

# Savvas Zannettou

Jaffalaan 5, 2628 BX  
Delft, Netherlands  
✉ [s.zannettou@tudelft.nl](mailto:s.zannettou@tudelft.nl)

---

## Experience

- 01/2022 – current **Assistant Professor**, Faculty of Technology, Policy and Management, Delft University of Technology.
- 11/2019 – 11/2021 **Postdoctoral Researcher**, Max Planck Institute for Informatics.
- 08/2018 – 02/2019 **Visiting Scholar**, University of Alabama at Birmingham.
- 01/2017 – 07/2017 **Research Intern**, Telefonica I+D, Barcelona.
- 06/2014 – 12/2014 **Research Intern**, NEC Labs Europe.

---

## Education

- 01/2016 – 10/2019 **PhD in Computer Science**, Cyprus University of Technology.
- 07/2014 – 12/2015 **MSc in Computer Science**, Cyprus University of Technology.
- 09/2010 – 07/2014 **BSc in Computer Science**, Cyprus University of Technology.

---

## Awards

- 06/2023 Best Senior Program Committee Member Award at AAAI ICWSM 2023 ([link](#)).
- 11/2022 Best Paper Honorable Mention Award at ACM CCS 2022 ([link](#)).
- 06/2022 Best Senior Program Committee Member Award at AAAI ICWSM 2022 ([link](#)).
- 10/2021 Best Paper Honorable Mention Award at CSCW 2021 ([link](#)).
- 06/2021 Best Reviewer Award at AAAI ICWSM 2021([link](#)).
- 06/2020 2x Best Paper Honorable Mention Awards at AAAI ICWSM 2020 ([link](#)).
- 06/2020 Best Reviewer Award at AAAI ICWSM 2020 ([link](#)).
- 09/2019 Contributor in the MPTCP Linux Kernel implementation, which received the ACM SIGCOMM Networking Systems Award ([link](#)).
- 05/2019 Best Paper Award, Cybersafety Workshop at the Web Conference 2019 ([link](#)).
- 10/2018 Distinguished Paper Award, ACM Internet Measurement Conference 2018 (IMC '18, [link](#)).

---

## Selected Publications

For a full list of my publications see my [Google Scholar profile](#): 61 publications with 5,007 citations and h-index 30 (as of February 2024)

### Conferences

- [1] **Savvas Zannettou**, Olivia Nemes-Nemeth, Oshrat Ayalon, Angelica Goetzen, Krishna P. Gummadi, Elissa Redmiles, and Franziska Roesner. Analyzing User Engagement with TikTok's Short Format Video Recommendations using Data Donations. In *CHI Conference on Human Factors in Computing Systems (CHI)*, 2024.
- [2] Salim Chouaki, Abhijnan Chakraborty, Oana Goga, and **Savvas Zannettou**. What News Do People Get on Social Media? Analyzing Exposure and Consumption of News through Data Donations. In *The WebConf*, 2024.

- [3] Karan Vombatkere, Sepehr Mousavi, **Savvas Zannettou**, Franziska Roesner, and Krishna P. Gummadi. TikTok and the Art of Personalization: Investigating Exploration and Exploitation on Social Media Feeds. In *The WebConf*, 2024.
- [4] Xinlei He, **Savvas Zannettou**, Yun Shen, and Yang Zhang. You Only Prompt Once: On the Capabilities of Prompt Learning on Large Language Models to Tackle Toxic Content. In *IEEE Symposium on Security and Privacy (S&P)*, 2024.
- [5] Yiting Qu, Xinyue Shen, Xinlei He, Michael Backes, **Savvas Zannettou**, and Yang Zhang. Unsafe diffusion: On the generation of unsafe images and hateful memes from text-to-image models. In *ACM SIGSAC Conference on Computer and Communications Security (CCS)*, 2023.
- [6] Yiting Qu, Xinlei He, Shannon Pierson, Michael Backes, Yang Zhang, and **Savvas Zannettou**. On the Evolution of (Hateful) Memes by Means of Multimodal Contrastive Learning. In *IEEE Symposium on Security and Privacy (S&P)*, 2023.
- [7] Pujan Paudel, Jeremy Blackburn, Emiliano De Cristofaro, **Savvas Zannettou**, and Gianluca Stringhini. LAMBRETTA: Learning to Rank for Twitter Soft Moderation. In *IEEE Symposium on Security and Privacy (S&P)*, 2023.
- [8] Alexandros Efstratiou, Jeremy Blackburn, Tristan Caulfield, Gianluca Stringhini, **Savvas Zannettou**, and Emiliano De Cristofaro. Non-Polar Opposites: Analyzing the Relationship Between Echo Chambers and Hostile Intergroup Interactions on Reddit. In *Proceedings of the International AAAI Conference on Web and Social Media, ICWSM '23*, 2023.
- [9] Felipe González-Pizarro and **Savvas Zannettou**. Understanding and Detecting Hateful Content using Contrastive Learning. In *Proceedings of the International AAAI Conference on Web and Social Media, ICWSM '23*, 2023.
- [10] Chen Ling, Krishna P Gummadi, and **Savvas Zannettou**. "Learn the Facts About COVID-19": Analyzing the Use of Warning Labels on TikTok Videos. In *Proceedings of the International AAAI Conference on Web and Social Media, ICWSM '23*, 2023.
- [11] Wai Man Si, Michael Backes, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, **Savvas Zannettou**, and Yang Zhang. Why So Toxic? Measuring and Triggering Toxic Behavior in Open-Domain Chatbots. In *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (Best Paper Honorable Mention)*, 2022.
- [12] Xinyue Shen, Xinlei He, Michael Backes, Jeremy Blackburn, **Savvas Zannettou**, and Yang Zhang. On Xing Tian and the Perseverance of Anti-China Sentiment Online. In *Proceedings of the International AAAI Conference on Web and Social Media, ICWSM '22*, 2022.
- [13] Kostantinos Papadamou, **Savvas Zannettou**, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Michael Sirivianos. "it is just a flu": Assessing the effect of watch history on youtube's pseudoscientific video recommendations. In *Proceedings of the International AAAI Conference on Web and Social Media, ICWSM '22*, 2022.
- [14] Max Aliapoulos, Antonis Papasavva, Cameron Ballard, Emiliano De Cristofaro, Gianluca Stringhini, **Savvas Zannettou**, and Jeremy Blackburn. The gospel according to Q: Understanding the QAnon conspiracy from the perspective of canonical information. In *Proceedings of the International AAAI Conference on Web and Social Media, ICWSM '22*, 2022.

- [15] Mohammad Hammas Saeed, Shiza Ali, Jeremy Blackburn, Emiliano De Cristofaro, **Savvas Zannettou**, and Gianluca Stringhini. TROLLMAGNIFIER: Detecting State-Sponsored Troll Accounts on Reddit. In *IEEE Symposium on Security and Privacy (S&P)*, 2022.
- [16] **Savvas Zannettou**. "I Won the Election!": An Empirical Analysis of Soft Moderation Interventions on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '21, Virtual Event, 2021.
- [17] Manoel Horta Ribeiro, Shagun Jhaver, **Savvas Zannettou**, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Robert West. Does Platform Migration Compromise Content Moderation? Evidence from r/The\_Donald and r/Incels. In *24th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, Virtual Event, 2021.
- [18] Fatemeh Tahmasbi, Leonard Schild, Chen Ling, Jeremy Blackburn, Gianluca Stringhini, Yang Zhang, and **Savvas Zannettou**. "Go eat a bat, Chang!": On the Emergence of Sinophobic Behavior on Web Communities in the Face of COVID-19. In *The Web Conference*, Virtual Event, 2021.
- [19] Antonis Papasavva, Jeremy Blackburn, Gianluca Stringhini, **Savvas Zannettou**, and Emiliano De Cristofaro. The Great Awakening: An Exploratory Study of QAnon on Voat. In *The Web Conference*, Virtual Event, 2021.
- [20] Chen Ling, Ihab AbuHilal, Jeremy Blackburn, Emiliano De Cristofaro, **Savvas Zannettou**, and Gianluca Stringhini. Dissecting the Meme Magic: Understanding Indicators of Virality in Image Memes. In *24th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, Virtual Event, 2021.
- [21] Yuping Wang, Fatemeh Tamahsbi, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, David Magerman, **Savvas Zannettou**, and Gianluca Stringhini. Understanding the Use of Fauxtography on Social Media. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '21, Virtual Event, 2021.
- [22] Manoel Horta Ribeiro, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, Gianluca Stringhini, Summer Long, Stephanie Greenberg, and **Savvas Zannettou**. The Evolution of the Manosphere Across the Web. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '21, Virtual Event, 2021.
- [23] Mohamad Hoseini, Philippe Melo, Manoel Júnior, Fabrício Benevenuto, Balakrishnan Chandrasekaran, Anja Feldmann, and **Savvas Zannettou**. Demystifying the Messaging Platforms' Ecosystem Through the Lens of Twitter. In *Proceedings of the ACM Internet Measurement Conference*, IMC '20, Virtual Event, 2020.
- [24] **Savvas Zannettou**, Mai ElSherief, Elizabeth Belding, Shirin Nilizadeh, and Gianluca Stringhini. Measuring and Characterizing Hate Speech on News Websites. In *Proceedings of the ACM Conference on Web Science*, WebSci '20, Virtual Event, 2020.
- [25] **Savvas Zannettou**, Tristan Caulfield, Barry Bradlyn, Emiliano De Cristofaro, Gianluca Stringhini, and Jeremy Blackburn. Characterizing the Use of Images in State-Sponsored Information Warfare Operations by Russian Trolls on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '20, Virtual Event, 2020.
- [26] **Savvas Zannettou**, Joel Finkelstein, Barry Bradlyn, and Jeremy Blackburn. A Quantitative Approach to Understanding Online Antisemitism. In *Proceedings of the*

*International AAAI Conference on Web and Social Media (Best Paper Honorable Mention)*, ICWSM '20, Virtual Event, 2020.

- [27] Alexandros Mittos, **Savvas Zannettou**, Jeremy Blackburn, and Emiliano De Cristofaro. "And We Will Fight For Our Race!" A Measurement Study of Genetic Testing Conversations on Reddit and 4chan. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '20, Virtual Event, 2020.
- [28] Kostantinos Papadamou, Antonis Papasavva, **Savvas Zannettou**, Jeremy Blackburn, Nicolas Kourtellis, Ilias Leontiadis, Gianluca Stringhini, and Michael Sirivianos. Disturbed youtube for kids: Characterizing and detecting inappropriate videos targeting young children. In *Proceedings of the International AAAI Conference on Web and Social Media (Best Paper Honorable Mention)*, ICWSM '20, Virtual Event, 2020.
- [29] Jason Baumgartner, **Savvas Zannettou**, Brian Keegan, Megan Squire, and Jeremy Blackburn. The Pushshift Reddit Dataset. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '20, Virtual Event, 2020.
- [30] Jason Baumgartner, **Savvas Zannettou**, Megan Squire, and Jeremy Blackburn. The pushshift telegram dataset. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '20, Virtual Event, 2020.
- [31] Antonis Papasavva, **Savvas Zannettou**, Emiliano De Cristofaro, Gianluca Stringhini, and Jeremy Blackburn. Raiders of the lost kek: 3.5 years of augmented 4chan posts from the politically incorrect board. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '20, Virtual Event, 2020.
- [32] **Savvas Zannettou**, Tristan Caulfield, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, and Guillermo Suarez-Tangil. On the Origins of Memes by Means of Fringe Web Communities. In *Proceedings of the Internet Measurement Conference (Distinguished Paper Award)*, IMC '18, Boston, MA, USA, 2018.
- [33] **Savvas Zannettou**, Tristan Caulfield, William Setzer, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. Who Let The Trolls Out? Towards Understanding State-Sponsored Trolls. In *Proceedings of the ACM Conference on Web Science*, WebSci '19, Boston, MA, USA, 2019.
- [34] **Savvas Zannettou**, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, and Gianluca Stringhini. Understanding Web Archiving Services and Their (Mis)Use on Social Media. In *Proceedings of the International AAAI Conference on Web and Social Media*, ICWSM '18, Stanford, CA, USA, 2018.
- [35] **Savvas Zannettou**, Tristan Caulfield, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. The Web Centipede: Understanding How Web Communities Influence Each Other Through the Lens of Mainstream and Alternative News Sources. In *Proceedings of the Internet Measurement Conference*, IMC '17, London, UK, 2017.
- [36] **Savvas Zannettou**, Michael Sirivianos, and Fragkiskos Papadopoulos. Exploiting path diversity in datacenters using mptcp-aware sdn. In *Proceedings of the IEEE Symposium on Computers and Communication*, ISCC' 16, Messina, Italy, 2016.

#### Journals

- [37] Alexandros Mittos, **Savvas Zannettou**, Jeremy Blackburn, and Emiliano De Cristofaro. Analyzing Genetic Testing Discourse on the Web Through the Lens of Twitter, Reddit, and 4chan. *ACM Transactions of the Web*, 1(1), 2020.

- [38] **Savvas Zannettou**, Michael Sirivianos, Jeremy Blackburn, and Nicolas Kourtellis. The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans. *Journal of Data and Information Quality (JDIQ)*, 11(3):10, 2019.
- Workshops**
- [39] **Savvas Zannettou**, Tristan Caulfield, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, and Jeremy Blackburn. Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web. In *Proceedings of the Web Conference Companion (Best Paper Award)*, WWW '19 Companion, San Francisco, CA, USA, 2019.
- [40] **Savvas Zannettou**, Barry Bradlyn, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, Haewoon Kwak, and Jeremy Blackburn. What is Gab? A Bastion of Free Speech or an Alt-Right Echo Chamber? In *Proceedings of the Web Conference Companion*, WWW '18 Companion, Lyon, France, 2018.
- [41] **Savvas Zannettou**, Sotirios Chatzis, Kostantinos Papadamou, and Michael Sirivianos. The Good, the Bad and the Bait: Detecting and Characterizing Clickbait on YouTube. In *IEEE Security and Privacy Workshops*, SPW '18, San Francisco, CA, USA, 2018. IEEE.

## Selected Press Coverage

- 2021 The Atlantic covered my work on soft moderation interventions on Twitter (<https://www.theatlantic.com/technology/archive/2021/03/twitters-birdwatch-aims-to-crowdsource-fact-checking/618187/>).
- 2021 The Washington Post covered our recent preprint studying the effect of community-based moderation interventions (<https://wapo.st/3cbiwAx>).
- 2020 The Washington Post covered our paper investigating the rise of Sinophobic behavior on Web communities in the face of the COVID-19 pandemic (<https://wapo.st/2UpEs2j>).
- 2019 The New York Times wrote an article mentioning our work on disturbing YouTube videos targeting young children (<https://nyti.ms/36zDgiF>).
- 2019 The Washington Post covered our work on detecting disturbing YouTube videos targeting young children (<https://wapo.st/3pBH1V>).
- 2018 The Washington Post covered (online and in print on the front page of the Business & Economy Section) our paper on a quantitative approach to understanding antisemitism online [26] (<https://wapo.st/3kEGSoN>).
- 2018 MIT Technology Review wrote an article about our paper on measuring memes [32] (<https://www.technologyreview.com/s/611332/this-is-where-internet-memes-come-from>).
- 2018 New Scientist covered our work characterizing Gab [40] (<https://www.newscientist.com/article/2161933-theres-an-alt-right-alt-twitter-and-its-filled-with-hate/>).
- 2017 The Atlantic wrote an article concerning online extremist communities, focusing on our findings from [35] showing how they have a disproportionate impact on mainstream social media. An edited version of the print article is available online at <https://www.theatlantic.com/politics/archive/2017/11/how-much-attention-should-extremists-get/546152/>.

## Service

- 2024 **Senior Program Committee Member**, AAAI ICWSM 2024

- 2024 **Program Committee Member**, The Web Conference 2024 (Web & Society Track), Usenix Security 2024, ICML 2024, ICLR 2024, ACM CCS 2024
- 2023 **Senior Program Committee Member**, AAAI ICWSM 2023
- 2023 **Program Committee Member**, The Web Conference 2023 (Web & Society Track), Usenix Security 2023, WSDM 2023, ICML 2023, ACM CCS 2023, IC2S2 2023
- 2022 **Senior Program Committee Member**, AAAI ICWSM 2022
- 2022 **External Reviewer**, NSF, Nature, WWTF, and NeurIPS 2022
- 2021 **Program Committee Member**, The Web Conference 2022, Web & Society Track
- 2021 **Guest Editor**, Special issue at IEEE Internet Computing
- 2021 **Data Chair**, AAAI ICWSM 2021
- 2021 **Program Committee Member**, AAAI ICWSM 2021
- 2021 **External Reviewer**, ACM CSCW
- 2020 **Program Committee Member**, AAAI ICWSM 2020
- 2020 **External Reviewer**, ACM CSCW and ACM CHI
- 2019 **External Reviewer**, AAAI ICWSM 2019

## Teaching

- 2023 **Responsible Lecturer**, TI3160TU: Natural Language Processing, Delft University of Technology
- 2021 **Co-Lecturer**, Data Networks Course, Saarland University (with Anja Feldmann, Oliver Gasser, and Devashish Gosain)
- 2021 **Co-Instructor**, Seminar on Data-driven Understanding of the Disinformation Epidemic, Saarland University (with Yang Zhang)
- 2020 **Co-Instructor**, Seminar on Data-driven Approaches on Understanding Disinformation, Saarland University (with Yang Zhang)
- 2016–2019 **Teaching Assistant**, Course on Computer Networks and Communications, Cyprus University of Technology
- 2016–2019 **Teaching Assistant**, Course on Advanced and Distributed Operating Systems, Cyprus University of Technology

## Invited Talks

- 2023 **Understanding and Mitigating Online Harms using AI**, Federal University of Minas Gerais (Belo Horizonte, Brazil)
- 2022 **Understanding and Detecting Hateful Content using Contrastive Learning**, CYENS Centre of Excellence (Nicosia, Cyprus)
- 2022 **Towards Understanding Soft Moderation Interventions on the Web**, University College London (London, UK)
- 2022 **Understanding Online Deception from State-Sponsored Trolls in Social Media Platforms**, Workshop on Active Defense and Deception at IEEE Euro S&P 2022 (Genoa, Italy)
- 2022 **Content Moderation Systems for Mitigating the Spread of False Information on the Web**, University of Amsterdam (Amsterdam, Netherlands)
- 2021 **Towards Understanding Moderation Interventions on the Web**, 74th Dies Natalis Conference (Groningen, Netherlands)
- 2019 **On the Origins of Memes by Means of Fringe Web communities**, University of Southern California, Information Sciences Institute (Los Angeles, USA).

- 2019 **Towards Understanding the Behavior of State-Sponsored Trolls and their Influence on the Web**, University College London (London, UK).
- 2019 **On the Origins of Memes by Means of Fringe Web communities**, ACM Sigmetrics (Arizona, USA).
- 2018 **Towards Understanding Small Fringe Web Communities and their Influence on the Web's Information Ecosystem**, Eurecat (Barcelona, Spain).

---

## References

- **Prof. Krishna P. Gummadi**  
(gummadi@mpi-sws.org)
- **Dr. Michael Sirivianos**  
(michael.sirivianos@cut.ac.cy)
- **Prof. Emiliano De Cristofaro**  
(e.decrisofaro@ucl.ac.uk)
- **Dr. Gianluca Stringhini**  
(gian@bu.edu)