

# **CSE291D – Sociotechnical Cybersecurity**

*Course introduction, definitions, harms, and course details*

UC San Diego

# whoami

osoft



# whoami



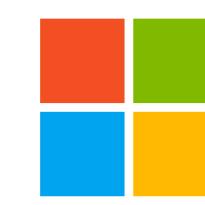
# whoami



# whoami

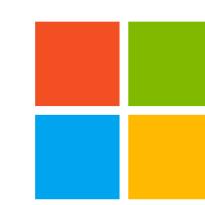


# whoami



Microsoft

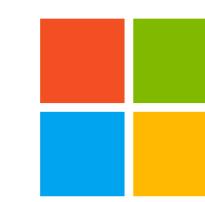
# whoami



Microsoft



# whoami



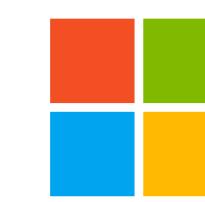
Microsoft



Google



# whoami



Microsoft



Google



# Also, I make theatre



Playwright Deepak Kumar, whose "House of India" play will make its world premiere at the Old Globe in 2025. (The Old Globe)

## 'House of India'

This world premiere play was written by San Diego resident Deepak Kumar, who is both a playwright and a computer science assistant professor at UC San Diego. The play is set in a family-owned Indian restaurant in a Cleveland-area strip mall, where Ananya has reluctantly taken over after her husband's death. She's faced with mounting bills and disagreement between her children over whether to stick to her husband's traditional menu or modernize with a quick-service, fusion concept. "It's about the American question of holding on to our traditions versus assimilating into the mainstream," Edelstein said. It will be directed by Zi Alikhan (Pasadena Playhouse's "Sanctuary City"). *May 10-June 1.*

# Course Staff



Deepak Kumar  
Assistant Professor, CSE  
[kumarde@ucsd.edu](mailto:kumarde@ucsd.edu)



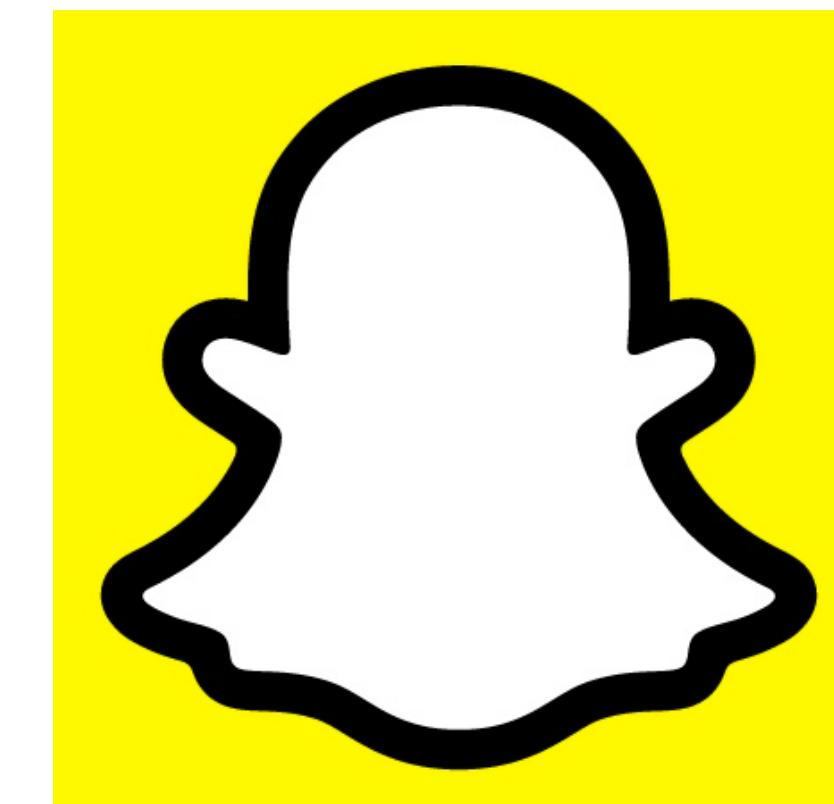
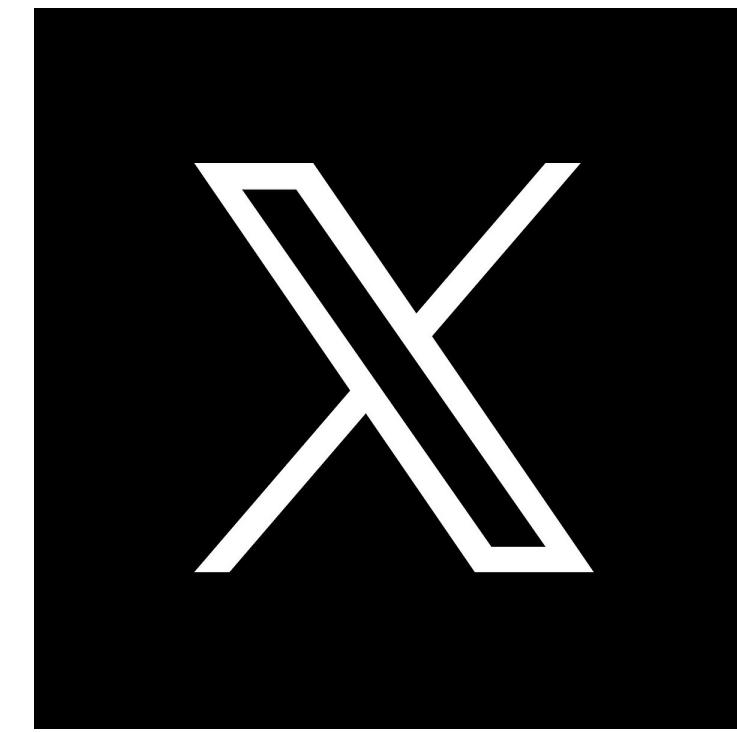
Muchan Li  
MS Student, CSE  
[mul005@ucsd.edu](mailto:mul005@ucsd.edu)

# Introductions

1. Your name, your degree program, when you moved to San Diego
2. Why are you enrolled in this class? What brought you here today?
3. What is one thing you're looking forward to this quarter (not class related?)

# **Let's poll the room**

# Let's poll the room

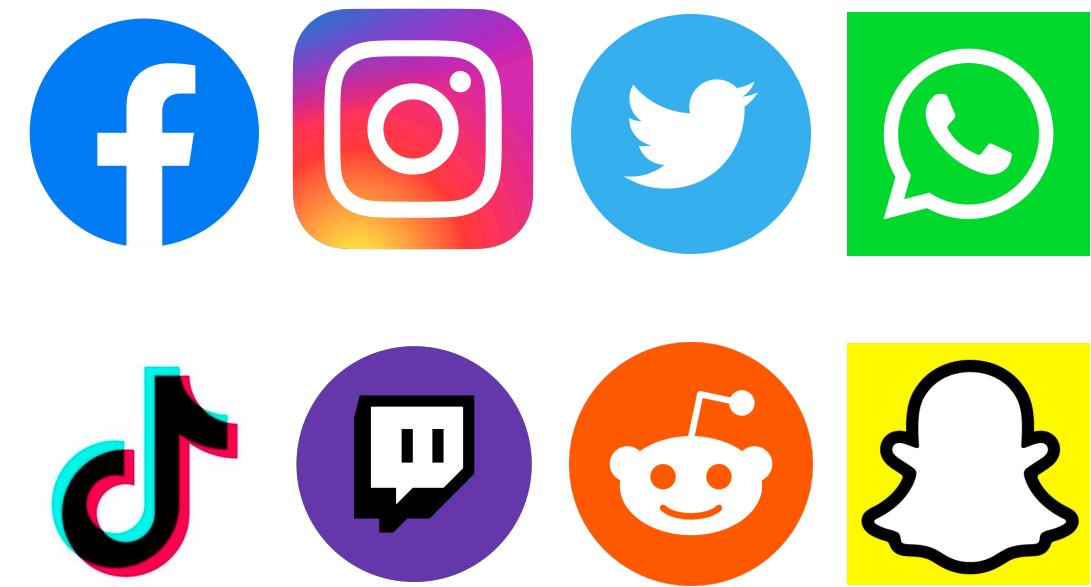


# The rise of sociotechnical systems

*Technology that interacts with personal, community, or societal interests*

# The rise of sociotechnical systems

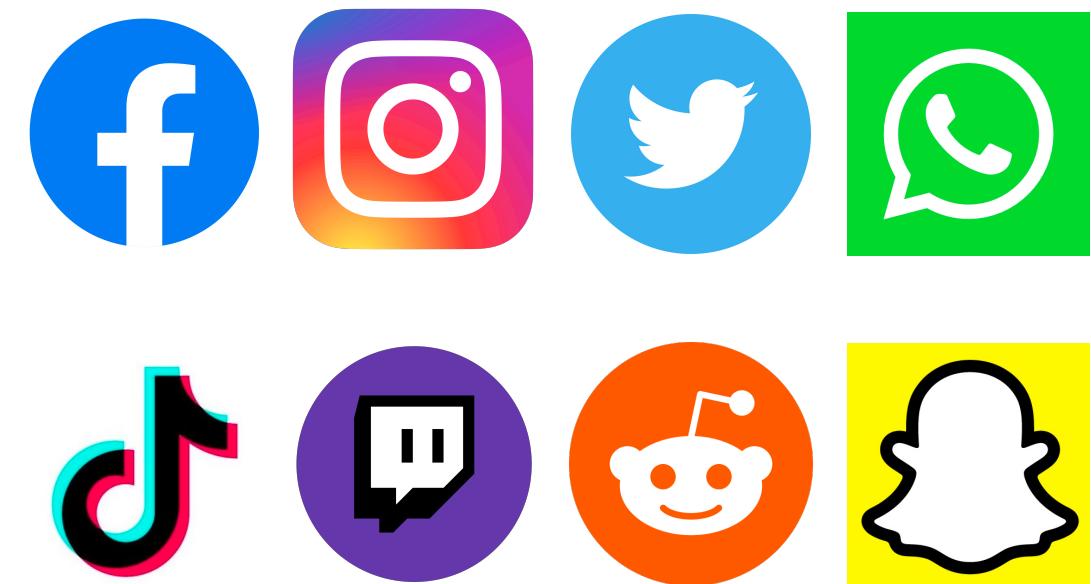
*Technology that interacts with personal, community, or societal interests*



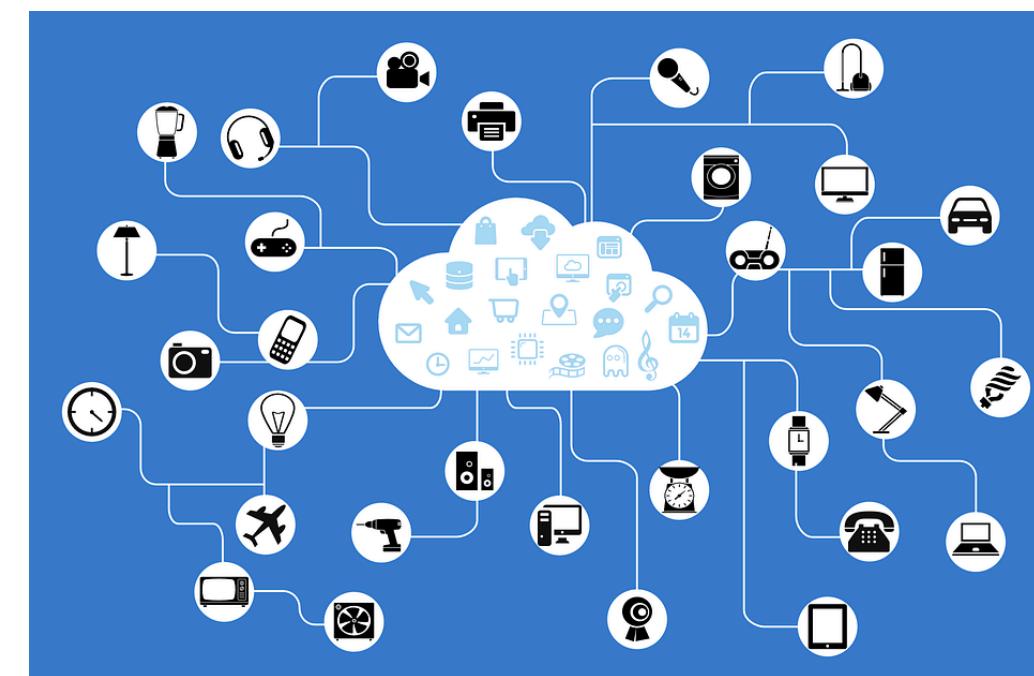
Social Media

# The rise of sociotechnical systems

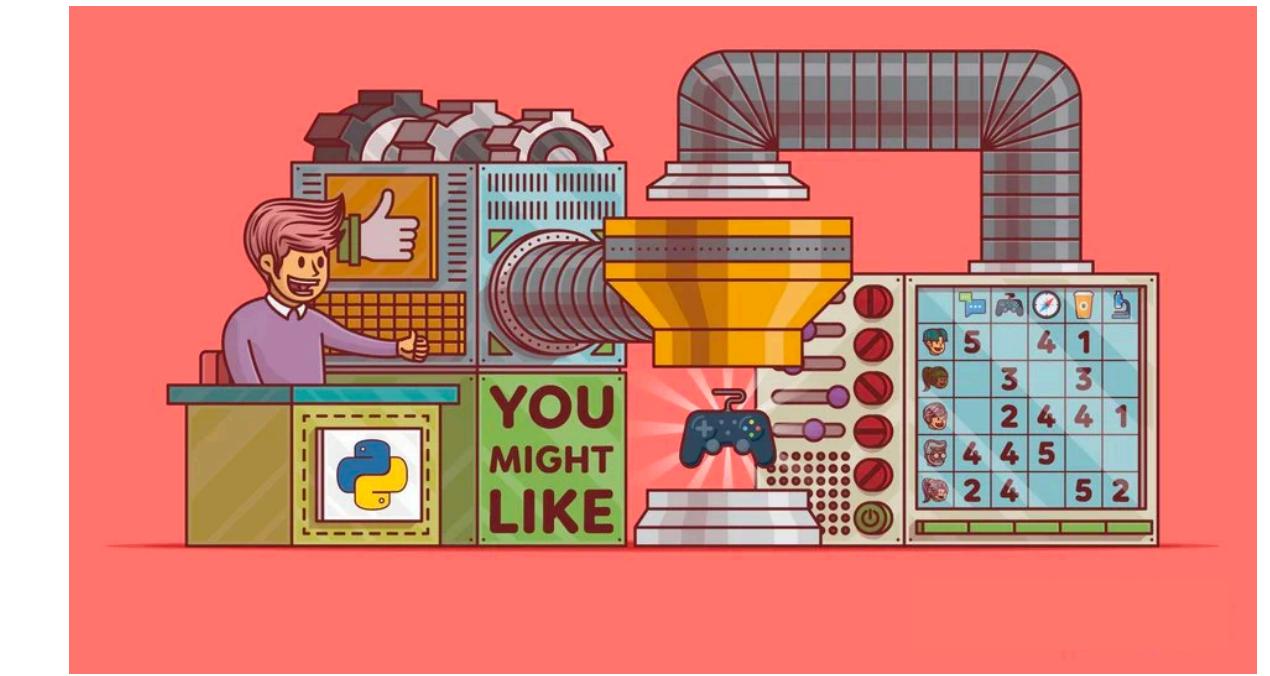
*Technology that interacts with personal, community, or societal interests*



Social Media



Internet of Things



Recommender Systems

# **Sociotechnical Cybersecurity**

# Sociotechnical Cybersecurity

How do computer systems fail in  
the presence of an adversary?

# Cybersecurity

# Defining Sociotechnical Cybersecurity

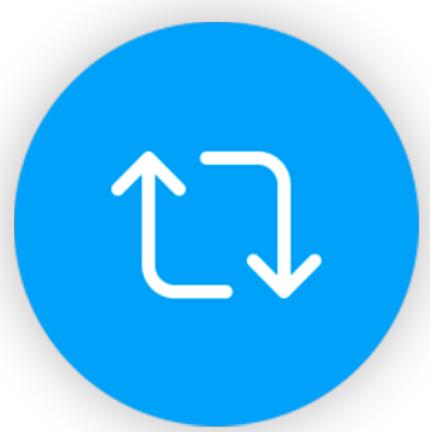
The study of how an adversary can use a computer system to cause societal-level harms.

# Sociotechnical systems enable security + safety threats

*Security + safety challenges emerge at scale*

# Sociotechnical systems enable security + safety threats

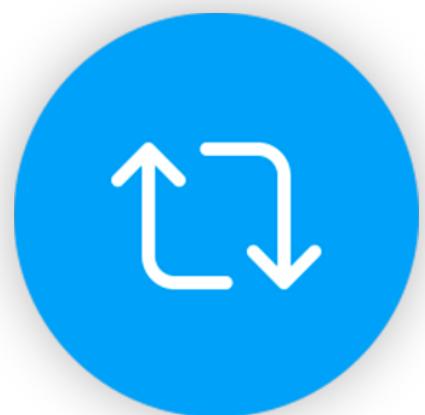
*Security + safety challenges emerge at scale*



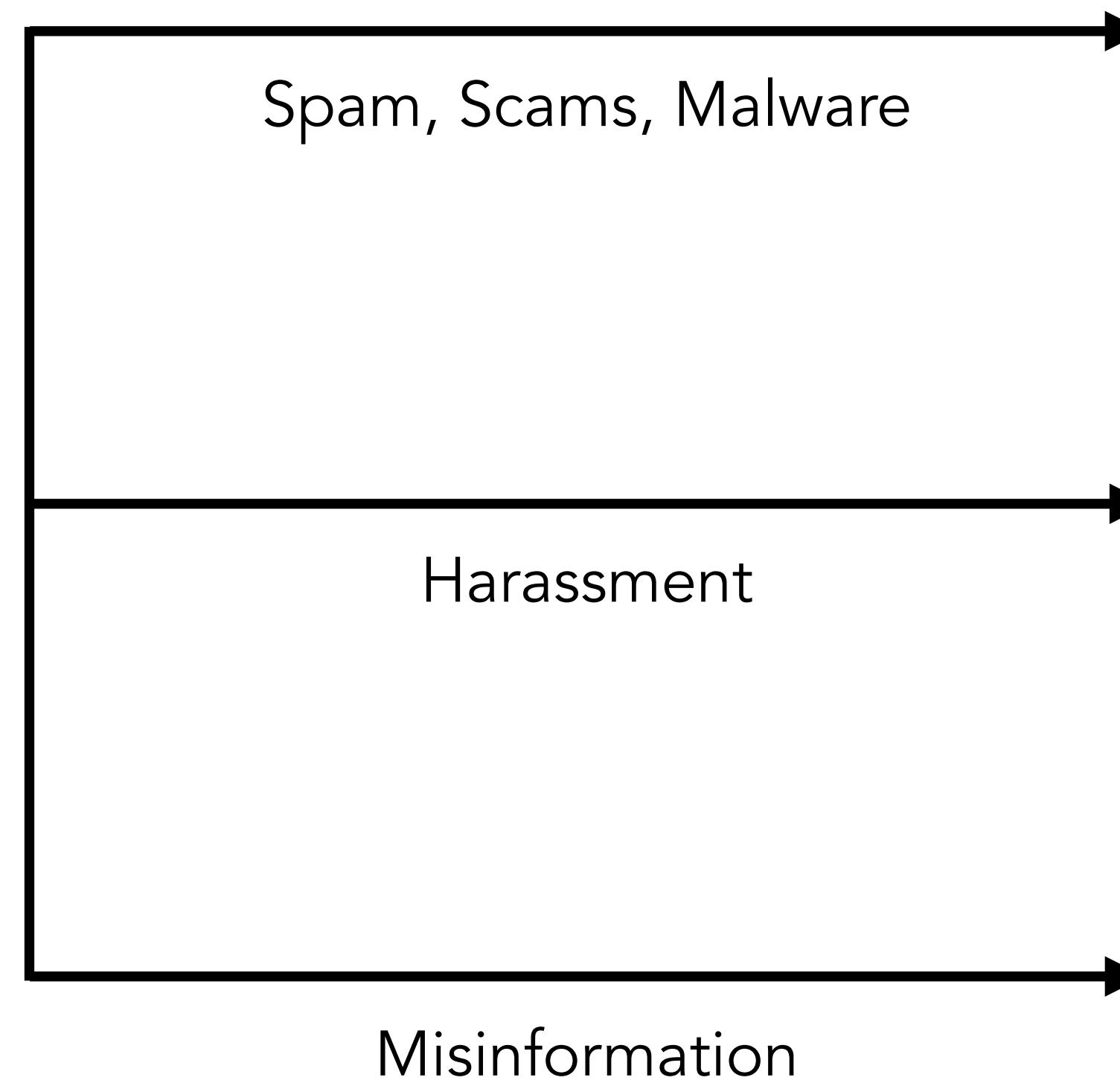
Repost

# Sociotechnical systems enable security + safety threats

*Security + safety challenges emerge at scale*



Repost



**10% of Twitter's active accounts are posting spam content, estimates GlobalData**

**@spam: The Underground on 140 Characters or Less \***

Chris Grier<sup>†</sup> Kurt Thomas<sup>\*</sup> Vern Paxson<sup>†</sup> Michael Zhang<sup>†</sup>

<sup>†</sup>University of California, Berkeley  
{grier, vern, mczhang}@cs.berkeley.edu

<sup>\*</sup>University of Illinois, Champaign-Urbana  
kathoma2@illinois.edu

**72 Hours of #Gamergate**

Digging through 316,669 tweets from three days of Twitter's two-month-old trainwreck

**Twitter users are twice as likely to retweet fake news stories than authentic ones**

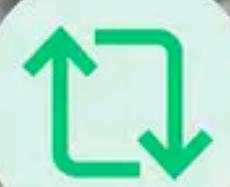
"There are real world and potentially negative consequences if decisions are going to be made based off falsity."

**Examining the Impact of Internet Research Agency Tweets in the 2016 U.S. Election**

# Social

Facebook

Re



# Search

spam

stories

ed off falsity."

rch  
on

# A non-exhaustive list of sociotechnical challenges

Violent & Criminal Behavior

Regulated Goods & Services

Offensive & Objectionable Content

User safety

Scaled / Automated Abuse

Deception & Fraud

# A non-exhaustive list of sociotechnical challenges

## Violent & Criminal Behavior

Regulated Goods & Services

Offensive & Objectionable Content

User safety

Scaled / Automated Abuse

Deception & Fraud

Extremist groups

Explicit threats of violence

Bomb-making instructions

Child sexual abuse material

Non-consensual sharing of intimate images

Human Trafficking

# A non-exhaustive list of sociotechnical challenges

Violent & Criminal Behavior

## Regulated Goods & Services

Offensive & Objectionable Content

User safety

Scaled / Automated Abuse

Deception & Fraud

Weapons dealing

Endangered animal products

Gambling and gaming

Selling access to nude images

# A non-exhaustive list of sociotechnical challenges

Violent & Criminal Behavior

Regulated Goods & Services

## Offensive & Objectionable Content

User safety

Scaled / Automated Abuse

Deception & Fraud

Slurs directed at protected classes

Support for racial supremacy movements

Imagery of fatal incidents

Imagery of animal cruelty

Pornography

# A non-exhaustive list of sociotechnical challenges

Violent & Criminal Behavior

Regulated Goods & Services

Offensive & Objectionable Content

## User safety

Scaled / Automated Abuse

Deception & Fraud

Suicide and self harm content

Harassment and bullying

Doxxing (e.g., publishing personally identifiable information)

Dangerous misinformation and endangerment

# A non-exhaustive list of sociotechnical challenges

Violent & Criminal Behavior

Regulated Goods & Services

Offensive & Objectionable Content

User safety

**Scaled / Automated Abuse**

Deception & Fraud

Spam

Distribution of malware

Distribution of spyware

Inauthentic behavior (e.g., fake accounts to deceive or manipulate users)

# A non-exhaustive list of sociotechnical challenges

Violent & Criminal Behavior

Regulated Goods & Services

Offensive & Objectionable Content

User safety

Scaled / Automated Abuse

**Deception & Fraud**

Loan scams

Fake charity solicitations

Account hijacking

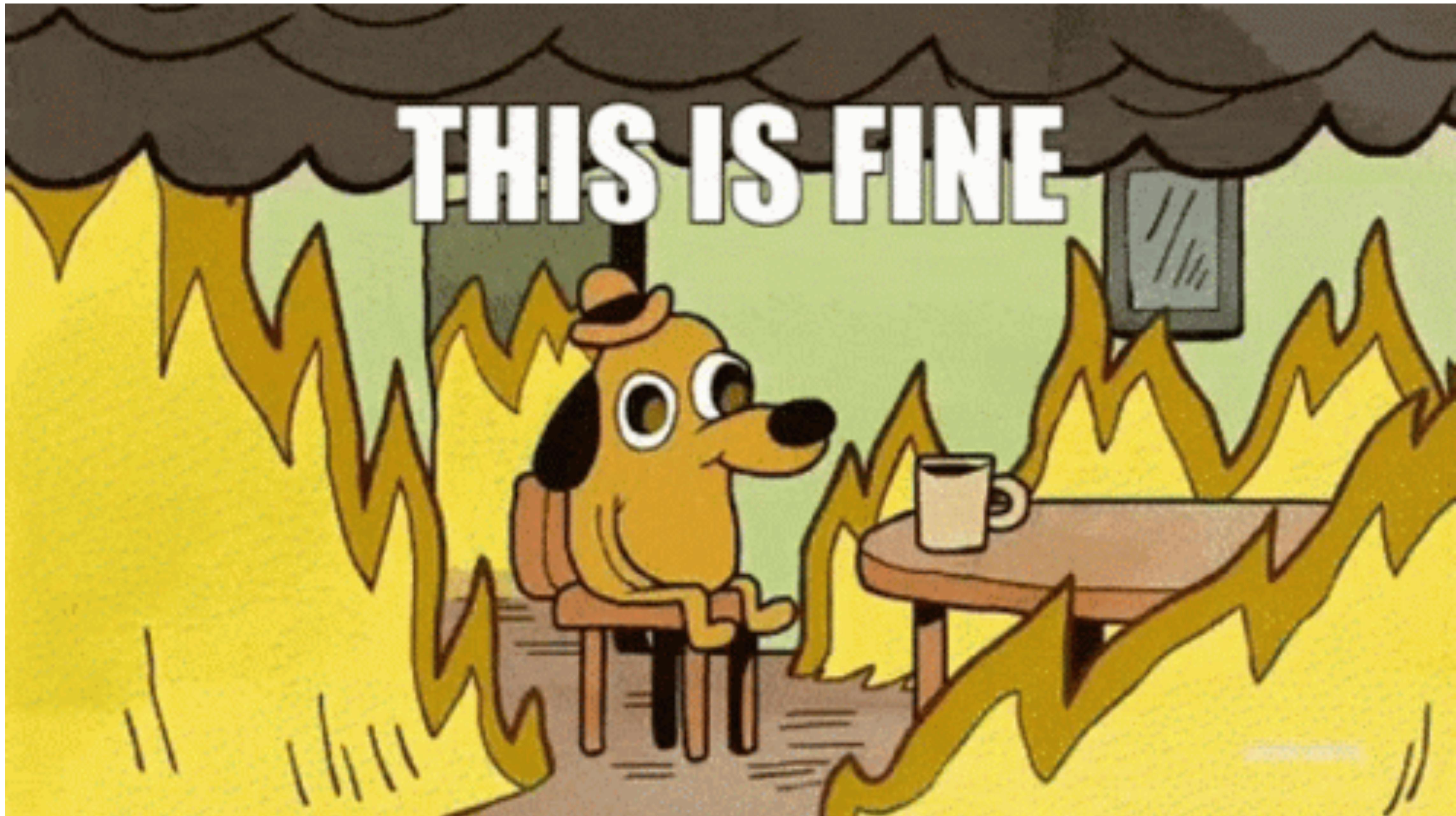
Fake names

Phishing

Posting photos / videos belonging to others

Defamation

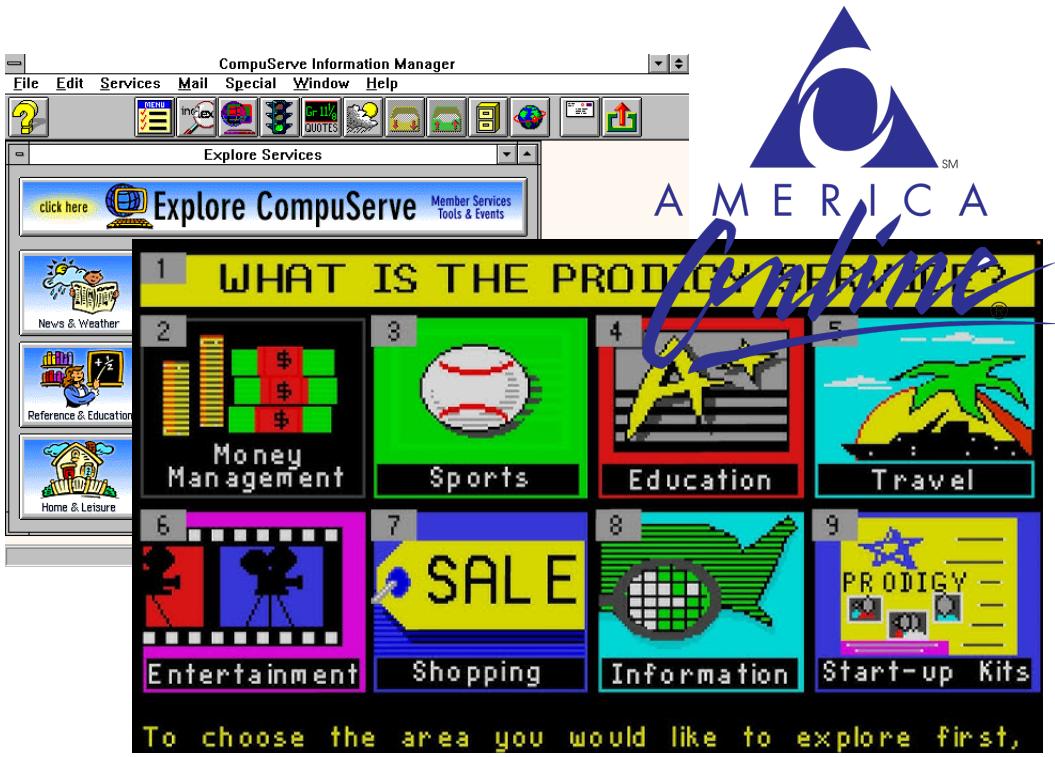
**THIS IS FINE**



# How did we get here?

**Content warning: Heavy  
material ahead**

# An incomplete history of our online content landscape

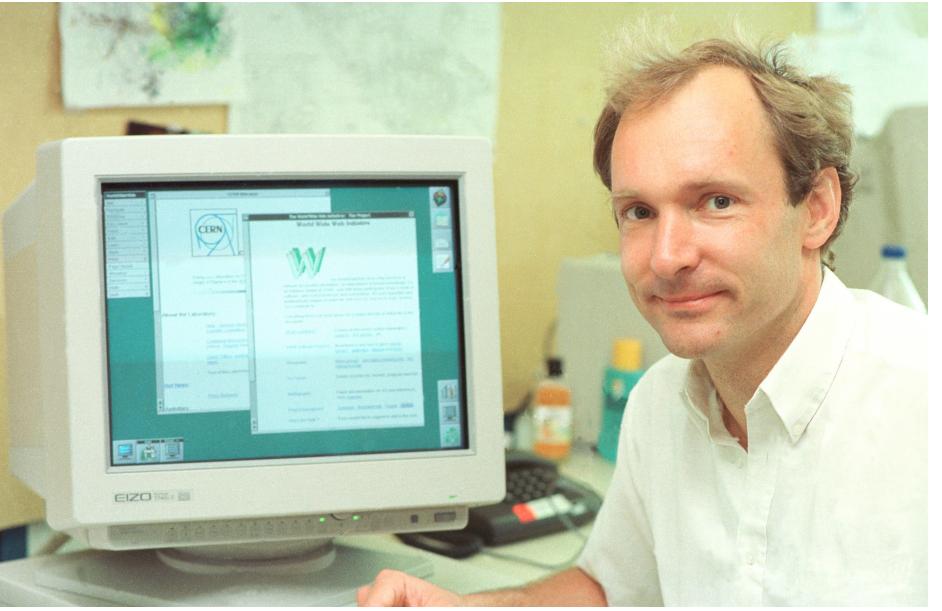


Late 1980s: Online services are born

# An incomplete history of our online content landscape



Late 1980s: Online services are born



1991: Tim Berners-Lee launches  
the first website

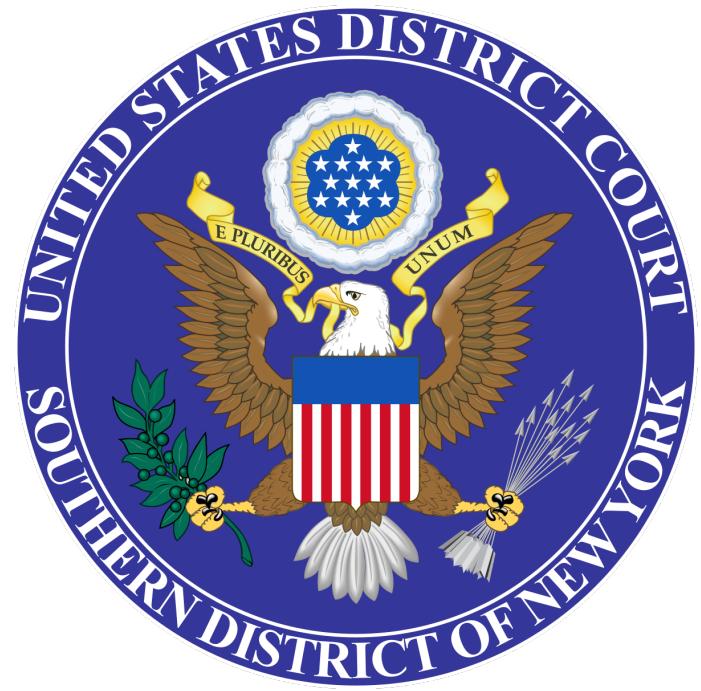
# An incomplete history of our online content landscape



Late 1980s: Online services are born



1991: Tim Berners-Lee launches the first website



1991: Cubby, Inc vs. CompuServe Inc.

# An incomplete history of our online content landscape



Late 1980s: Online services are born



1991: Tim Berners-Lee launches the first website



1991: Cubby, Inc vs. CompuServe Inc.

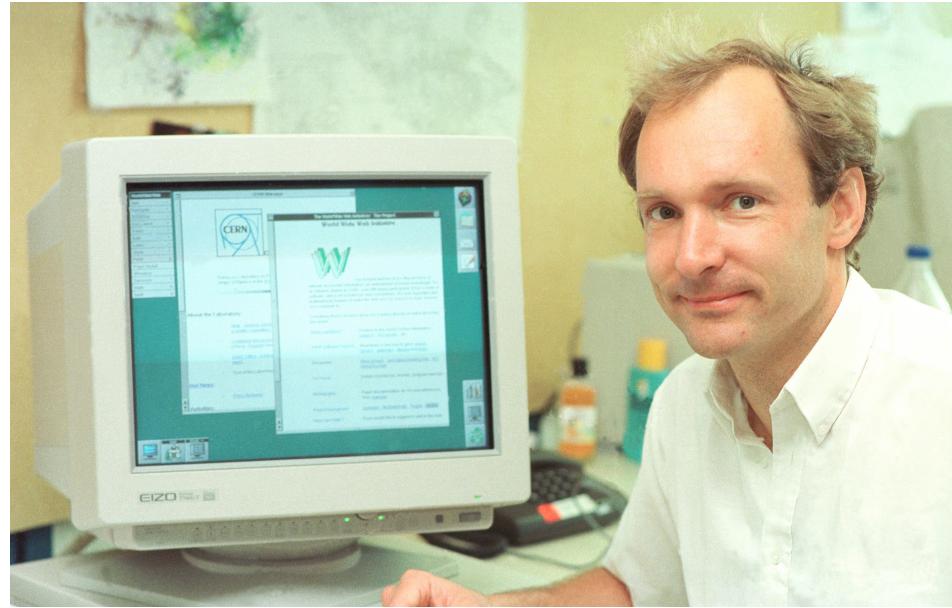
A newspaper clipping from a publication like USA Today. The headline reads 'The Internet? Bah!' in large bold letters. Below it, a sub-headline says 'Hype Alert: Why cyberspace isn't, and will never be, all that's cracked up to be'. The article is signed 'By CLIFFORD STOLL'. To the right of the main text is a small sidebar with the text 'pretense of completeness. La...'. The full text of the article is cut off at the bottom.

1995: People think the Internet is dumb and is doomed to fail

# An incomplete history of our online content landscape



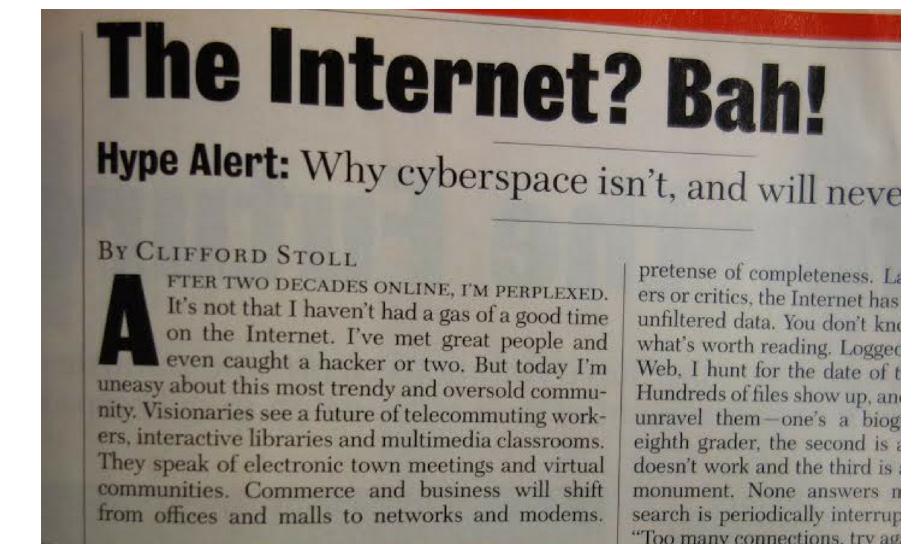
Late 1980s: Online services are born



1991: Tim Berners-Lee launches the first website



1991: Cubby, Inc vs. CompuServe Inc.

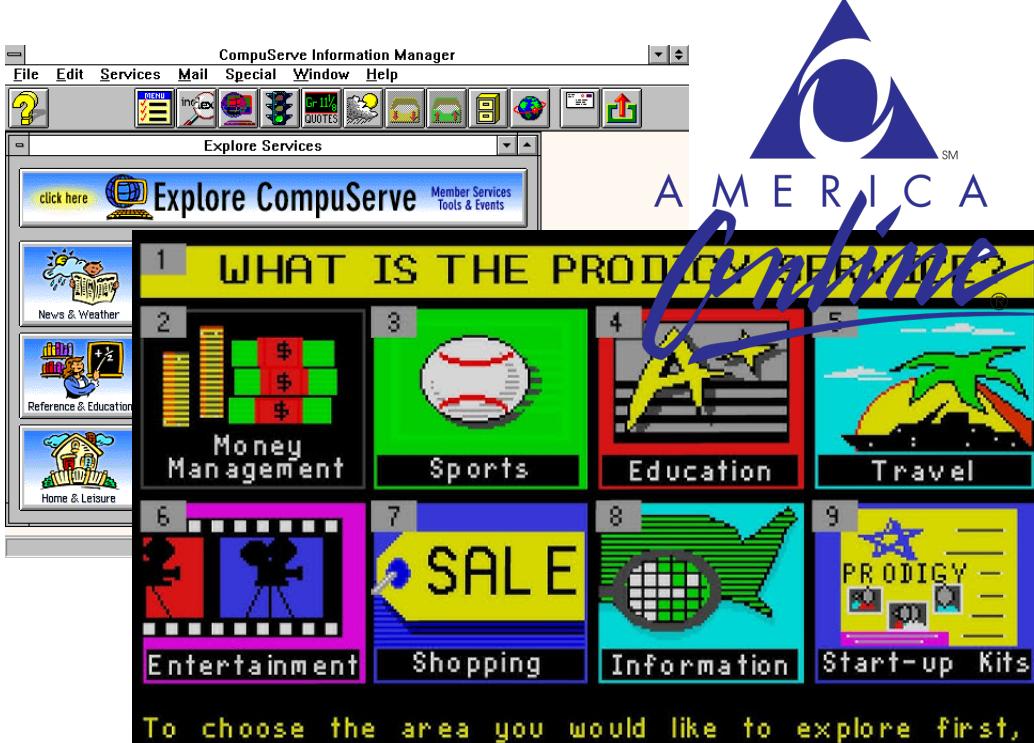


1995: People think the Internet is dumb and is doomed to fail



1995: Stratton Oakmont vs. Prodigy, Inc.

# An incomplete history of our online content landscape



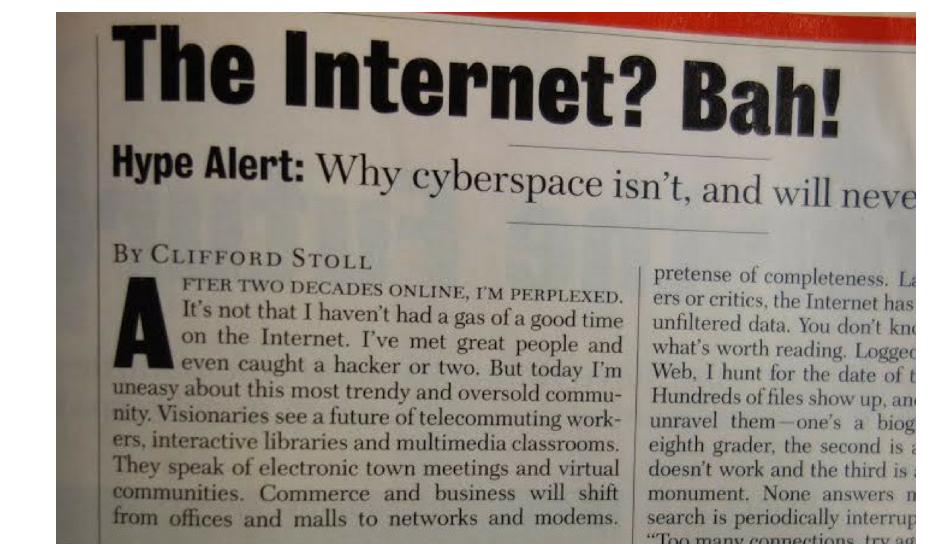
Late 1980s: Online services are born



1991: Tim Berners-Lee launches the first website



1991: Cubby, Inc vs. CompuServe Inc.



pretense of completeness. La  
ers or critics, the Internet has  
unfiltered data. You don't know  
what's worth reading. Logged  
on the Web, I hunt for the date of t  
Hundreds of files show up, and  
unravel them—one's a blog  
eighth grader; the second is a  
doesn't work and the third is a  
monument. None answers n  
search is periodically interrup  
"Too many connections, try ag

1995: People think the Internet is dumb and is doomed to fail



1995: Stratton Oakmont vs. Prodigy, Inc.



1996: CDA + Section 230

# An incomplete history of our online content landscape

Late 1990s – Early 2000s: A proliferation of online social spaces

# An incomplete history of our online content landscape



1994

Late 1990s – Early 2000s: A proliferation of online social spaces

# An incomplete history of our online content landscape



1994



1997

Late 1990s – Early 2000s: A proliferation of online social spaces

# An incomplete history of our online content landscape



1994



1997



2003

Late 1990s – Early 2000s: A proliferation of online social spaces

# An incomplete history of our online content landscape



1994



1997



2003



2003

Late 1990s – Early 2000s: A proliferation of online social spaces

# An incomplete history of our online content landscape



1994



1997



2003



2003



2003

Late 1990s – Early 2000s: A proliferation of online social spaces

# An incomplete history of our online content landscape



1994



1997



2003



2003



2003

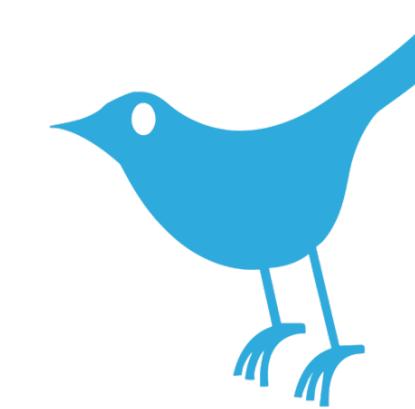
Late 1990s – Early 2000s: A proliferation of online social spaces



2004



2005



2006

# An incomplete history of our online content landscape

The screenshot shows the eBay homepage with the classic logo. At the top, there's a navigation bar with links for 'home', 'my eBay', 'site map', and 'sign in'. Below the logo, a banner for 'billpoint' offers '\$1 Off every time you use eBay Online Payments!'. The main search bar has a placeholder 'Search' and a 'Smart Search' link. A checkbox for 'Search titles and descriptions' is also present.

## About eBay

### ▼ Press Releases

[December 1999](#)  
[November 1999](#)  
[October 1999](#)  
[September 1999](#)  
[August 1999](#)  
[July 1999](#)  
[June 1999](#)

[May 1999](#)  
[April 1999](#)  
[March 1999](#)  
[February 1999](#)  
[January 1999](#)  
[December 1998](#)  
[November 1998](#)  
[October 1998](#)  
[September 1998](#)  
[August 1998](#)  
[July 1998](#)  
[June 1998](#)

[May 1998](#)  
[April 1998](#)  
[March 1998](#)  
[February 1998](#)  
[January 1998](#)

[Company Overview](#)  
[eBay Community](#)  
[eBay Foundation](#)  
[Contact Information](#)  
[Investor Relations](#)  
[Jobs](#)



## Press Releases

You've come to the right source for a complete archive of company press releases. Feel free to peruse eBay's many announcements about new product features, business initiatives, on-site promotions, business partnerships, user services, and much more.

- [01/29/99](#) - eBay Soars To New Heights
- [01/29/99](#) - eBay's response to the investigation by the New York City Department of Consumer Affairs
- [01/25/99](#) - Compaq and eBay Offer Easier Access To Leading Online Trading Community
- [01/25/99](#) - A Picture Is Worth A Thousand Words
- [01/15/99](#) - eBay™ Launches The Most Comprehensive Trust And Safety Upgrades To The World's Largest Person-To-Person Trading Site
- [01/05/99](#) - Mark McGwire and Sammy Sosa 1998 Home Run Balls To Be Auctioned Online and On-Live

### eBay Soars To New Heights

-- *Secures Position as the Second Most Popular Site on the Internet*

San Jose, Calif. Jan. 29, 1999 -

eBay™ (NASDAQ: EBAY; [www.ebay.com](http://www.ebay.com)), the world's leading person-to-person online trading community, today announced that it is ranked the second most popular site on the Internet, when measured by total user minutes per month (696 million total user minutes), according to Media Metrix December 1998 Web report.

1999 – First use of “Trust & Safety” from an online platform

# An incomplete history of our online content landscape

The screenshot shows the eBay homepage with a banner for Billpoint. Below it, the 'Press Releases' section is highlighted. The main heading 'Press Releases' is in red. A paragraph explains the purpose of the archive. A list of press releases from December 1999 follows, with links for each month from January to November.

Buyers! \$1 Off every time you use eBay Online Payments!

home | my eBay | site map | sign in  
Browse Sell Services Search Help Community  
overview news chat newsletter library charity eBay store about eBay  
Smart Search  
Search titles and descriptions

About eBay  
Press Releases  
December 1999  
November 1999  
October 1999  
September 1999  
August 1999  
July 1999  
June 1999  
May 1999  
April 1999  
March 1999  
February 1999  
January 1999  
December 1998  
November 1998  
October 1998  
September 1998  
August 1998  
July 1998  
June 1998  
May 1998  
April 1998  
March 1998  
February 1998  
January 1998  
Company Overview  
eBay Community  
eBay Foundation  
Contact Information  
Investor Relations  
Jobs

## Press Releases

You've come to the right source for a complete archive of company press releases. Feel free to peruse eBay's many announcements about new product features, business initiatives, on-site promotions, business partnerships, user services, and much more.

- [01/29/99](#) - eBay Soars To New Heights
- [01/29/99](#) - eBay's response to the investigation by the New York City Department of Consumer Affairs
- [01/25/99](#) - Compaq and eBay Offer Easier Access To Leading Online Trading Community
- [01/25/99](#) - A Picture Is Worth A Thousand Words
- [01/15/99](#) - eBay™ Launches The Most Comprehensive Trust And Safety Upgrades To The World's Largest Person-To-Person Trading Site
- [01/05/99](#) - Mark McGwire and Sammy Sosa 1998 Home Run Balls To Be Auctioned Online and On-Live

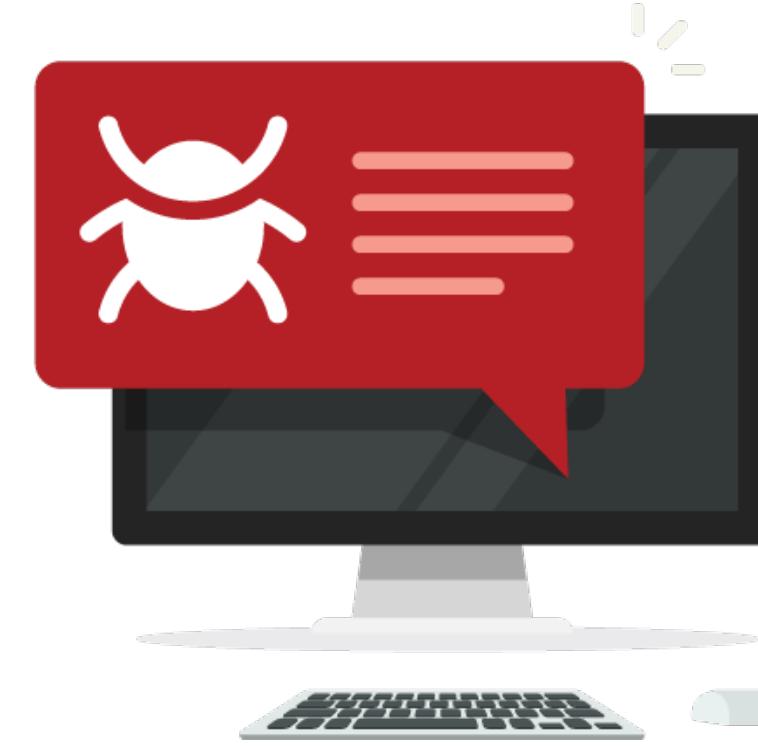
### eBay Soars To New Heights

-- Secures Position as the Second Most Popular Site on the Internet

San Jose, Calif. Jan. 29, 1999 -

eBay™ (NASDAQ: EBAY; [www.ebay.com](http://www.ebay.com)), the world's leading person-to-person online trading community, today announced that it is ranked the second most popular site on the Internet, when measured by total user minutes per month (696 million total user minutes), according to Media Metrix December 1998 Web report.

1999 – First use of “Trust & Safety” from an online platform



2000s – 2010s: Spam, Malware, Phishing, Account Hijacking

# An incomplete history of our online content landscape

Buyers! \$1 Off every time you use eBay Online Payments!

About eBay  
▼ Press Releases



## Press Releases

You've come to the right source for a complete archive of company press releases. Feel free to peruse eBay's many announcements about new product features, business initiatives, on-site promotions, business partnerships, user services, and much more.

- [01/29/99](#) - eBay Soars To New Heights
- [01/29/99](#) - eBay's response to the investigation by the New York City Department of Consumer Affairs
- [01/25/99](#) - Compaq and eBay Offer Easier Access To Leading Online Trading Community
- [01/25/99](#) - A Picture Is Worth A Thousand Words
- [01/15/99](#) - eBay™ Launches The Most Comprehensive Trust And Safety Upgrades To The World's Largest Person-To-Person Trading Site
- [01/05/99](#) - Mark McGwire and Sammy Sosa 1998 Home Run Balls To Be Auctioned Online and On-Live

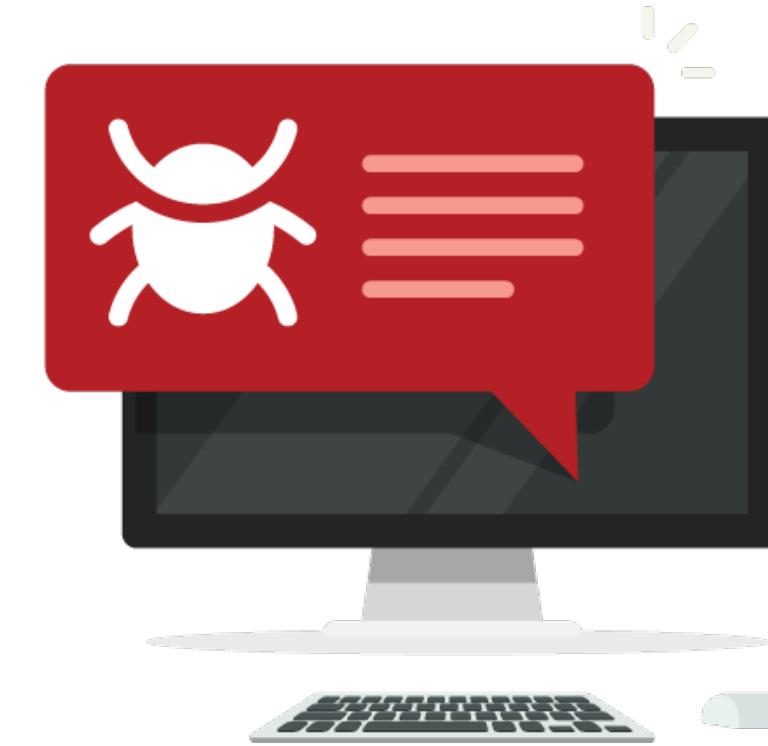
### eBay Soars To New Heights

-- Secures Position as the Second Most Popular Site on the Internet

San Jose, Calif. Jan. 29, 1999 -

eBay™ (NASDAQ: EBAY; [www.ebay.com](http://www.ebay.com)), the world's leading person-to-person online trading community, today announced that it is ranked the second most popular site on the Internet, when measured by total user minutes per month (696 million total user minutes), according to Media Metrix December 1998 Web report.

1999 – First use of “Trust & Safety” from an online platform



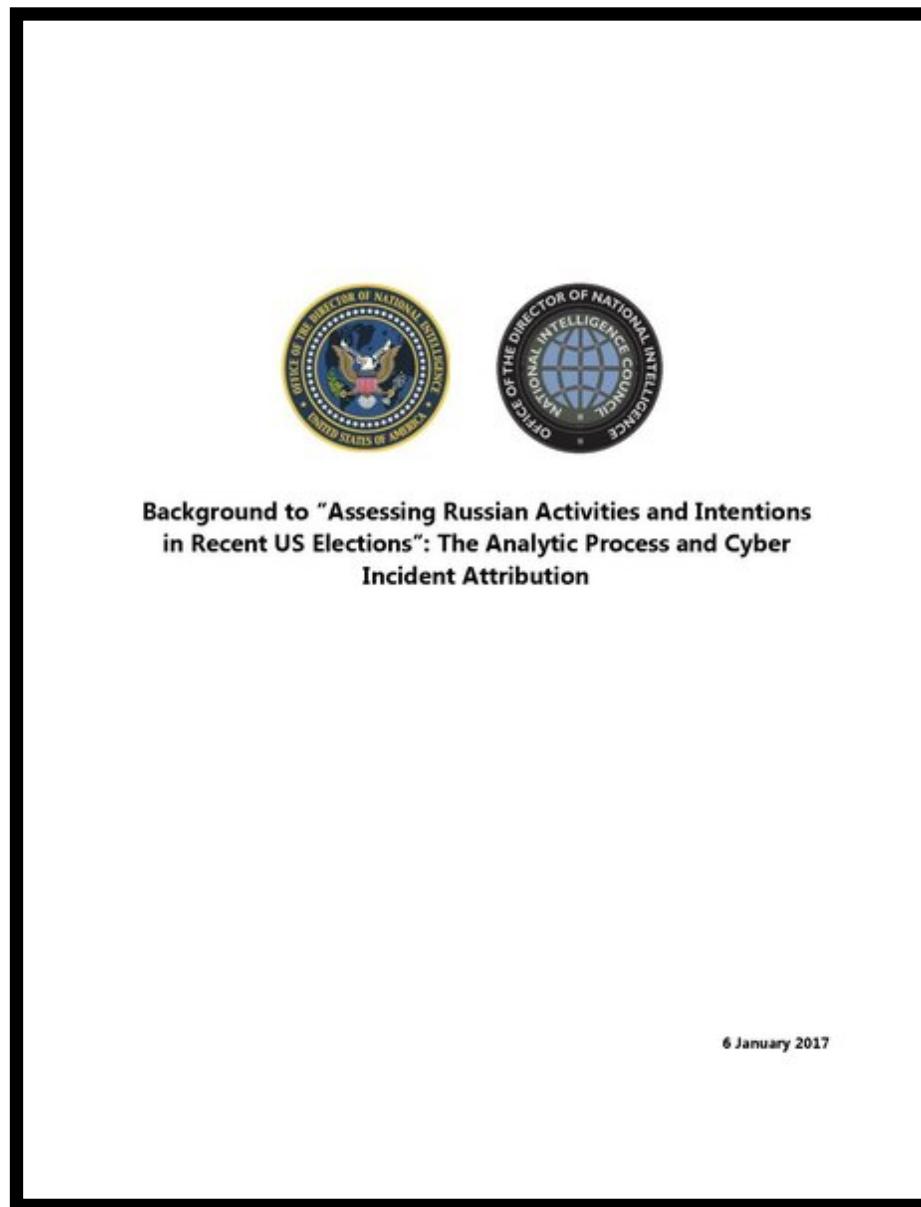
2000s – 2010s: Spam, Malware, Phishing, Account Hijacking

PhotoDNA™  
The Next Chapter in Protecting Children Online

FBI DIGITAL CRIMES UNIT facebook Microsoft

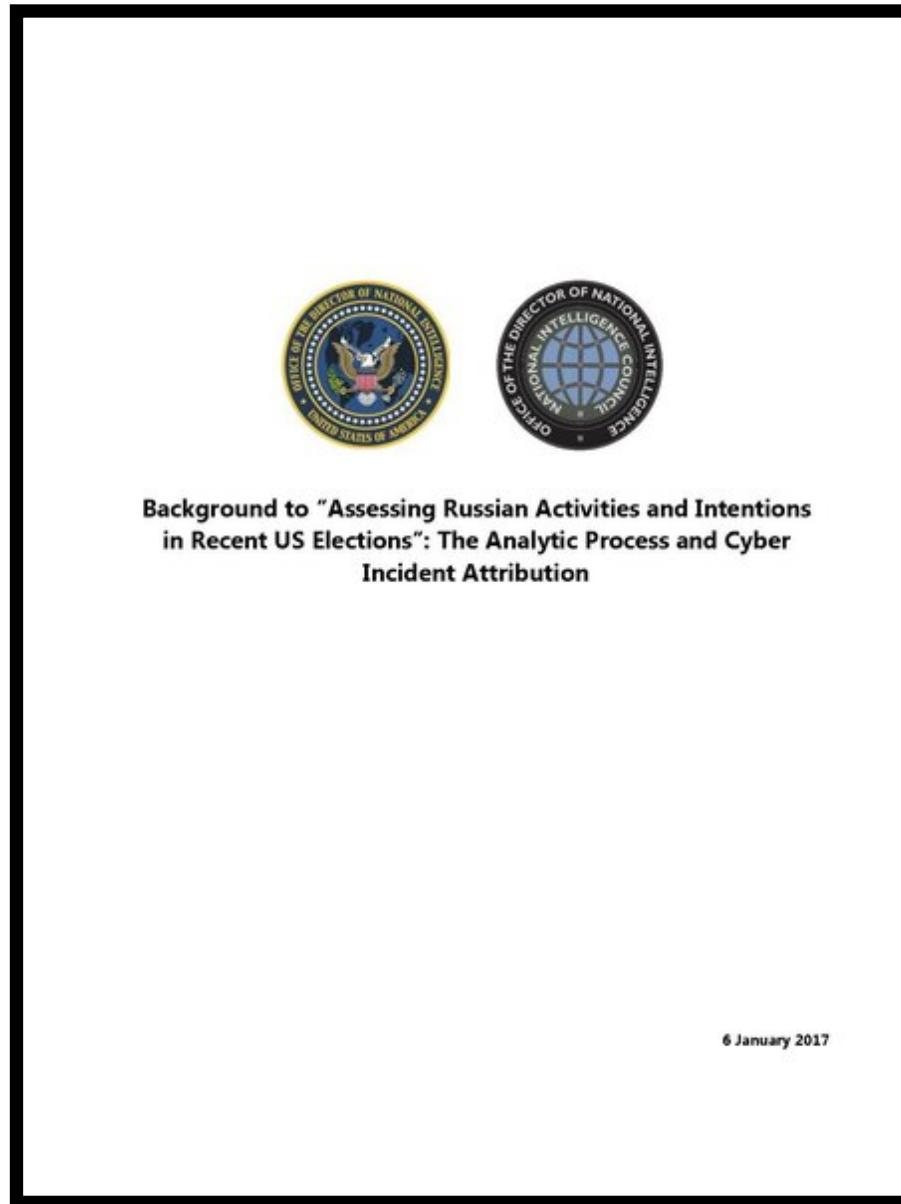
2009: Microsoft releases PhotoDNA to identify CSAM

# An incomplete history of our online content landscape



2016 – Russian digital misinformation interference in US election

# An incomplete history of our online content landscape



2016 – Russian digital misinformation interference in US election

**'The bleakest of worlds': how Molly Russell fell into a vortex of despair on social media**

London teenager died from an act of self-harm in 2017 after the darker side of online life overwhelmed her

2017 – Molly Russell takes her own life after exposure to self-harm content

# An incomplete history of our online content landscape

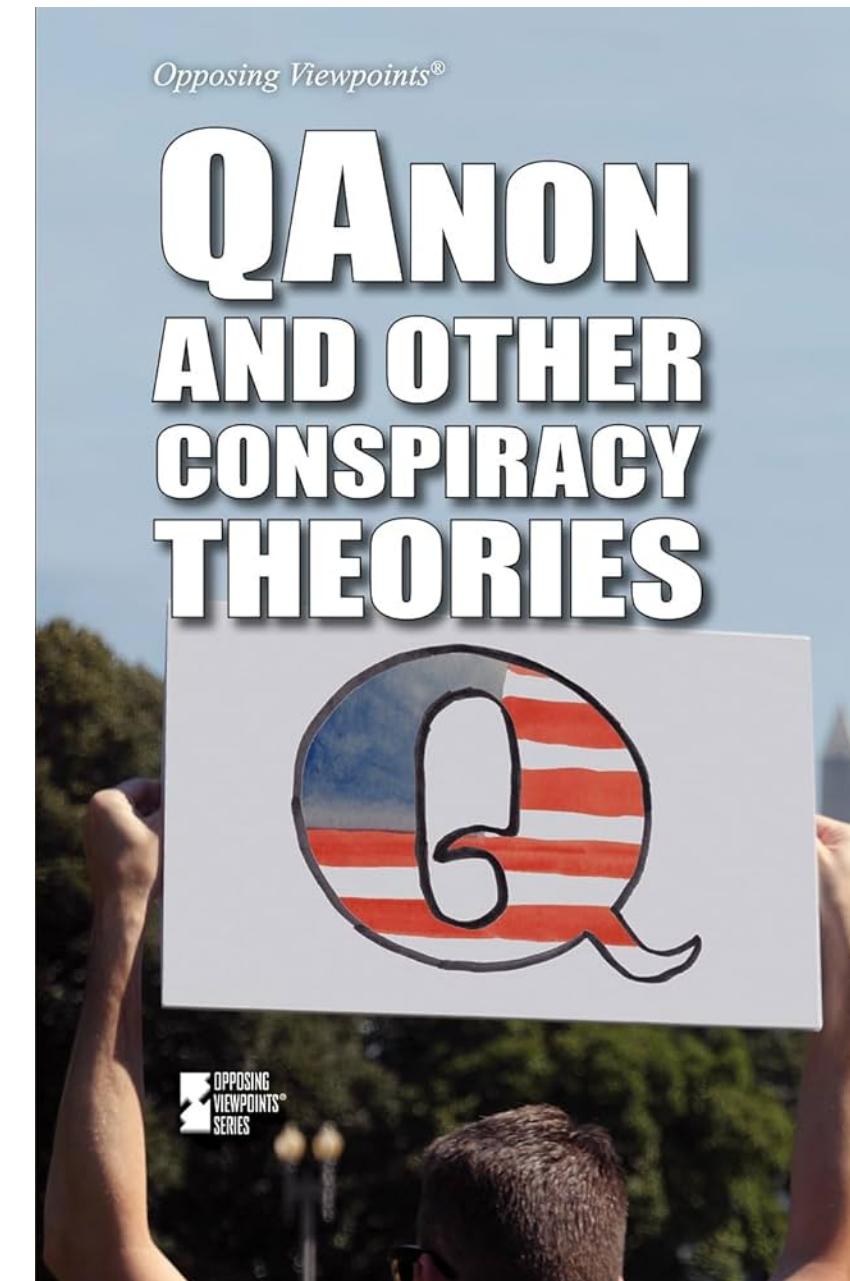


2020 – COVID-19 Infodemic

# An incomplete history of our online content landscape



2020 – COVID-19 Infodemic



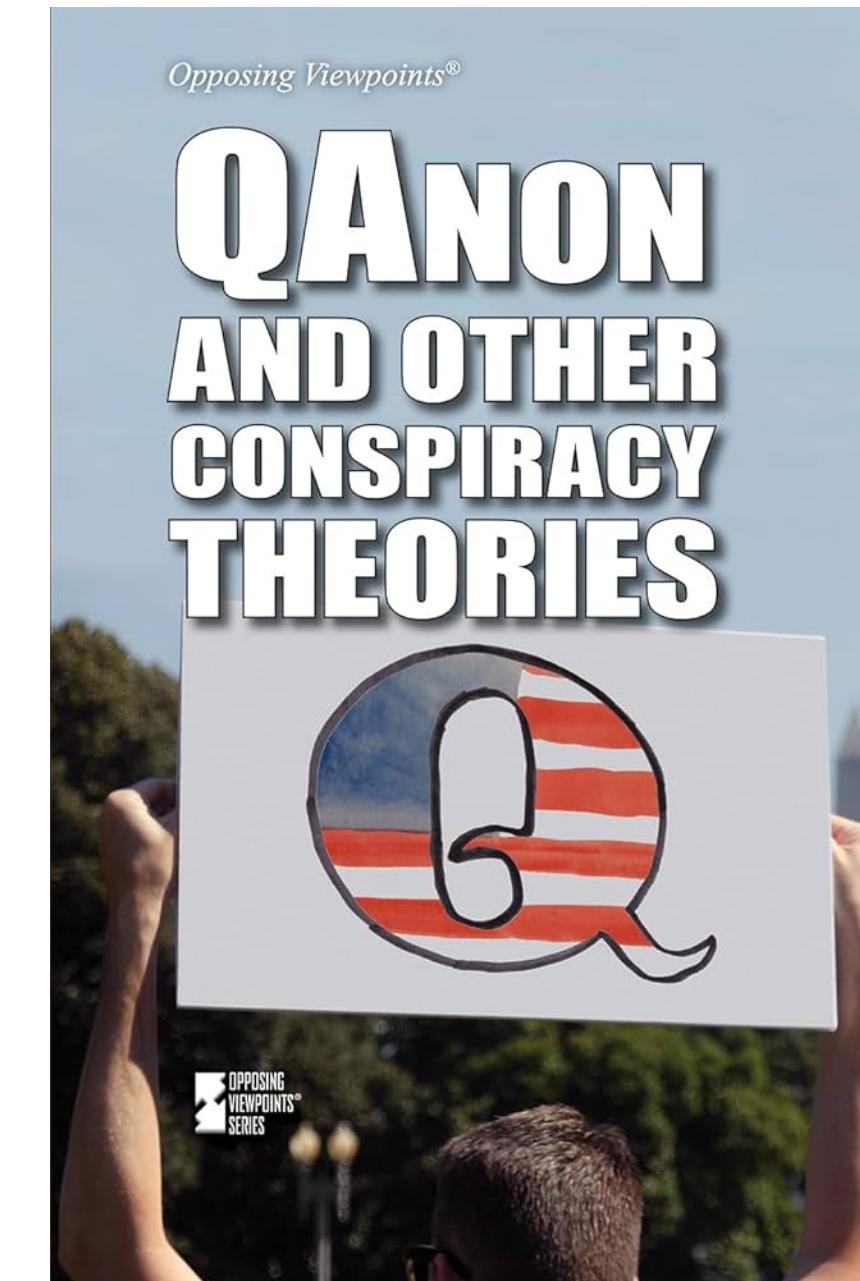
Conspiracy theories proliferate online, in part leading to the insurrection at the US Capitol on Jan 6, 2021



# An incomplete history of our online content landscape



2020 – COVID-19 Infodemic



Conspiracy theories proliferate online, in part leading to the insurrection at the US Capitol on Jan 6, 2021



# Online Safety: 1990 – 2024



# The world is on fire – and we are to blame

What are we going to do about it?

- Central questions of this course (and of the world)
  - What are the sociotechnical problems that are exacerbated by or enabled by technology?
  - What is our role as computer scientists in tackling these problems?
  - What are the promises and the limitations of using technology to solve these issues?
  - How should computer scientists study and understand these problems today?

# Course Logistics

**This is a research oriented class.**

# Structure of this course

- Attendance (5% of your grade)
- Reading + Participation (20% of your grade)
- Discussion leading a paper with the class (25% of your grade)
- Term project (50% of your grade)

# Attendance

- Attendance is **mandatory** – it will be checked every class within the first 10 minutes of the seminar
  - You can miss up to 2 classes without telling me
- There will be assigned readings every session starting **10/8**
  - You must have read and be prepared to discuss the material at the beginning of class
  - I will cold-call at the start of each paper discussion – you can skip 3 times per quarter

# Reading + Participation

- There will be assigned readings every session starting **10/8**
  - You must have read and be prepared to discuss the material at the beginning of class
  - I will cold-call at the start of each paper discussion – you can skip *3 times per quarter*

# Paper Presentations + Discussion Leading

- Starting **10/8**, the course structure will move away from lecture and towards a discussion oriented format
  - One student will be assigned as “discussion lead” per paper for that day
  - Expectation is that you will create a presentation to guide discussion that will take approximately 1/2 of class time (~40 minutes)
  - 5 minutes is on summarizing findings of the paper, but swiftly move to discussion
  - **You must send the slides to me and Muchan 3 days in advance of your assigned date or you will take a late penalty (10% off for every 24h)**

# Paper Presentations + Discussion Leading

- To facilitate selection of topics, please fill out this Google Form by next class, **10/2:** <https://forms.gle/mrNDT2D4JQ7X29Fu7>
- *Paper assignments will go out the evening of 10/2, and updated on the website*
- If you don't fill it out, I'll assume you're okay with any topic (and even presenting in week 3)!

# Term Project – Guidelines

- You will form groups of 2 – 3 and work on an original research project over the course of the term
  - Intention is to give you experience *getting a flavor* for research in this area
- Two options:
  - Original research
  - “On-Track” research
    - Choose a platform, choose a harm, go and study the dynamics of that harm on that platform

# Term Project – Timeline

- Project spec will be released by 10/2
- Milestones
  - Project Intention Document – Due **10/8 by 12:30pm**
  - Midpoint Check-in and Meeting – Due **11/5 by 12:30pm**
  - Final Presentation – Last week of the quarter
  - Final writeup – Due **12/11 Anywhere-on-Earth**

# Term Project – Checkins

- Every Thursday session will only have one paper to present – the rest of the time, you will meet with your groups and I'll go around to chat for 5-minutes about progress + status updates
- Informal check-ins where we can talk about any progress, any blockers, and brainstorm more about your idea
- *Simulates what a weekly meeting with me might look like*
- If you are already doing research broadly in this area, **talk to me to see if it will fit for the requirements of this course.**

# Course Resources

- Class website: <https://kumarde.com/cse291-fa24>
- Canvas: <https://canvas.ucsd.edu/courses/58787>
- Piazza: [https://piazza.com/ucsd/fall2024/cse291\\_fa24\\_d00/home](https://piazza.com/ucsd/fall2024/cse291_fa24_d00/home)
- Syllabus: <https://docs.google.com/document/d/1Fk-cdqF1cGAiIY5fGkGO0CfwDKOxPX44snAgaFPr4w8/edit>

# Campus Resources

- The material in this class can be emotionally challenging to engage with
  - CAPS – free confidential psychological counseling and crisis for all UCSD Students
  - Self Help Library (through CAPS)

# Immediate Logistics

- By 10/2 @ 12:30pm PT
  - Fill out topic preferences document: <https://forms.gle/mrNDT2D4JQ7X29Fu7>
  - Project spec released for term project
- By 10/1
  - Fill out “First Day Survey” <https://canvas.ucsd.edu/courses/58787/quizzes/188179>
- By 10/8 @ 12:30pm PT
  - Project intention document due for term project (group selection + general research idea)