be educated on ensuring that they are maintaining up to date student lists, monitors, and users to ensure their data size does not grow larger than it needs to be which could result in slower queries. Another concern with data size and querying if the creation of monitors is very generic then in turn it will generate too many notifications that would make the client need to sort through more notification than needed.

In addition, the client must be aware of physical security of the devices, maintaining users in the system, creating monitors, reading notifications, maintaining student information in the institutions and ensure they are protected from unauthorized use. They will be responsible for creating policies and directives around their use of the application and will not be aided by SafeGuard other than providing access to the technology for them to use. The institution needs to whitelist their desired IP addresses with the application so access to the application is limited to those specific devices. Passwords must be strong and meet basic requirements such as lowercase, uppercase, numbers, special characters, and a minimum length. An access control restricted by IP address is used for the client's access by their network, JSON web token is used that has an expiration. All clients using the application requires authentication and authorization on each request. The application uses the HTTPS protocol for client requests.

4. Data Flow Diagram

Figure 1 depicts basic flow of data through SafeGuard application's system and its interaction with external entities. It has two inputs: Contact (the user of the social media platform that system is trying to monitor) and User (the person who is monitoring contact) and output: Notification. As soon as a contact posts a message, text, tweet, etc., and a user has this person as their contact, SafeGuard application checks the message for match with any keywords that the user was tracking, and if detected, the system would output an alert message for the user.

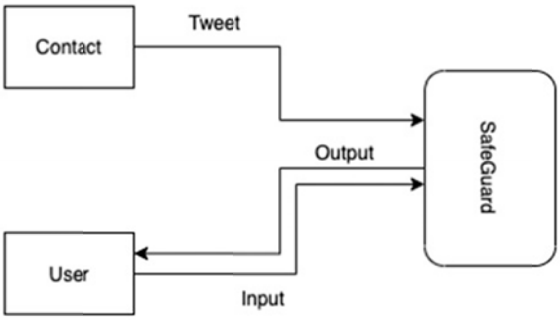

Figure 1. Contextual data flow diagram

Figure 2 shows a more detailed flow of the data through the system. As soon as a contact posts a tweet, SafeGuard application is supposed to fetch that tweet from the Twitter API (or mock twitter/webhook for this prototype currently) and handles the response and stores that tweet as a post in MongoDB document database. It then fetches the monitors that a user has and inputs that contact's post into the ElasticSearch database to look for the list of monitors (keywords) in the particular post. If a match is detected, meaning that a monitor is found, then SafeGuard generates a notification and sends that to the user on their SafeGuard website account.

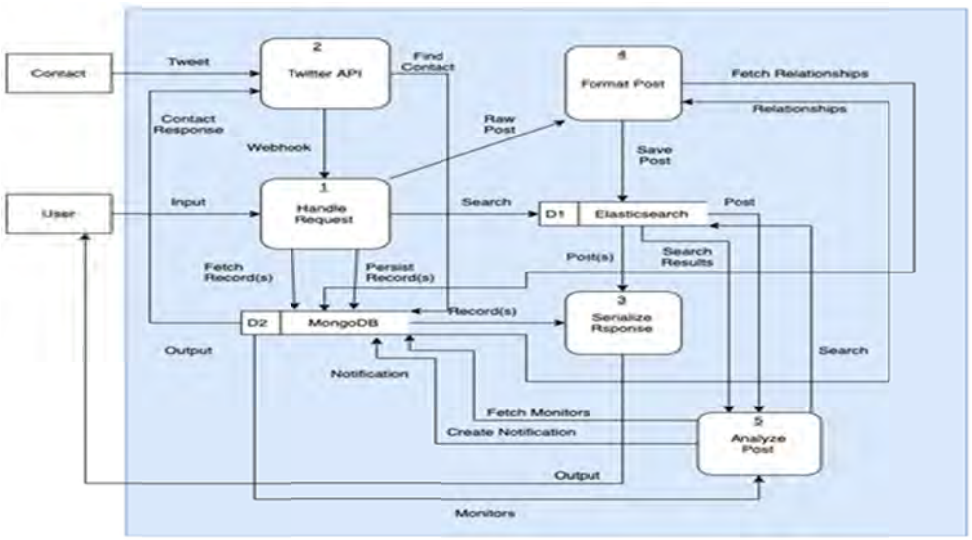

Figure 2. Data flow diagram

This is the process for analyzing a post to look for a match of monitors in the post's text notifying a user immediately upon match detection. A User can go to the notifications page to view, change status of (read, unread, resolved, unresolved), and delete a notification. A User can go to the posts page to view all the posts made by their contacts and can use the contact page to view, update, or delete a contact. A User can also go to the monitors page to add or delete monitors (keywords). In addition, users page can also be updated in order to add, update, and delete users and admins in the system (granted that they have admin privileges to do so).

Figure 3 shows details of the process used for authentication and fetching user records. Figure 4 shows the process for saving a tweet from a contact into MongoDB database.

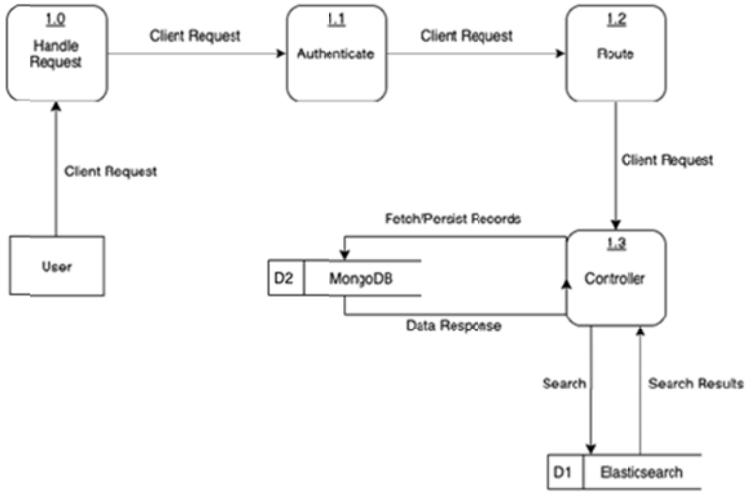

Figure 3. Authentication and fetching user record

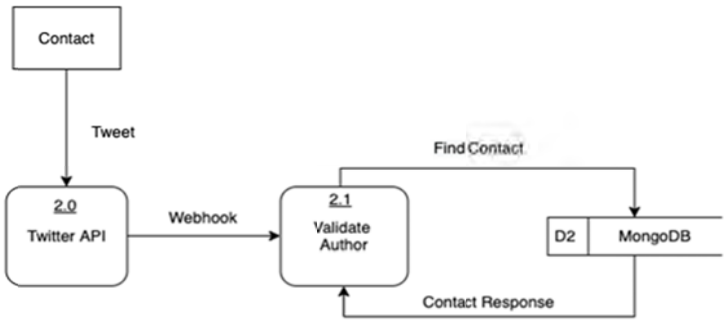

Figure 4. Saving a tweet into MongoDB

5. Application Working

Following are the screen shots (Menu options) of the user interface of the initial SafeGuard prototype that helps to understand working of the proposed application. The user of the application "Login Page" Figure 5, will determine the level of privileges for the person login to SafeGuard based on the login credentials. The level of users are "User" and "Admin", as shown in Figure 6.

Figure 5. Login page

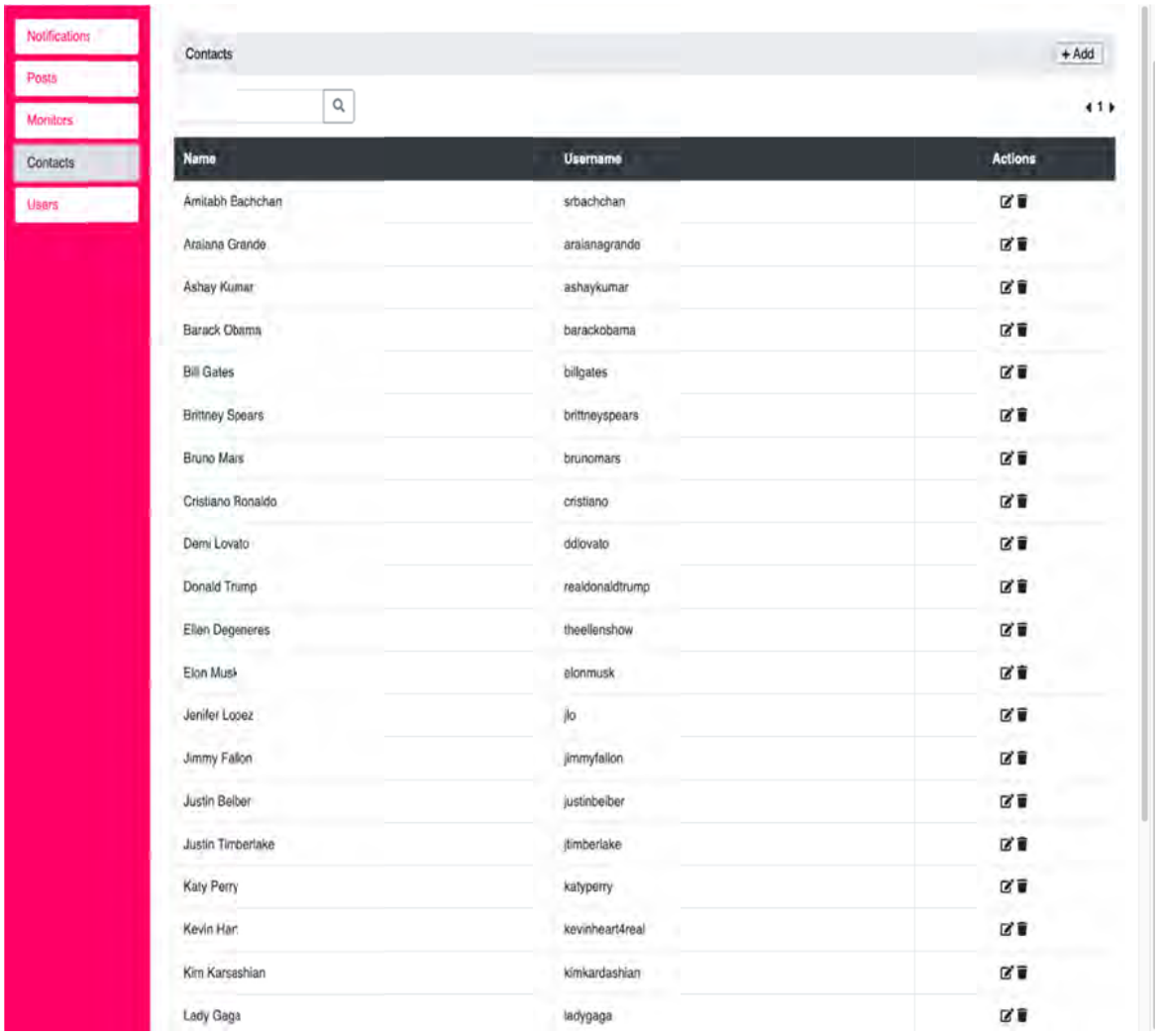

Figure 6. Users view page

The User is any member of the academic institution management team, who will be managing all accounts, notifications, and posts. The Admin account is for a person responsible for managing the system and has all administration privileges. The Monitor is the list of Keywords that the academic institution wants to look for in all posts, text messages etc. on multimedia platforms. Once monitor discovers a match, an alert is generated in the form of notifications for the User after which any form of necessary action can be taken for example request for human intervention through law enforcement.

The institution management will assume responsibility for physical security, maintaining users in the system, creating monitors, reading notifications, maintaining contact (student) information, and alerting higher authorities when there is a need. If the keywords are very general, then more notification would be generated than needed thus requiring more time to sort through these notifications and determine which are more serious and relevant, so it is advisable to use very specific keywords. If a match is detected, meaning that a keyword is found, then the SafeGuard generates a notification and sends that to the user on their SafeGuard website account. A User can go to the notifications page, Figure 7, to view, change status of (read, unread, resolved, unresolved), and delete a notification made. A User can go to the posts page, Figure 8, in order to view all the posts made by their contacts as well as to view, update, and delete a contact.

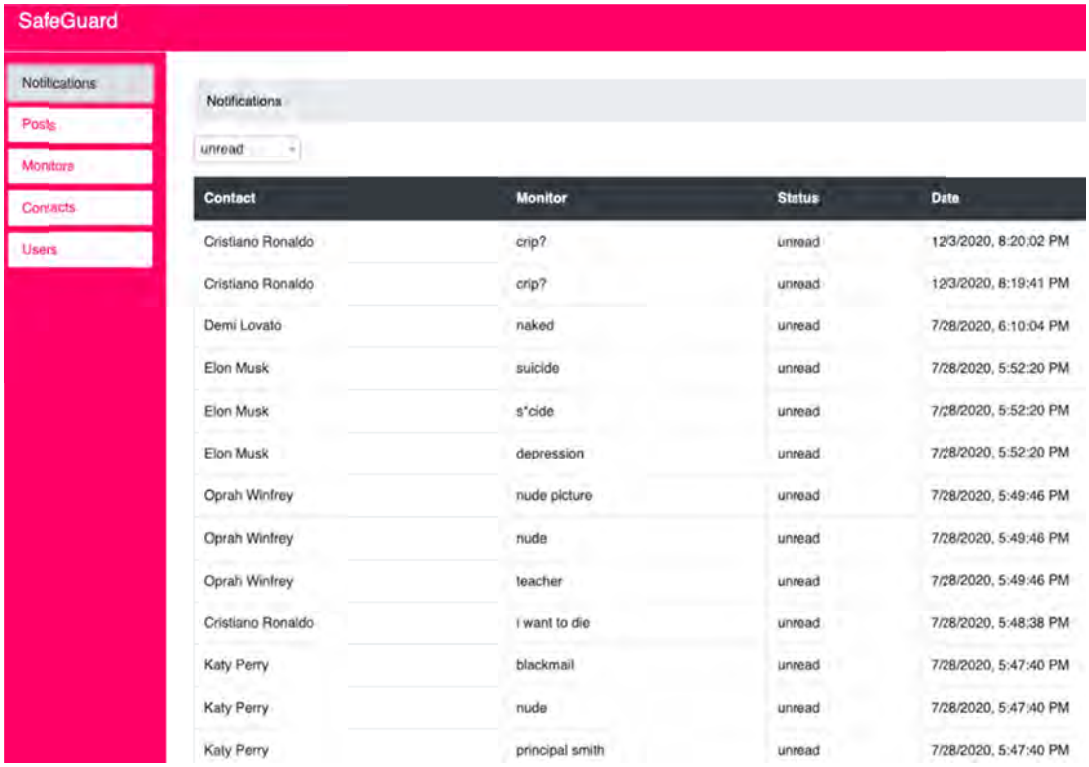

Figure 7. Notification page

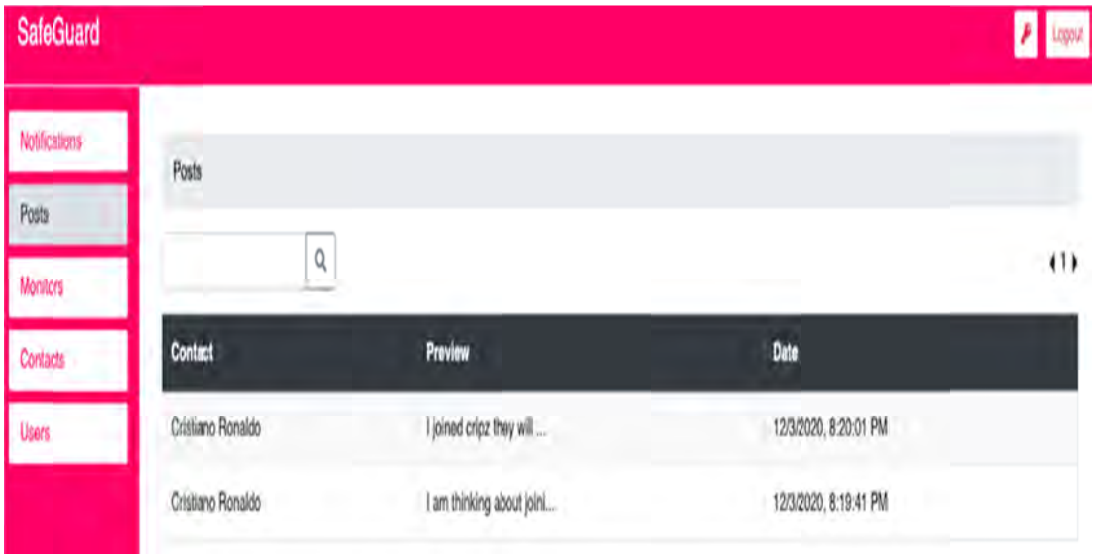

Figure 8. Posts page

6. Conclusion and Recommendation

This paper presents “SafeGuard” a web-based application that is used to monitor social media posts and detect keywords (Monitors) and phrases to generate alerts for human intervention. Monitors can be set up to detect cyberbullying, threats and emotional distress thus allowing an institution to be proactive rather than reactive. The SafeGuard App needs to expand upon the functionality by adding more features such as adding weightage factor to monitors, so as to favor looking for certain keywords more than others such as the keyword “rape” having more weight as a monitor than the keyword “slap”. By adding weights, it would help the system filter out dangerous posts more finely and accurately than before.

References

Burk, B. N., DiRenzo, A., & Rachelle, H. F. (2019). Who is a Bully Anyway? Examining Perceptions of Bullying among Parents and Children. *Journal of Park and Recreation Administration, 37*(2), 88–98. https://doi.org/10.18666/JPRA-2019-8816

Donegan, R. (2012). Bullying and cyberbullying: History, statistics, law, prevention and analysis. *The Elon Journal of Undergraduate Research in Communications, 3*(1), 33–42.

Eriksen, I. M. (2018). The power of the word: Students' and school staff's use of the established bullying definition. *The Journal of Educational Research, 60*(2), 157–170. https://doi.org/10.1080/00131881.2018.1454263

Eslea, M., & Smith, P. (2012). Pupil and parent attitudes toward bullying in primary schools. *European Journal of Psychology of Education, 15*(2), 207–219. https://doi.org/10.1007/BF03173175

Fekkes, M., Pijpers, F. I. M., & Verloove-Vanhorick, S. (2005). Bullying: Who does what, when and where? Involvement of children, teachers and parents in bullying Behavior. *The Journal of Health Education Research, 20*(1), 81–91. https://doi.org/10.1093/her/cyg100

Juvonen, J., & Gross, E. (2008). Extending the school grounds? Bullying experiences in cyberspace. *Journal of School Health, 78*(9), 496–505. https://doi.org/10.1111/j.1746-1561.2008.00335.x

Kowalski, R. M., & Limber, S. P. (2007). Electronic Bullying Among Middle School Students. *Journal of Adolescent Health, 41*, 22–30. https://doi.org/10.1016/j.jadohealth.2007.08.017

Lenhart, A., Maddeen, M., & Hitlin, P. (2005). *Youth are Leading the Transition to a Fully Wired Mobile nation.* Pew Internet & American Life Project, Teens and Technology.

Mishna, F., Saini, M., & Solomon, S. (2009). Ongoing and online: Children and youth's perceptions of cyber bullying. *Children and Youth Services Review, 31*(12), 1222–1228. https://doi.org/10.1016/j.childyouth.2009.05.004

Papanikolaou, M., Chatzikosma, T., & Kleio, K. (2011). Bullying at school: The role of family. *Procedia - Social and Behavioral Sciences*, 29, 433–442. https://doi.org/10.1016/j.sbspro.2011.11.260

Paullet, K., & Pinchot, J. (2014). Behind the Screen Where Today's Bully Plays: Perceptions of College Students on Cyberbullying. *Journal of Information Systems Education, 25*(1), 63–69.

Schneider, S. K., O'Donnell, L., Stueve, A., & Coulte, R. (2012). Cyberbullying, School Bullying, and Psychological Distress: A Regional Census of High School Students. *American Journal of Public Health, 102*(1), 171–177. https://doi.org/10.2105/AJPH.2011.300308

Stacey. E. (2009). Research into Cyberbullying: Student Perspectives on Cybersafe Learning Environments. *Informatics in Education, 8*(1), 115–130. https://doi.org/10.15388/infedu.2009.08

Tenenbaum, L., Varjas, K., Meyers, J., & Parris, L. (2011). Coping strategies and perceived effectiveness in fourth through eighth grade victims of bullying. *The Journal of School Psychology International, 32*(3), 263–287. https://doi.org/10.1177/0143034311402309

Copyrights