

ENhancing seCurity and privAcy in the Social wEb: a user-centered approach for the protection of minors



A few words about protecting children

Michael Sirivianos,
Cyprus University of Technology



**ReCRED H2020 Project Clustering
Workshop**

Athens, January 2018



Horizon 2020
European Union funding
for Research & Innovation

Funded by the Horizon H2020 Framework Programme of the European Union under grant agreement no 691025.

theguardian

home > UK > society law scotland wales northern ireland education med all

Cyberbullying

One in five young people has suffered online abuse, study finds

Survey of more than 4,700 teenagers reveals almost half think cyberbullying is a bigger issue than drug abuse among young people



A fifth of teenagers who had experienced cyberbullying said it had made them consider suicide. Photograph: Graeme Robertson for the Guardian

newsbeat

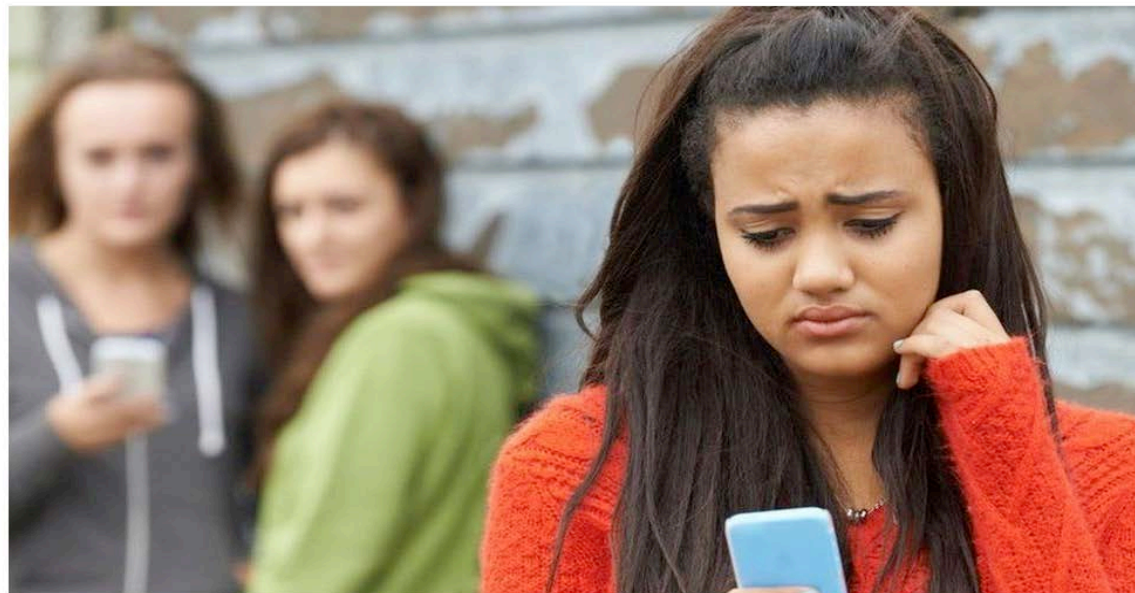
Latest

Popular

Topics

Cyberbullying 'worse than face-to-face' abuse, suggests global research

LIFE | 22 Sep 2015



One-in-five young people has been cyberbullied, according to a research across 11 countries.

Most Popular

Cyber bullying is increasing – study

GAUTENG: *Youths from Grade 8-12 are victims*

BULLYING of young people between Grade eight and 12 was on the rise, especially cyber bullying, according to a study released by Unisa's Youth Research Unit (YRU) on Sunday.

"The YRU research study highlighted that almost three in every 10 pupils (34.4%) who participated in the study were victims of bullying," stated the study.

"Whereas four in 10 pupils (38.1%) were aware of a friend being bullied, about a quarter (23.3%) admitted to having bullied other pupils."

The study included a total of 3 371 Gauteng high school pupils and formed part the Schools Community Engagement research project.

It showed high rates of traditional forms of bullying and the emergence of cyber bullying as a new phenomenon.

"This was evident from 55.3% of learners who had experienced emotional bullying while almost one in every five (16.9%) had experienced cyber bullying."

The authors of the study believed the emergence of cyber bullying needed to be noted as having an impact on the lives of young people,

Snip

The study included 3 371 Gauteng high school pupils and showed high rates of traditional forms of bullying and cyber bullying as the new phenomenon.

families and communities.

If ignored, it could reach "crisis levels" which would impact negatively on broader society.

The study found that 67.7% of participants felt bullying had increased over the past two years, with 40.3% feeling cyber bullying had increased in the same period.

Of the pupils who were cyber bullied, 53.6% received upsetting messages, 49% had rumours and gossip spread about them, and 48% were called names.

"Other cyber bullying incidents reported includes exposure to sexual remarks (24.5%), unflattering and suggestive personal photos spread online (13.3%) and being bullied, recorded and photos/videos distributed online (8.7%)."

The figures also showed evidence of "sexting" as a new phenomenon emerging among young people.

The result of cyber bullying saw

74.5% of victims avoid chat rooms, 25.5% of victims avoided MXit.

The majority of bullying took place through SMSs and cell-phone-accessed social networking.

Traditional bullying took place through name calling, physical encounters, threats and intimidation, with the majority of bullying taking place off school premises.

Professor Deon Tustin, executive director of the Bureau of Market Research in which the YRU was situated, said the depression and sadness caused by bullying affected young person's emotional development and functioning.

This could contribute to youth being trapped in a vicious cycle of exploitation and not being able to cope psychologically within the learning environment and broader society, he said.

It could also lead to self-destructive behaviour.

The study did find that 51.6% of victims told their parents, peers, teachers or NGOs, while 48.7% said appropriate action was taken to prevent bullying.

Lack of action perpetuated violence as bullies faced no consequences, potentially increasing risk for victims. – *Sapa*.

OFFERS FANTASY BINGO DATING JOBS BUYSSELL COMPETITIONS HOROSCOPES CARTOONS CROSSWORDS COOKIE POLICY

EMAIL ALERTS
Get the latest stories direct to your inbox >

Mirror
WEBSITE OF THE YEAR

Our new FREE Mirror apps are here! >

News - Politics Football Sport - Celebs - TV & Film Weird News
Most read ★ Videos 📄 Quizzes

TRENDING MILLY DOWLER BEDROOM TAX NATIONAL LOTTERY ILLEGAL IMMIGRANTS WEATHER
Technology Money Travel Fashion Mums

M · News · UK News · Crime

Facebook 'sexual predator' posed as teenager on social networking site so he could groom young girls

19:53, 29 NOV 2015 | UPDATED 20:34, 29 NOV 2015 | BY ABBY BOLTER

Police said Aled Clement showed victims bundles of cash - and was "aggressive in his advances"

    319 SHARES |  COMMENTS

Enter your e-mail for our daily newsletter

Subscribe

★ Recommended In UK News



ADAM JOHNSON

Adam Johnson trial: Live updates as ex-Sunderland star appears in court accused of sexual activity with a child

❑ Popular among young people and children

❑ Threats:

- Cyberbullying
- Online sexual abuse
- Distribution/Possession of child abusive material
- Fake information propagation



□ OSNs allow cyberbullies/sexual predators to:

- Reach countless children
- Remain relatively anonymous
- Avoid being held accountable for their actions



- ❑ Equip children and custodians with a set of tools
 - Safeguard their safety and privacy
 - Raise awareness regarding potential threats

- ❑ Extend and use the following techniques
 - Sentiment and affective analysis
 - Fake activity detection in OSN's
 - Content detection and protection



- ☐ Funded by the EU under H2020
- ☐ Call Identifier: H2020-MSCA-RISE-2015



ARISTOTLE
UNIVERSITY OF
THESSALONIKI



encase.socialcomputing.eu

- ❑ Adoption of the developed technology by stakeholders (OSNs, parents, educators, etc.)

- ❑ Create new and lasting research collaborations
 - Transfer of Knowledge (ToK)
 - Improvement of R&I potential

- ❑ Provide new career perspectives through R&I



- ❑ Create strong industry-academia cooperation

- ❑ Conduct interdisciplinary research
 - User experience design
 - Data mining and machine learning
 - Security and privacy

- ❑ Enrich researchers' existing background and business circle
 - Active participation in ambitious research
 - Exchange of researchers to various partners
 - Participation to networking events organized by ENCASE

To protect minors from malicious actors in Online Social Networks

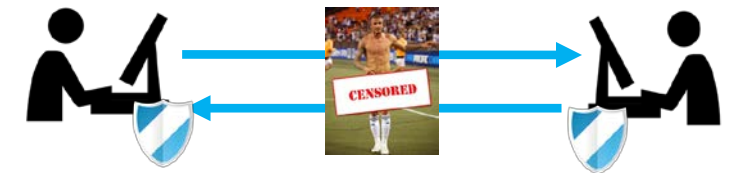
1. Cyberbullying, malicious and distressed behavior detection



2. Fake Activity Detection and suppression



3. Sensitive Content Detection and Protection



- ❑ Understand security and privacy concerns of OSN users

- ❑ Research methods for the analysis of OSN information to:
 - Perform user profiling
 - Perform sentiment and affective analysis
 - Reveal users under distress

- ❑ Extend state of the art in:
 - Fake account detection
 - Audience boosting detection
 - False information propagation



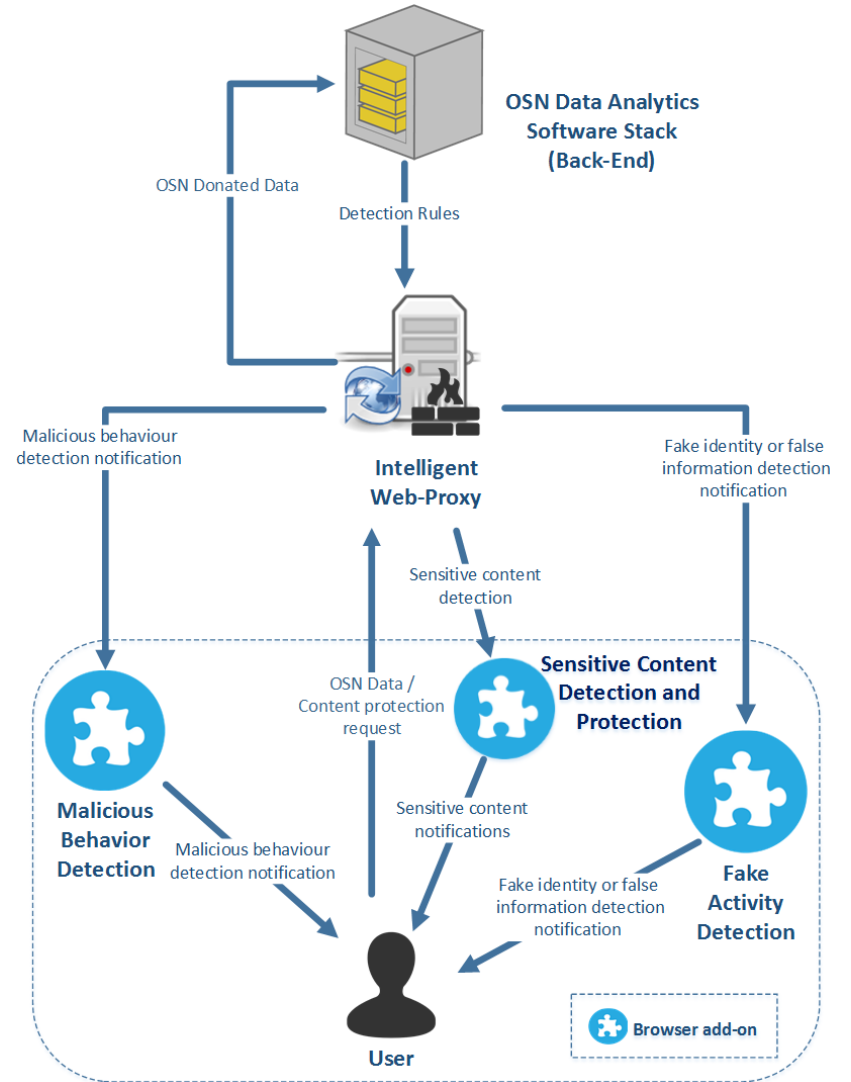
- Design and implement web-based user interfaces that:
 - Discourage users from befriending suspicious users
 - Warn custodians about online abuse of minors
 - Discourage users from sharing sensitive information

- Design and implement effective browser-based content protection by employing
 - Watermarking
 - Steganography
 - Advanced encryption techniques



Architectural components:

- Front-end
- Intelligent Web-Proxy
 - Malicious behavior detection
 - Fake identity and false information detection
 - Sensitive content detection and protection
- Data Analytics Software Stack (Back-end)
 - generates detection rules for the Web Proxy



- ❑ Placing functionality (filters, text replacement, etc.) in the **Intelligent Web-Proxy** instead of the browser add-ons when it can be correctly and efficiently implemented
 - We wish to prevent minors from modifying or disabling the functionality of the ENCASE system through the browser add-ons



DO's



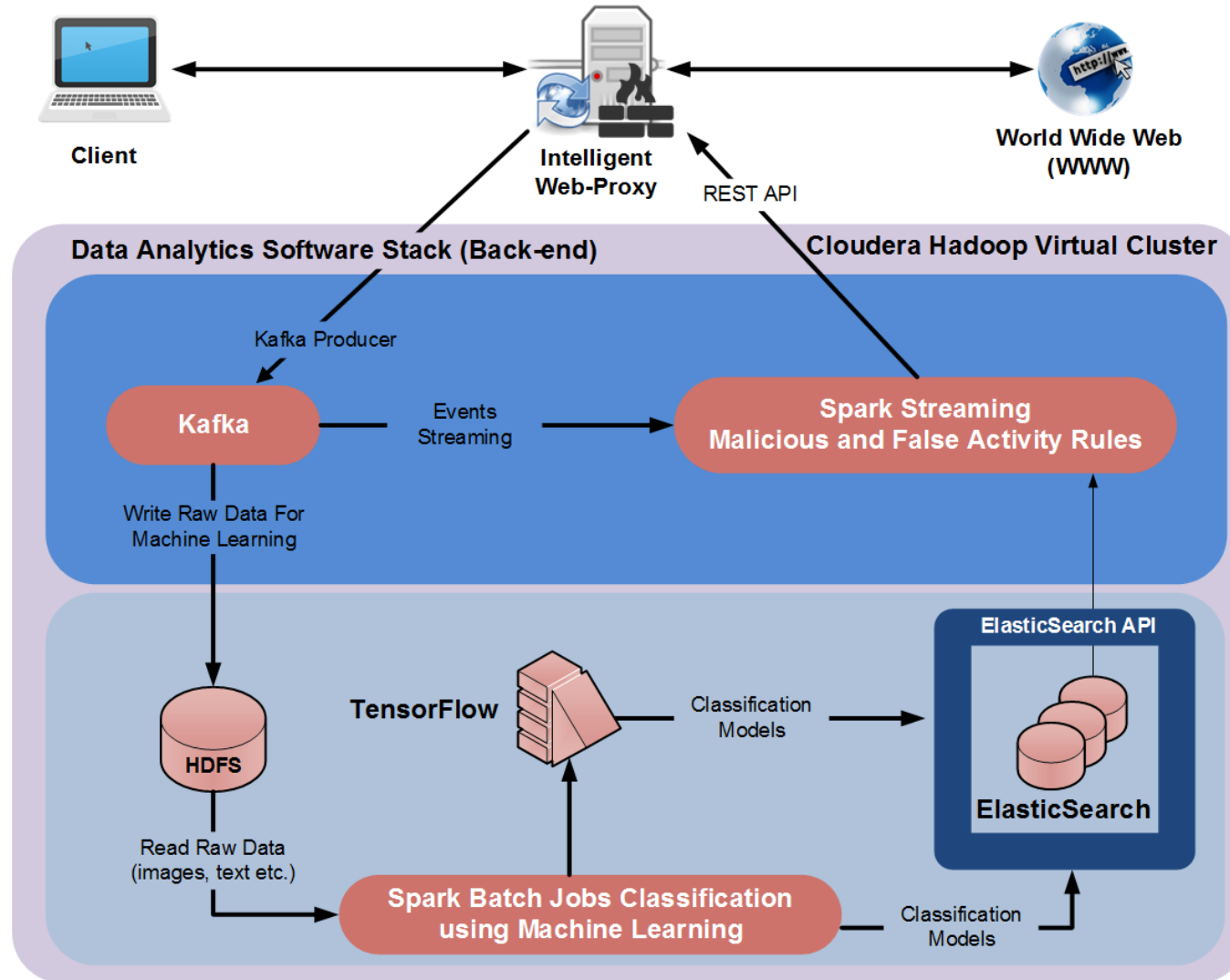
DONT's

- ENCASE aims to provide the ability to seamlessly support multiple types of clients (desktop browsers, mobile apps, etc.) with minimal user or client platform configurations
- Placing functionalities on the Intelligent Web-Proxy means less requests from the Back-end
- Solution to single point of failure. The functionalities placed in the Intelligent Web-Proxy are available even if the Back-end is down

□ Rules and trained classifiers are generated in the **Back-End**.

- Trained classifiers will be placed on the Intelligent Web-Proxy only if they can run efficiently
- The Back-end collects data from all the Intelligent Web-Proxies to generate detection rules or trained classifiers





☐ Warning, flagging and feedback functionality on the **browser add-ons**

- Warnings are displayed to the user (minor) through the browser add-ons
- The users are able to flag content through the browser add-ons in case the Intelligent Web-Proxy failed to identify so
- The users are able to give their feedback based on the activity detected by the Intelligent Web-Proxy



- Visibility of content. Define what the parent and Back-end can “see”
 - The parent is able to set up the **Visibility settings** of content using a fine-grained console and always with the consent of the minor
 - This way we enable various levels of monitoring for parents and the Back-end with the consent of the child, while keeping the child fully aware of what his parents and the Back-end can see (wall, chat messages, friends list, etc.)
 - The minor can see what options the parent has selected and she is be able to confirm or decline these options anytime



The Web Centipede: Understanding How Web Communities Influence Each Other Through the Lens of Mainstream and Alternative News Sources

Savvas Zannettou^{*}, Tristan Caulfield[†], Emiliano De Cristofaro[†], Nicolas Kourtellis[‡], Ilias Leontiadis[‡], Michael Sirivianos^{*}, Gianluca Stringhini[†], and Jeremy Blackburn⁺

^{*}Cyprus University of Technology, [†]University College London, [‡]Telefonica Research,
⁺University of Alabama at Birmingham

sa.zannettou@edu.cut.ac.cy, {t.caulfield,e.decrisofaro,g.stringhini}@ucl.ac.uk,
{nicolas.kourtellis,ilias.leontiadis}@telefonica.com, michael.sirivianos@cut.ac.cy, blackburn@uab.edu

The first systematic study of how fake news spread from fringe communities to mainstream social networks, ACM Internet Measurement Conference '17

Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web

Savvas Zannettou^{*}, Tristan Caulfield[†], Emiliano De Cristofaro[†],
Michael Sirivianos^{*}, Gianluca Stringhini[†], Jeremy Blackburn[‡]

^{*}Cyprus University of Technology, [†]University College London, [‡]University of Alabama at Birmingham
sa.zannettou@edu.cut.ac.cy, {t.caulfield,e.decrisofaro,g.stringhini}@ucl.ac.uk
michael.sirivianos@cut.ac.cy, blackburn@uab.edu

State-sponsored Internet Research Agency trolls are not that influential except when it comes to sharing Russia Today links

- ❑ It was founded in 2004. Started operating in 2007
- ❑ Located in the city center of Limassol
- ❑ Ranked 350-400 by Times Higher Education
- ❑ Six faculties:
 - Geotechnical Sciences and Environmental Management
 - Management and Economics
 - Communication and Media Studies
 - Health Sciences
 - Fine and Applied Arts
 - **Engineering and Technology**
- ❑ Approximately 3K students



Thank you!



ARISTOTLE
UNIVERSITY OF
THESSALONIKI



facebook.com/ENCASE.H2020



encase.socialcomputing.eu



twitter.com/ENCASE_H2020