

# AGENDA

## Thursday, September 28, 2023

7:30 – 8:30	Registration and Continental Breakfast
8:30 – 8:45	Opening Remarks (McCaw Hall)  Jeffrey T. Hancock, Harry and Norman Chandler Professor of Communication, Stanford University  Faculty Director, Stanford Internet Observatory and Social Media Lab
8:45 – 9:15	Reflecting on the State of "Misinformation" Research During a Turbulent Time Kate Starbird, Associate Professor, Human Centered Design & Engineering, University of Washington
9:15 – 9:30	Break
9:30 - 10:30	Select one of the below, simultaneously held, sessions to attend.

Panel: Mental Health and Al	Lightning Talks: Policy Proposals	Workshop: Transparency Reporting with Ofcom
McCaw Hall Mainstage	Fisher Conference Center Lane/Lyons/Lodato	Fisher Conference Center Barnes/McDowell/Cranston

#### Mental Health and Artificial Intelligence (McCaw Hall Mainstage)

Chaired by the Stanford Institute for Human-Centered Artificial Intelligence (HAI), this panel will explore how applying AI to mental healthcare could expand access to underserved communities and reduce costs, and what challenges need to be overcome before realizing these benefits.

Moderated by Ahmad Rushdi, Stanford HAI

- Johannes Eichstaedt, Stanford University
- Diyi Yang, Stanford University
- o Ravi Iyer, University of Southern California
- Rob Morris, Koko

Mainstage sessions will be recorded. Research sessions, panels and lightning talks are on the record. Workshops are off the record. Session descriptions and a list of non-presenting authors are on the online agenda.

**Lightning Talks: Policy Proposals** (Fisher Conference Center Lane/Lyons/Lodato) Lightning talks feature five-minute, rapid fire presentations with time for questions. *Moderated by Elena Cryst, Stanford Internet Observatory* 

- Human + Al Processes for Trust & Safety: Human in the Loop or Human on Appeal?
   Theodoros Evgeniou, INSEAD
- Frequency of Pro-Suicide Website Searches Associated with Death by Suicide in the U.S.
   Kerry Keyes, Columbia University
- In Devices We Trust User Survey on Factors that are Important for Trusting Smart Devices and Influencing Technology Purchase Decisions
   Narsi Gangula, Google
- Designing for Prosociality A Framework and Review of Digital Interventions for Promoting Online Prosocial Behavior
   David Grüning, Prosocial Design Network; Heidelberg University; GESIS – Leibniz Institute for the
- Social Sciences
   Embedding Societal Values into Social Media Algorithms
   Michael S. Bernstein, Stanford University
- French Risk Managers Facing Information Warfare
   Caroline Rabourdin, Pulpe de Com
- Protecting Democracy Online in 2024 and Beyond Megan Shahi, Center for American Progress
- Is the Notion of Systemic Risk Adapted to the Challenges of Technology Regulation?
   Florence G'Sell, Sciences Po and Stanford Cyber Policy Center

**Transparency Reporting with Ofcom** (Fisher Conference Center Barnes/McDowell/Cranston)
This workshop will bring together Trust & Safety practitioners, academics, researchers and members of civil society in a conversation around transparency reporting and meaningful metrics. The discussions will inform Ofcom's development of its transparency requirements for platforms, allowing key stakeholders to provide input early into Ofcom's regulatory strategy. The workshop will feature a brainstorming session on meaningful metrics for different kinds of online platforms, a discussion of how transparency reporting can best benefit different audiences, and an analysis of lessons learned from platforms' first transparency reports provided under the European Union's Digital Services Act. *Facilitated by Eleanor Simmance and Sheryl Wilson, Ofcom* 

10:30 – 11:00 **Break** 

11:00 – 12:30 Select one of the below, simultaneously held, sessions to attend.

**Lightning Talks: Trust & Safety in Search**McCaw Hall Mainstage

**Research Session: Policy Proposals in the Public Sector**Fisher Conference Center
Lane/Lyons/Lodato

Workshop: Instagram: Exploring Tradeoffs in Ranking Algorithms Fisher Conference Center Barnes/McDowell/Cranston

#### **Trust & Safety in Search** (McCaw Hall Mainstage)

A special lightning talk session that looks at the Trust & Safety issues unique to search products and novel ways their harms can be reduced.

Moderated by Ronald Robertson, Stanford Internet Observatory

- Evaluating Online Legal Information on SERPs and LLMs: Looking Beyond Generic Content Margaret Hagan, Stanford Law School
- Safeguarding Without Sanitizing Search: Modular Policies, Defaults and User Controls Rafael Burde, Google
- Sowing "Seeds of Doubt": A Study of Election and Medical Misinformation Trust in the United States and Brazil
   Beth Goldberg, Jigsaw
- Market or Markets? Investigating Google Search's Market Shares Under Vertical Segmentation
   Desheng Hu, University of Zurich and Northeastern University
- Explaining Country-Level Variation in Governments' Takedown Requests to Google, 2011-2022
   Zarine Kharazian, University of Washington
- All About Trust: Measurement Challenges and Possibilities
   Sujata Mukherjee, Google
- Safe and Trusted Ads Ecosystem by Preventing Invalid Traffic Puneet Sangal, Google
- A Comparative Analysis of the Prevalence of Disinformation across Major Social Media Platforms in Poland, Slovakia, and Spain Kashyap Puranik, TrustLab

#### Policy Proposals in the Public Sector (Fisher Conference Center Lane/Lyons/Lodato)

Five studies on the role of the public sector in online safety.

Moderated by Justin Hendrix, Tech Policy Press

- Measurement of Age Assurance Technologies
   George Billinge, Ofcom; Wei Lynn Ng, UK's Information Commissioner's Office
- When Seconds Matter: Harnessing Data to Safeguard Student Life Kathleen Boehle, Securly
- Securing Online Spaces: A Comparative Overview of Singapore's New Online Safety Legislation and Operational Insights
  - Makoto Hong, Singapore Infocomm Media Development Authority
- Combating Non-Consensual Intimate Imagery in the Caribbean: Overcoming Legislative Gaps with Creative Approaches
  - Desiree Valentine and Michelle d'Auvergne, Eastern Caribbean Supreme Court
- Artificial Intelligence Malpractice
   Chinmayi Sharma, Fordham Law School

#### **Instagram: Exploring Tradeoffs in Ranking Algorithms**

(Fisher Conference Center Barnes/McDowell/Cranston)

Social media platforms use ranking algorithms to distribute online content and help people discover information. In light of the growing public debate on the role of algorithmic amplification in society, this workshop is designed to break down the "black box" around how ranking works on Instagram and facilitate a dialogue within the trust and safety community about its current approach to ranking and ways to improve transparency and user agency. The workshop will feature a brainstorming session about pros and cons of different ranking methods, facilitate a candid exchange about the challenges and tradeoffs platforms face, and provide an avenue to inform industry approaches. Facilitated by Sarah Shirazyan, Instagram / Stanford Law School; Liz Arcamona and Yvonne Lee, Instagram.

12:30 – 1:30 **Lunch** (Ford Gardens)

1:30 – 2:30 Select one of the below, simultaneously held, sessions to attend.

Panel: Evaluating Digital Literacy Interventions Across Platforms

McCaw Hall Mainstage

Lightning Talks: Trust & Safety Tooling Fisher Conference Center Lane/Lyons/Lodato

#### **Evaluating Digital Literacy Interventions Across Platforms** (McCaw Hall Mainstage)

Moderated by Jeffrey T. Hancock, Stanford University

- Beth Goldberg, Jigsaw
- Sander Van der Linden, University of Cambridge
- Melanie Welch, American Library Association
- Sindy Braun, Stanford University

#### Trust & Safety Tooling (Fisher Conference Center Lane/Lyons/Lodato)

Lightning talks feature five-minute, rapid fire presentations with time for questions.

Moderated by Naomi Shiffman, Oversight Board

- Empowering Internet Users Through Source Credibility Ratings
   Sarah Brandt, NewsGuard
- Auditing Tools for AI Regulations Alisar Mustafa, CivicSync
- Preventing Tech Facilitated Gender Based Violence NCII Abuse Larry Magid, SWGfL; stopncii.org
- Governance for End-to-End Encrypted Communities
   Armin Namavari, Cornell University
- Insights from The Global Disinformation Policy Database
   Ryan Williams, Global Disinformation Lab at University of Texas

- Towards Better Automated Content Moderation in Low-Resource Languages
   Gabriel Nicholas, Center for Democracy & Technology
- Discovering Patterns and Trends of Online Child Sexual Abuse
   Rahul Krishna Gajula, Technological University of Dublin

#### 2:30 - 3:00 **Break**

3:00 – 4:00 Select one of the below, simultaneously held, sessions to attend.

# **Lightning Talks: Content Moderation** and **Detection**

McCaw Hall Mainstage

### Research Session: Al and Trust & Safety Fisher Conference Center Lane/Lyons/Lodato

#### **Content Moderation and Detection** (McCaw Hall Mainstage)

Lightning talks feature five-minute, rapid fire presentations with time for questions.

Moderated by Marietje Schaake, Stanford Cyber Policy Center

- A Scalable Framework for Mitigating Harm on Social Media Jennifer Allen, MIT Sloan School of Management
- Future Challenges for Online, Crowd-sourced Content Moderation: Evidence from Twitter's Community Notes
  - Sharanya Majumder, Brookings Institution
- No-Media CSAM Networks Operating Inside Social Media: Recognition Phase Carolina Christofoletti, University of S\u00e4o Paulo
- The Rise in Online Sexual Extortion Schemes Targeting Minors: The Cross-Platform Complexity,
   Drivers, and Characteristics
  - Avi Jager, ActiveFence
- Gender Blind Spots in Content Moderation of the Far-Right Eviane Leidig, Tilburg University
- Now We Got Bad Blood: Exploring Harassment Dynamics Within the Online Taylor Swift Fandom
   Cristina López G., Graphika
- Exploring the Use of LLMs in Moderating Mental Health Content on Social Media: Insights from BlueFever
  - Karen Mosoyan, Blue Fever
- Challenging "Problematic" Internet Usage Ashley Marie Walker, Google
- Toward a Common Baseline Understanding of Trust and Safety Terminology
   David Sullivan, Digital Trust and Safety Partnership
- Content Moderator Coping Strategies: A Qualitative Study Ruth Spence, Middlesex University

#### Artificial Intelligence and Trust & Safety: New Risks, New Solutions

(Fisher Conference Center Lane/Lyons/Lodato)

Five studies examining how artificial intelligence may mitigate or exacerbate harms.

Moderated by David Thiel, Stanford Internet Observatory

- Hidden Attributes of Human and LLM Conversationalists: Implications for Diversity in Society Nathanael Fast, University of Southern California
- Filtering Technologies and the Fairness of Natural Language Systems
   Eddie Yang, University of California San Diego
- Generative AI and Its Potential Intersections with Online CSA Tim O'Gorman, Thorn
- Trust and Safety Issues Arising from the Exclusion of African Datasets in AI Training Models
   Arthur Gwagwa, Utrecht University Ethics Institute
- AI Technology and the Impact to Kids' Mental Health Courtney Gregoire, Microsoft; Candice Odgers, UC Irvine

#### 4:00 – 5:30 **Happy Hour and Poster Session** (Ford Gardens)

During happy hour, poster presenters will be on hand to discuss their research.

- The Anatomy of [In]Direct Swarming and Its Impact on Democracies
   Eirliani Abdul Rahman, Harvard University
- From Flowers to Fascism: The Cottagecore to Tradwife Pipeline on Tumblr Oliver Allen, Northeastern University
- The Unseen A+ Student: Evaluating the Performance and Detectability of Generative AI Models in High School Assignments
   Matyas Bohacek, Stanford University
- AIPasta: Modelling & Evaluating Threats of Language Models in Multilingual CopyPasta Campaigns Saloni Dash, University of Washington
- Not So Black and White: Perceptions of Racism on Social Media Maggie Harrington and Jennifer Eberhardt, Stanford University
- Loaded Language and Conspiracy Theorizing Emily Klein, Rensselaer Polytechnic Institute
- The Impact of Platform Architecture Characteristics and Moral Values on Polarization in Social Media: The Context of Vigilantism
  - Shalini Kapali Kurumathur, The University of Texas at San Antonio
- MediaWise en Español: A Longitudinal Digital Media Literacy Intervention to Improve Misinformation Discernment in English and Spanish
   Angela Lee, Stanford University
- The Algorithmic Amplifier: A Study on TikTok, Adolescents, and Misinformation Dissemination Sunny Liu, Stanford University
- Ranking the Best of Trust and Safety Practices
   Shubhi Mather, UC Berkeley

- Effects of Online Ideological Group Rhetoric on Perceptions of Credibility
   Ivan Ramirez, University of Oklahoma
- Meta-problems: Government Access to Data in the Metaverse Simon Sun, Indiana University, Maurer School of Law
- Defining Effective Data Privacy Legislation
   Darshan Vijaykumar, Duke University
- Eyes of the Street: Surveillance Urbanism from Above and Below in Prospect Park and Surrounding Neighborhoods
   Mia Winther-Tamaki, Columbia University

#### 5:00 – 6:00 **Moderated Content Live!** (Fisher Conference Center Lane/Lyons/Lodato)

A live recording of the Stanford Law School and Cyber Policy Center's *Moderated Content*, a podcast about content moderation, with Evelyn Douek and Alex Stamos. The recording will feature special guests engaging in wonky conversation about the regulation of what you see, hear and do online.



from Stanford Law School

### Friday, September 29, 2023

#### 8:00 – 9:00 **Registration and Continental Breakfast**

9:00 – 10:00 Select one of the below, simultaneously held, sessions to attend.

Three Coordinated Studies on Misinfo Judgements	Research Session: Civic and Harmful Content	Lightning Talks: Generative AI
McCaw Hall Mainstage	Fisher Conference Center	Fisher Conference Center
	Lane/Lyons/Lodato	Barnes/McDowell/Cranston

# A Platform-University Collaboration: Three Independent but Coordinated Studies on Crowdsourced Misinformation Judgments (McCaw Hall Mainstage)

Moderated by Henry Silverman

- o Jennifer Allen, MIT Sloan School of Management
- o Zeve Sanderson, NYU's Center for Social Media and Politics
- o Paul Resnick, University of Michigan

#### Civic and Harmful Content (Fisher Conference Center Lane/Lyons/Lodato)

Five studies examining the classification, spread, or effects of civic, false, and violent content. *Moderated by Adam Berinsky, Massachusetts Institute of Technology* 

- Classifying the Spread of Violent Memes on Social Media Rod Abhari, Northwestern University
- The Effect of Online Misinformation Exposure on False Election Beliefs Ross Dahlke, Stanford University
- Identifying Informative Civic Content on Facebook Karen Jusko, Meta
- Insights from a Comparative Study on the Variety, Velocity, Veracity, and Viability of Crowdsourced and Professional Fact-Checking Services
   Andy Zhao, Cornell University
- Identifying Search Directives on Social Media Ronald Robertson, Stanford Internet Observatory

### **Generative Artificial Intelligence** (Fisher Conference Center Barnes/McDowell/Cranston)

Lightning talks feature five-minute, rapid fire presentations with time for questions.

Moderated by John Perrino, Stanford Internet Observatory

- Generative AI, Elections and Information Disorder in Africa
   Hannah Ajakaiye, FactsMatterNG; JSK Journalism Fellow, Stanford University
- Should AI (Help) Feed the Trolls?
   Susan Benesch, Dangerous Speech Project

- Accuracy and AI: Towards the Consideration of Accuracy as a Legal Requirement Between Data,
   Safety and Performance
  - Elisabetta Biasin, University of Leuven Centre for IT & IP Law
- Public Opinion about Political Campaigns Using Deepfakes
   Josh Goldstein, Center for Security and Emerging Technology at Georgetown University
- From Detection to Counterspeech: Generative AI in Content Moderation
   John Wihbey, Northeastern University
- Large Language Models Can Rate News Outlet Credibility
   Kai-Cheng Yang, Observatory on Social Media, Indiana University

#### 10:00 - 10:30 **Break**

10:30 – 11:30 Select one of the below, simultaneously held, sessions to attend.

Research Session: The Experience of Reporting Abuse Online	Panel: Self-Harm Contagion	Lightning Talks: Misinformation and News
McCaw Hall Mainstage	Fisher Conference Center Lane/Lyons/Lodato	Fisher Conference Center Barnes/McDowell/Cranston

#### The Experience of Reporting Abuse Online (McCaw Hall Mainstage)

Four studies on product design and victim experience of user reporting flows. *Moderated by Eli Sugarman, Hewlett Foundation* 

- A Comparative Analysis of Platform Reporting Flows
  - Alex Leavitt, UC Berkeley/Roblox; Kat Lo, Meedan
- An Analysis of Financial Sextortion Victim Posts Published on Reddit.com/r/Sextortion
   Jacques Marcoux, Canadian Centre for Child Protection
- "Is Reporting Worth the Sacrifice of Revealing What I've Sent?": Privacy Considerations When Reporting on End-to-End Encrypted Platforms
   Leijie Wana, University of Washington
- Online Sexual Exploration and Risks of Exploitation: Elevating the Voices of LGBTQ+ Youth Amanda Goharian, Thorn

#### **Self-Harm Contagion** (Fisher Conference Center Lane/Lyons/Lodato)

Chaired by the Stanford Center for Youth Mental Health and Wellbeing, this session will look at how harmful social contagions spread online and impact young users, both for the better or for the worse. *Moderated by Vicki Harrison, Stanford Center for Youth Mental Health and Wellbeing* 

- Matt Soeth, Spectrum Labs
- Ryn Linthicum, TikTok
- o Anjuli Corzine, Columbia University
- o Sunny Liu, Stanford Social Media Lab

Misinformation and News (Fisher Conference Center Barnes/McDowell/Cranston)

Lightning talks feature five-minute, rapid fire presentations with time for questions.

Moderated by Daphne Keller, Stanford Cyber Policy Center

- Content Modeling in Multi-Platform Multilingual Social Media Data Arman Setser, Graphika
- The Role of Advertisers and Platforms in Monetizing Misinformation:
   Descriptive and Experimental Evidence
   Wajeeha Ahmad, Stanford University
- RCTs Prove Instagram Campaigns Can Have Lasting Counter-Disinfo Effect
   Debra Lavoy, Reality Team
- Partisan Conflict Over Content Moderation is More Than Disagreement About Facts
   Ruth Elisabeth Appel, Stanford University
- Breaking Echo Chambers Using a Sound-Based Nudge Technology
   Kazutoshi Sasahara, Tokyo Institute of Technology
- The Role and Influence of Perceived Experts in an Anti-Vaccine Misinformation Community Mallory Harris, Stanford University
- Political Gambit of Censoring Under the Guise of Preventing Disinformation: Regulatory Developments in Turkey
   Burak Haylamaz, Centre for Information Policy Leadership
- o Exposure to Misinformation: What Do (and Don't) We Know?
  - Ryan Moore, Stanford University
- The Power of Storytelling: Shujaaz's Journey to Segment Social Media Audiences for Positive Change on Sexual and Reproductive Health.
   Anastasia Mirzoyants, Shujaaz Inc.

#### 11:30 - 1:00 Lunch

Attendees are welcome to enjoy their lunch in the gardens or bring their lunch to the breakout.

# 11:45 – 12:45 Industry <> Researcher Collaboration: Sharing insights from the Tech Coalition Safe Online Research Fund (Fisher Conference Center Lane/Lyons/Lodato)

The Tech Coalition Safe Online Research Fund focuses specifically on efforts around online child sexual exploitation and abuse. The lessons learned from this initiative are broadly applicable to the Trust & Safety community working at the intersection of independent research and industry policy and practice. In this session, you will hear about the unique way companies and independent researchers engage through the Tech Coalition Safe Online Research Fund, hear insights and impact from the independent research to date, and have the chance to exchange with collaborators in this novel initiative.

Moderated by Natalie Shoup, Safe Online Initiative; Kathryn Grant, TikTok; Marija Manojlovic, Safe Online; Kay Chau, Tech Coalition

1:00 – 2:00 Select one of the below, simultaneously held, sessions to attend.

Research Session: Scaling Content Moderation to New Harms

McCaw Hall Mainstage

Research Session: Global Attitudes and Online Harms

Fisher Conference Center Lane/Lyons/Lodato

Workshop: Applications of the Meta Content Library for Trust & Safety Research

Fisher Conference Center Barnes/McDowell/Cranston

#### **Scaling Content Moderation to New Harms** (McCaw Hall Mainstage)

Four studies on new ways of thinking about content moderation.

Moderated by Charlotte Willner, Trust & Safety Professional Association

- Creating and Testing Classifiers for Civic Health Kylan Rutherford, Civic Health Project
- Large-Scale Inclusion Training for Online Community Moderators
   Xuan Zhao, Stanford University
- Multistakeholder Content Governance in South Korea
   Inyoung Cheong, University of Washington
- Artificial Intelligence for the Prevention and Detection of Online Sexual Exploitation and Abuse of Children and Adolescents: What We Can Learn from Conversations Between Offenders and Reports of Victims
  - Lina Maria Saldarriaga, Aulas en Paz

#### Global Attitudes and Online Harm (Fisher Conference Center Lane/Lyons/Lodato)

Four studies on the ways online harms are handled in diverse global contexts.

Moderated by Tongtong Zhang, Stanford Internet Observatory

- Responding to Online Sexual Abuse and Exploitation of Children in the Philippines: Understanding Facilitator-Perpetrator Engagement and the Role of Technology in Enabling Demand Maggie Brennan, Dublin City University
- Harms and Remedies Across Countries
   Sarita Schoenebeck, University of Michigan
- How Gen Z Navigate Eating and Health Pathways Rachel Xu, Google
- Right to be De-listed? User Reactions to Delisting Public-Interest Content from Search Results
   Rebecca Umbach, Google

#### Application of the Meta Content Library for Trust & Safety Research

(Fisher Conference Center Barnes/McDowell/Cranston)

This workshop will introduce Trust & Safety researchers to the Meta Content Library and demo new data fields and functionalities available in both the User Interface and the API.

Facilitated by the Meta Research Partnerships Team: Michael Zoorob, Jessica Bachman, Yair Rubinstein, Christina Fan

2:00 – 2:15 **Break** 

2:15 – 3:15 Select one of the below, simultaneously held, sessions to attend.

Research Session:Trust & Safety TeachingWorkshop: TikTok ResearchPlatform PolicyConsortiumAPI Workshop for AcademicsMcCaw Hall MainstageFisher Conference Center<br/>Lane/Lyons/LodatoFisher Conference Center<br/>Barnes/McDowell/Cranston

#### **Platform Policy** (McCaw Hall Mainstage)

Five studies on ways platforms can and do self-regulate around complex spaces in content moderation.

Moderated by Riana Pfefferkorn, Stanford Internet Observatory

- Speculative Futures of Restorative Online Content Moderation Anna Gibson, Northeastern University
- Trust and Safety for Immersive Technologies
   Michael Karanicolas, UCLA Institute for Technology, Law & Policy
- Antisemitism Across Platforms: Networked Online Hate on the Fringes and in the Mainstream Stephen Rea, Anti-Defamation League
- Studying the Online Deepfake Community Brian Timmerman, NYU
- Designing for the Unknown: Intervening in Potential Safety Incidents
   Naomi Yarin, Lyft

#### Trust & Safety Teaching Consortium (Fisher Conference Center Lane/Lyons/Lodator))

Join members of the consortium in a discussion of how to grow and enhance the consortium and make it as useful as possible.

- o Alex Stamos, Stanford Internet Observatory
- Shelby Grossman, Stanford Internet Observatory
- Jeffrey T. Hancock, Stanford Internet Observatory
- o Inbal Goldberger, ActiveFence
- Brian Murphy, Georgetown University
- Michael Swenson
- Alex Leavitt, UC Berkeley | Roblox
- Leslie Taylor, Genpact
- Kerry Keyes, Columbia University
- o Sarah Shirazyan, Instagram | Stanford Law School

#### **TikTok Research API Workshop for Academics**

(Fisher Conference Center Barnes/McDowell/Cranston)

In this workshop, academic researchers will learn about TikTok's Research API and have the opportunity to walk through sample queries. Hands-on participation is limited to academic researchers who have pre-registered for temporary API access. Others may observe. Facilitated by Shiva Hullinakatte, Max Baryshevstev, Maurice Turner, and Kathryn Grant, TikTok

3:15 – 3:30	Break
3:30 - 4:00	Closing Fireside Chat: Yoel Roth and Janine Zacharia (McCaw Hall) Introduction by Alex Stamos, Stanford Internet Observatory
4:00 - 5:30	Happy Hour (Ford Gardens)

#### Resources

Thank you for joining us at the second annual Trust & Safety Research Conference. A detailed conference agenda is available at <a href="https://conferences.law.stanford.edu/tsrc/agenda">https://conferences.law.stanford.edu/tsrc/agenda</a>.

Recordings of the mainstage sessions will be posted to the Stanford Cyber Policy Center YouTube Channel and website after the event.

Conference proceedings have been published through the *Journal of Online Trust and Safety*, available at <a href="http://tsjournal.org">http://tsjournal.org</a>.

### **Event Sponsor**



The Stanford Internet Observatory wishes to thank the William and Flora Hewlett Foundation for their generous grant in support of this event.



io.stanford.edu