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

Cybersafety Threats and Intelligent Parental Advice Tools for Protection in Social Networks

Michael Sirivianos
Assistant Professor
ENCASE Coordinator

Cyprus University of Technology

MSCA Artificial Intelligence Cluster

Brussels, December 10, 2019

European Commission | Horizon 2020 European Union funding for Research & Innovation

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

Consortium and Budget

- Research funding under the Marie Skłodowska-Curie RISE Action
- Total project funding: € 2,110,500
- Funding for Cyprus Univ. of Technology: € 633,780

Cyprus University of Technology
Telefónica I+D
CyRIC
UCL
LS TECH
Signal GeneriX Ltd
Advanced Signal Solutions
ROMA TRE
UNIVERSITÀ DEGLI STUDI

Research and Innovation Achievements

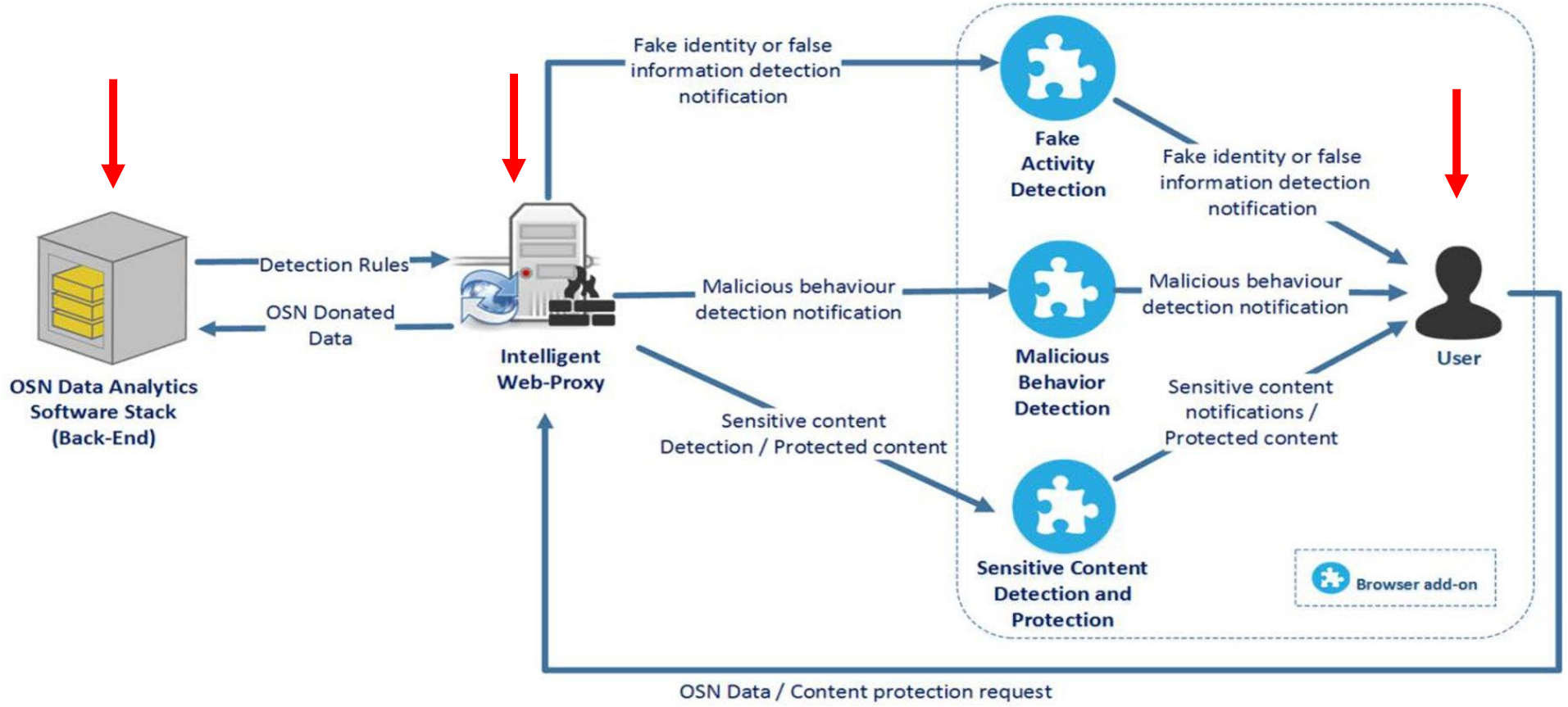
- Detecting and characterizing aggressive behavior
- Fake news and hateful memes studies
- State-sponsored trolls investigation
- A suite of parental advice tools

Astroscreen

YouTube
reddit
4chan
twitter

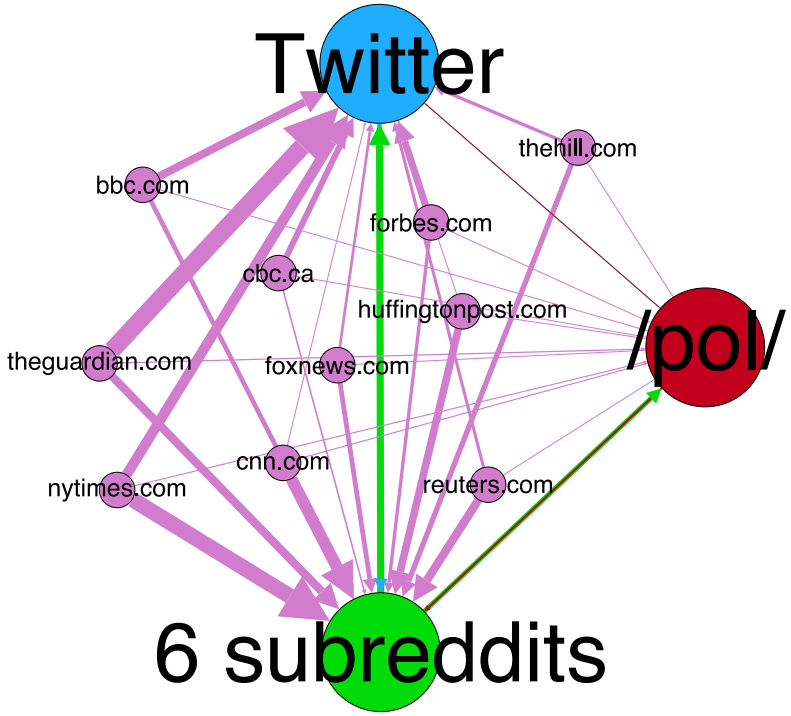
ENCASE Architecture

- Our architecture aims to bring custodians, educators, and children closer together
- Provide advice and facilitate cybersafety discussions between children and their guardians

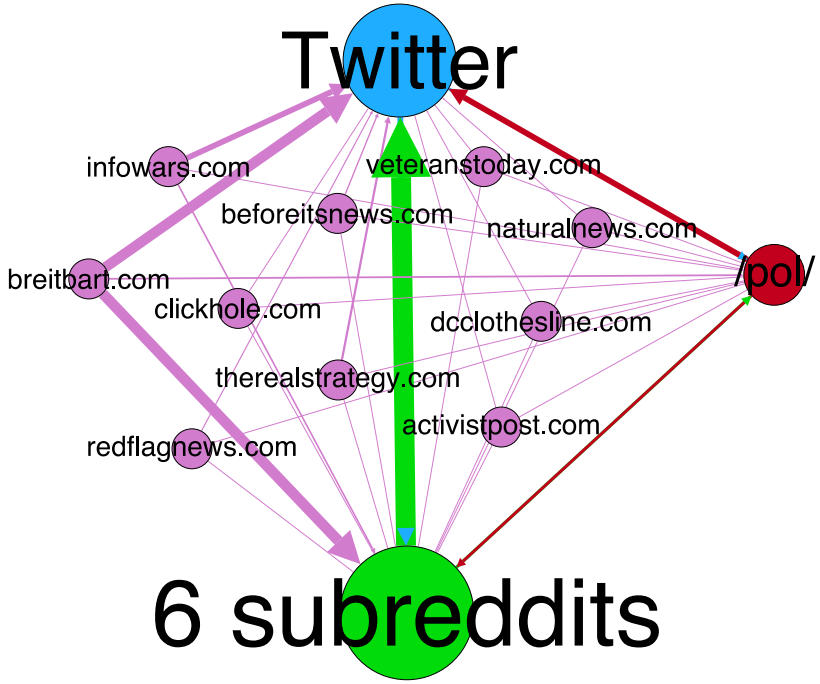


First study on how disinformation spreads among Web communities

- Looked into fringe web communities, such as 4chan/pol, reddit/The_Donald, gab
- Innovative use of Hawkes processes to infer causal relationships between Web posts



Mainstream News Sources



Alternative News Sources

Publication Highlights

- 39 scientific publications
- Published in top tier venues, e.g.,
  - ACM IMC
  - ACM TWEB
  - ACM CSCW
  - AAAI ICWSM
  - ACM CCS
- Best Papers: 4

"Kek, Cucks, And God Emperor Trump: A Measurement Study Of 4chan's Politically Incorrect Forum And Its Effects On The Web". Gabriel Emile Hine, Jeremiah Onaolapo, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Riginos Samaras, Gianluca Stringhini And Jeremy Blackburn. International AAAI Conference On Web And Social Media (ICWSM), 2017.
In The News: Nature.com, La Stampa, Motherboard Vice, Paideia News, The Conversation, Repubblica, VICE, Sky News, Welt.
Awards: Best Paper Runner Up.

"On The Origins Of Memes By Means Of Fringe Web Communities". Savvas Zannettou, Tristan Caulfield, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, And Guillermo Suarez-Tangil. ACM Internet Measurement Conference (IMC), 2018.
In The News: QUARTZ, Business Insider, Nerdcore, MIT Technology Review, VICE, InfoWars, Dailydot, Indy100, Thinkprogress, Iflscience, Hornet, Washington Post, BBC News, The Conversation, Rappler, El Pais, MotherJones, Raw Story, Bellingcat, Infowars, I-D Vice.
Awards: Distinguished Paper.

Publication Highlights

"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 And Jeremy Blackburn. Workshop On Computational Methods In CyberSafety, Online Harassment And Misinformation, 2019. Long Version

In The News: The Register, The Conversation, Houston Chronicle, US News.

Policy: Regulating Disinformation With Artificial Intelligence, EU Parliament Research Service.

Awards: Best Paper.

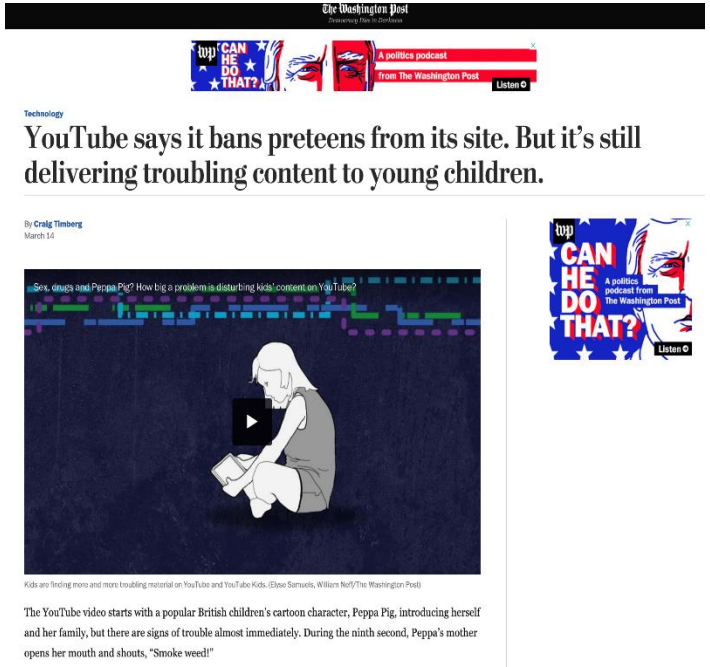
"Detecting Cyberbullying And Cyberaggression In Social Media". Despoina Chatzakou, Ilias Leontiadis, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, Athena Vakali, Nicolas Kourtellis. ACM Transactions on the Web (TWEB), 2019.

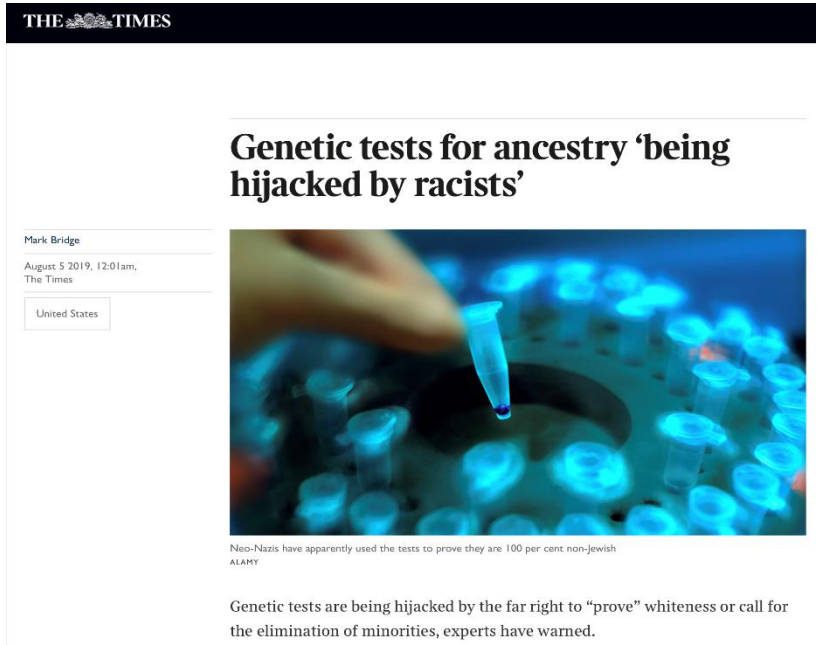
In The News: The Register, BBC, NewsBusters, Cyprus Mail, Sky News.

Communication Highlights

- Our work was covered by The New York Times, Washington Post, BBC, and many others

- 69 international/mainstream press mentions





Communication Highlights

RADIO SHOW NEWS VIDEOS STORE BREAKING NEWS CONTACT

NEW SCIENTIFIC STUDY PROVES THE LEFT CAN'T MEME

Most effective memes originate on 4chan and The Donald subreddit

Paul Joseph Watson | Infowars.com - JUNE 13, 2018

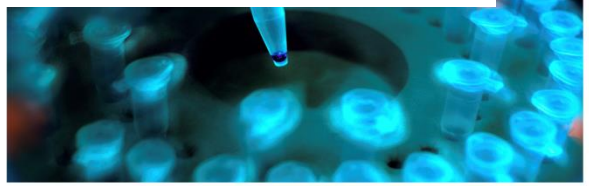
Technology

YouT
delive

By Craig Timberg
March 14

Kids are finding more and more troubling material on YouTube and YouTube Kids. (Elyse Samuels, William Neff/The Washington Post)

The YouTube video starts with a popular British children's cartoon character, Peppa Pig, introducing herself and her family, but there are signs of trouble almost immediately. During the ninth second, Peppa's mother opens her mouth and shouts, "Smoke weed!"



Neo-Nazis have apparently used the tests to prove they are 100 per cent non-Jewish
ALAMY

Genetic tests are being hijacked by the far right to "prove" whiteness or call for the elimination of minorities, experts have warned.

Impact on Policy Makers and Regulators

- A **European Parliament Research Service** report cites 13 ENCASE publications
- A **UK Parliamentary committee** discussed our *Disturbed YouTube for Kids* study
- The **Canadian Security Intelligence Service** published a report referring to our *Web Centipede* study

UK Parliament

HOUSE OF COMMONS

European Parliament

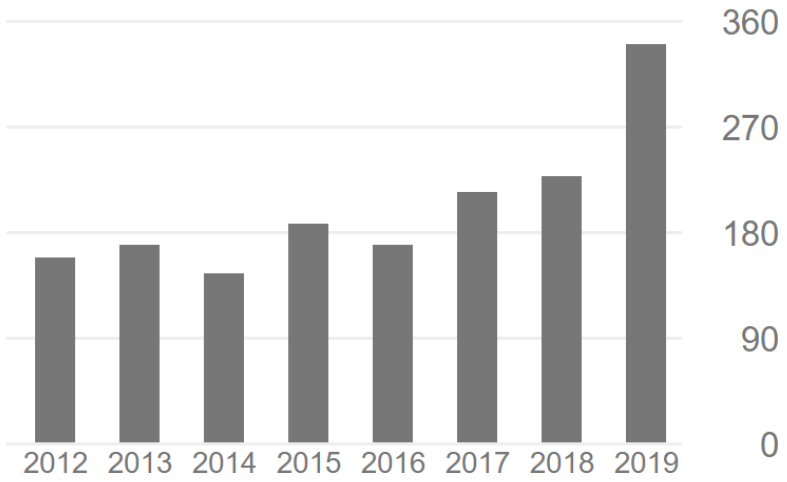
Canadian Security Intelligence Service

Impact on our Careers

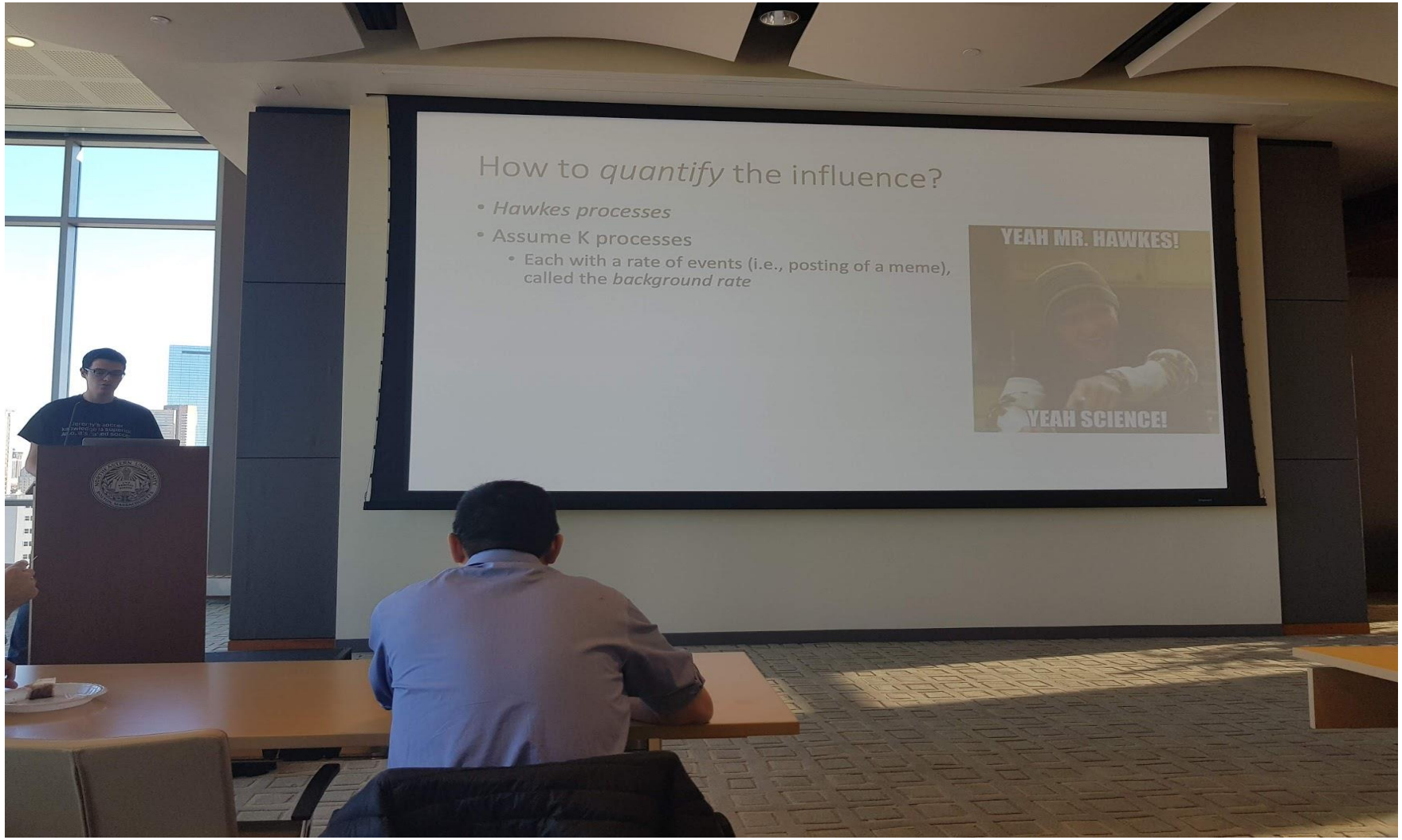
- Largest R&I project I have coordinated

- Savvas Zannettou, now at MPI-SWS. Kostantinos Papadamou graduates next year

Citations	2169	1288
h-index	19	17
i10-index	25	24



Impact on our Careers



Impact on the communities we studied

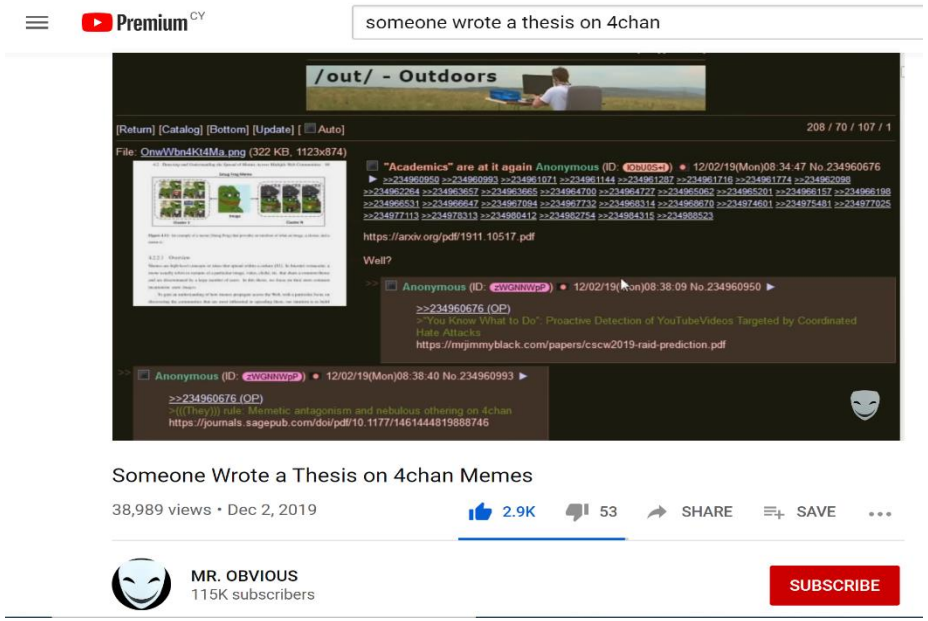
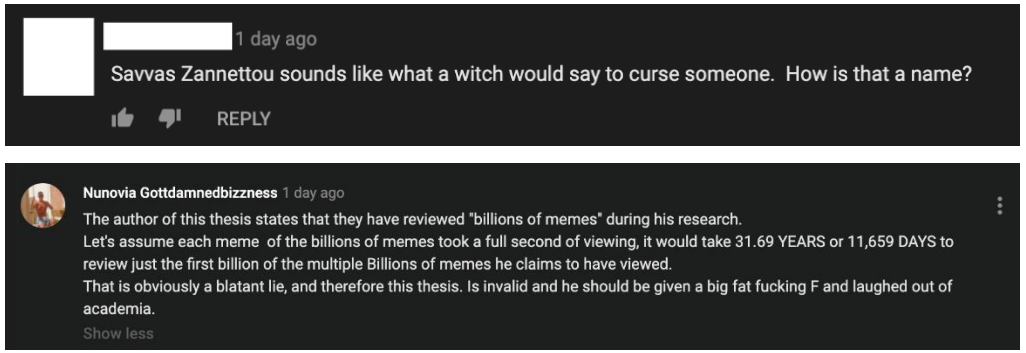


Figure 29: A very comfy Savvas Zannettou



Thank you!

https://encase.socialcomputing.eu