



The University of Texas at Austin
School of Journalism and Media
Moody College of Communication

February 2023

Platform Accountability

Dhiraj Murthy, Ph.D.
Professor, The University of Texas at Austin



THE COMPUTATIONAL
MEDIA LAB



Platform Accountability



How well do **social media** platforms manage, moderate and curate content online and prevent/remove harmful content and user accounts?



What are Social Platforms?



- Profile-based
- Friends, followers, and feeds
- Comments and posts
- Social networks
- Social interaction
- Venues to build relationships
 - Public or Private



It's not always obvious...



- Messaging platforms
 - WhatsApp
- Video sharing platforms
 - YouTube
- 'Authentic' sharing platforms
 - BeReal
 - Snapchat
- Payment platforms
 - Venmo



Why Accountability?

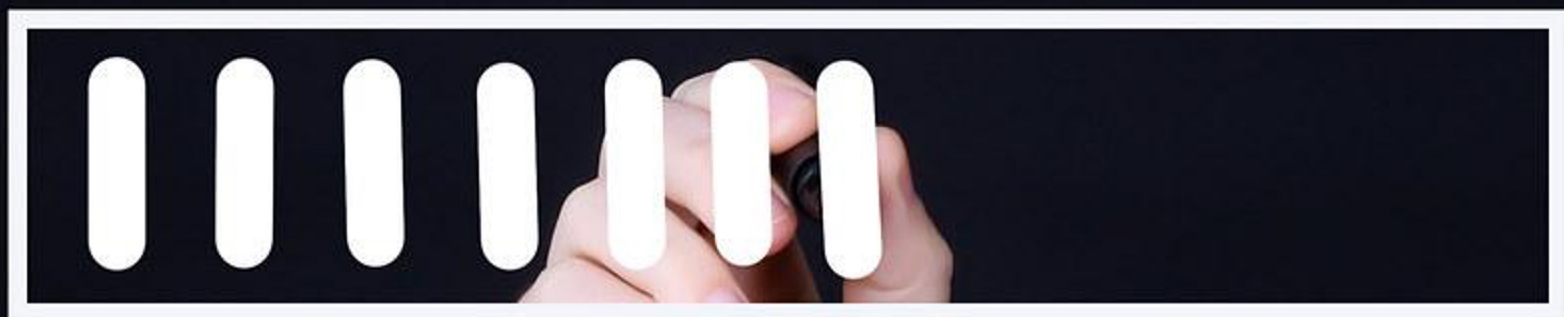


THE WHITE HOUSE
WASHINGTON

“Although tech platforms can help keep us connected, create a vibrant marketplace of ideas, and open up new opportunities for bringing products and services to market, they can also divide us and wreak serious real-world harms.”

Source: Readout of White House Listening Session on Tech Platform Accountability 09/22

ACCOUNTABILITY



LOADING...

Why is Accountability So Tough?





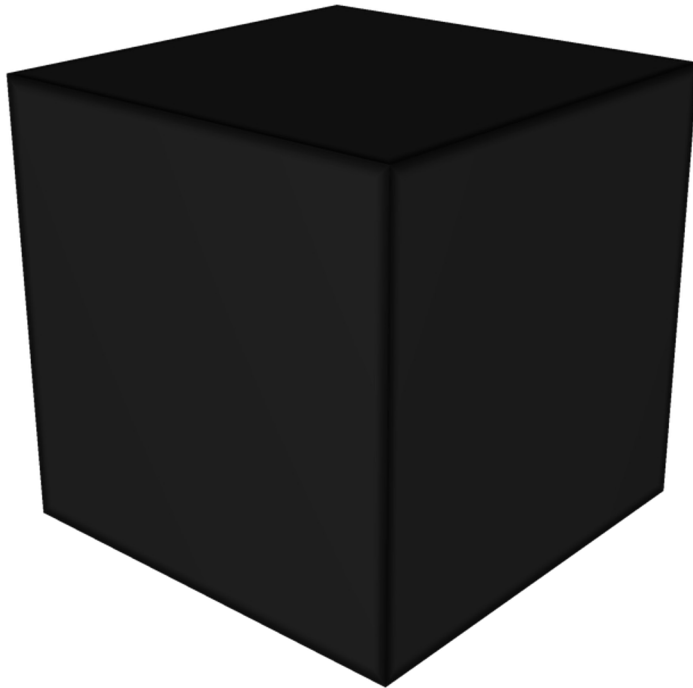
Accountability is Tough



- Why is it so difficult to conduct accountability assessments of tech companies?
 - Algorithms are complex; and confusing (even to experts)
 - Code-based barriers to transparency



Algorithms



MOST Social Media Algorithms are protected or “proprietary”

- A Black Box

What does this mean for accountability?

- “Algorithms are not neutral” (Kitchin, 2017)



Algorithms



Lack of transparency:

Algorithms generally cannot be scrutinized by anyone outside the platforms themselves

Creates barriers to meaningful accountability



Critical Algorithm Studies



- Why are algorithms problematic?
 - Can adversely impact online and offline life
 - What makes them problematic?

(Fuller, 2008; Feuz et al., 2011; Gillespie, 2014; Kitchin & Dodge, 2011; Ruppert et al., 2013)

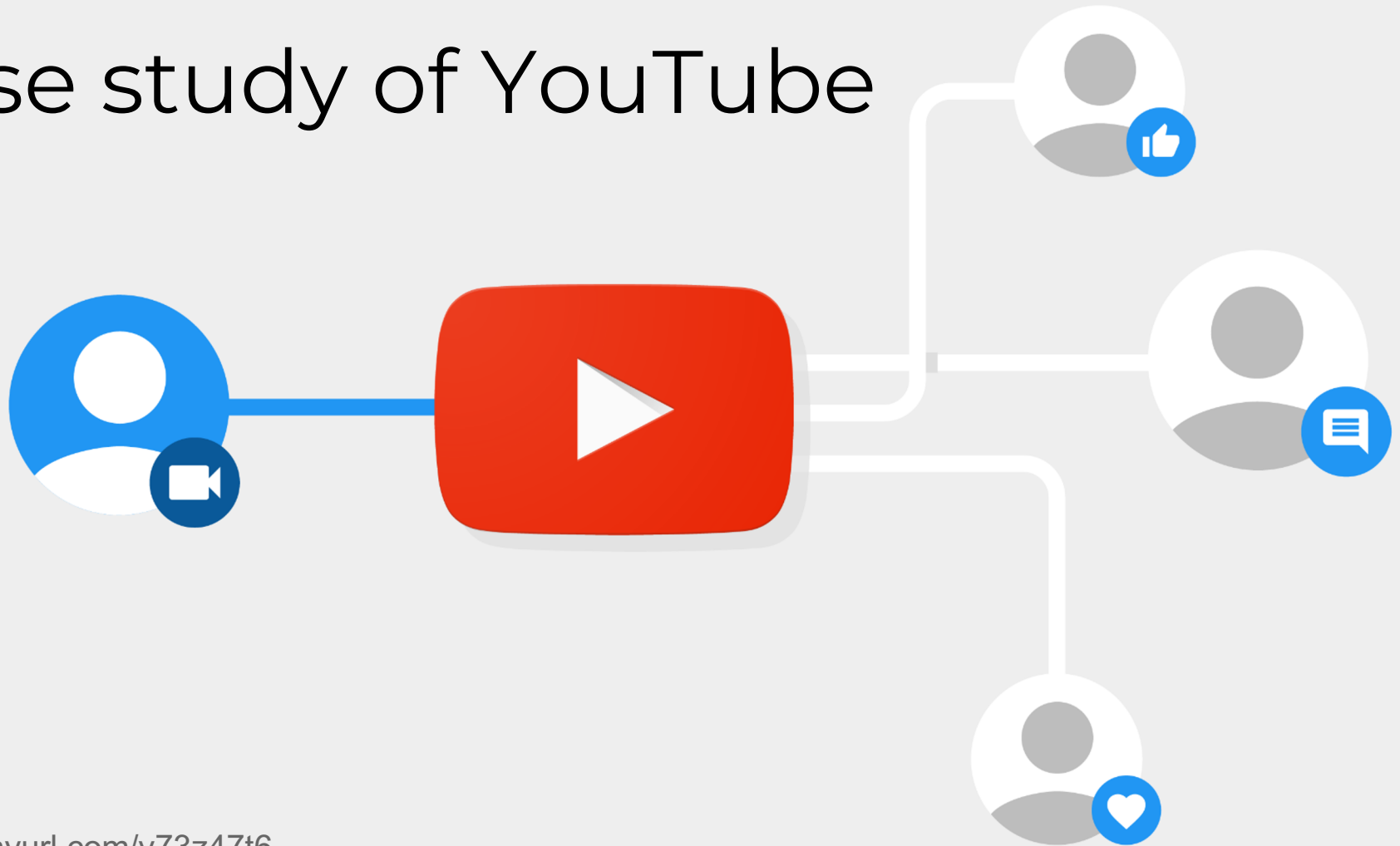


Critical Algorithm Studies



“...seeks to open the black box of processors and arcane algorithms to understand how [...] lines and routines of code [...] work in the world by instructing various technologies how to act.” (Kitchin & Dodge, 2011)

Case study of YouTube



<https://tinyurl.com/v73z47t6>



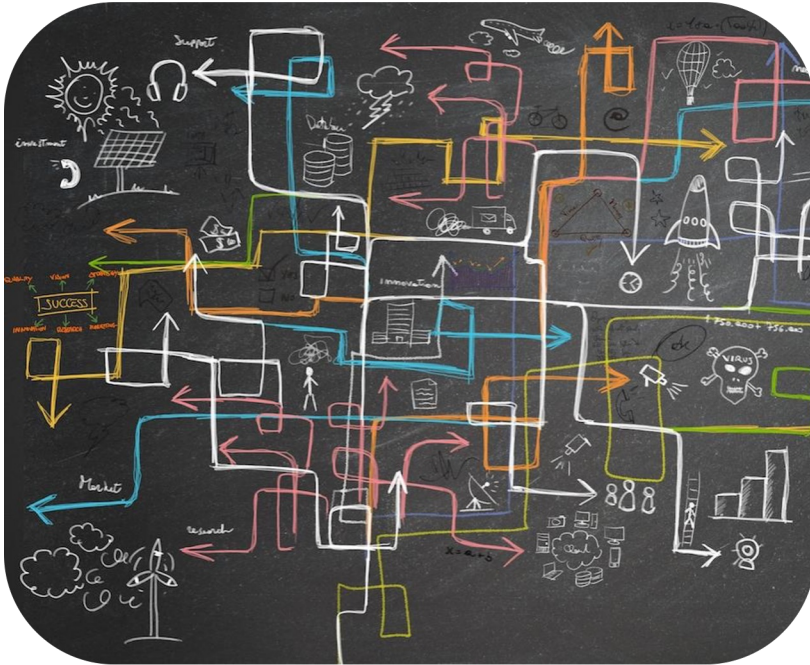
YouTube Algorithm Always Changes



YouTube states, “we run hundreds of experiments each year to make this and every other discovery feature better [...] and current algorithms will change over time.”



YouTube's Algorithms are Complex



Christos Goodrow, an engineering director at YouTube, noted

“...it takes over **1 million lines of code** to figure out what videos to recommend,”
(Computerphile, 2014).



YouTube's Algorithms Made Content Accessible

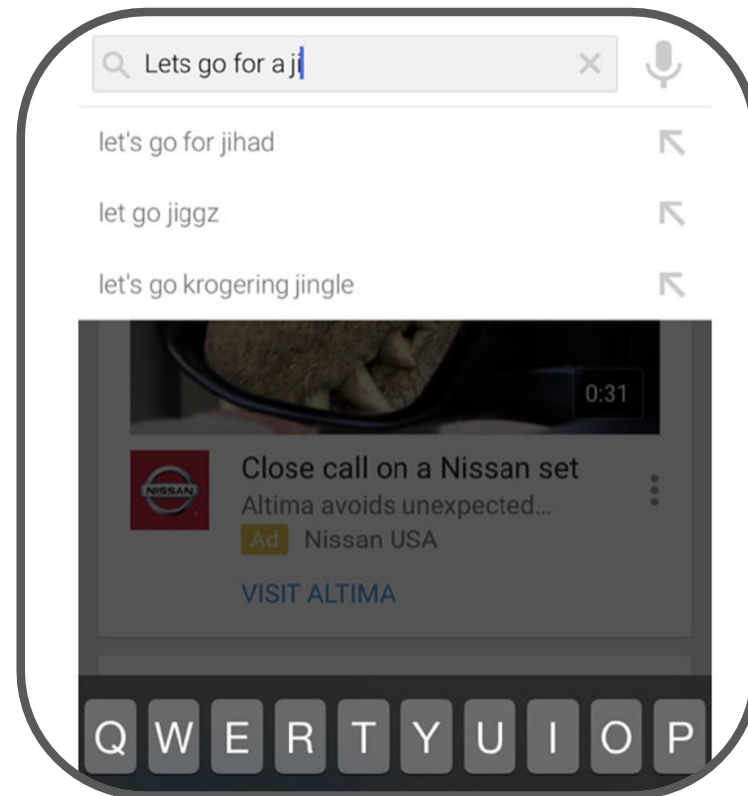


YouTube “remain[ed] easy to find content on YouTube that violate[d] the company’s community guidelines against hate speech and/or explicitly promote[d] terrorism” (Neumann, 2013, p. 442).

Neumann, P. R. (2013). Options and strategies for countering online radicalization in the United States. *Studies in Conflict & Terrorism*, 36(6), 431-459.



YouTube's
algorithms even
suggested extreme
content





Calls for Platform Accountability



Former U.S. President Barack Obama suggested that technology companies need to work to silence ISIS online and make “alternative accounts” to ISIS’ narrative available

How did I do an accountability study of YouTube?

ACCOUNTABILITY



YouTube Accountability



How can we hold YouTube accountable?

- Despite use of a **Black Box** Algorithm

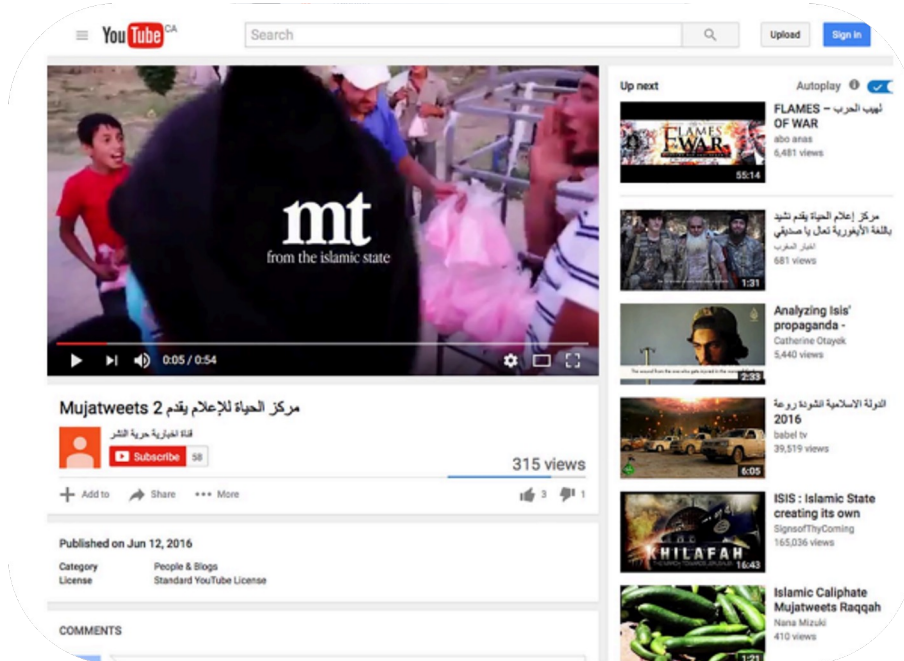
How can we build methods of platform accountability?

- Reverse engineering

Rieder et al. (2018)



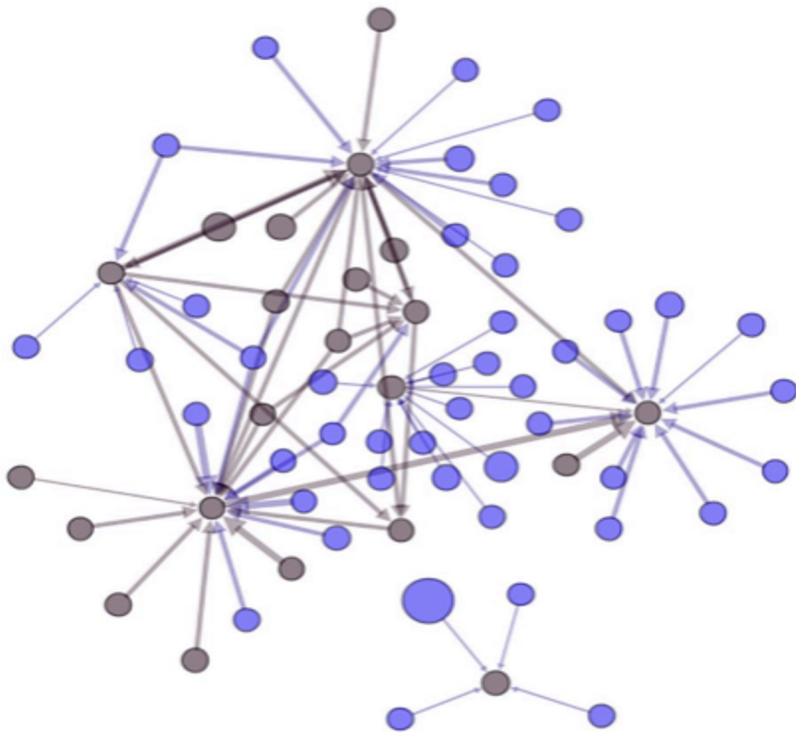
Case Study



- YouTube was a source of extremist content
- ISIS actively recruited through the platform



YouTube Study: Methods



Built a video network from 2016:

- 15,021 videos
- 190,087 video recommendations

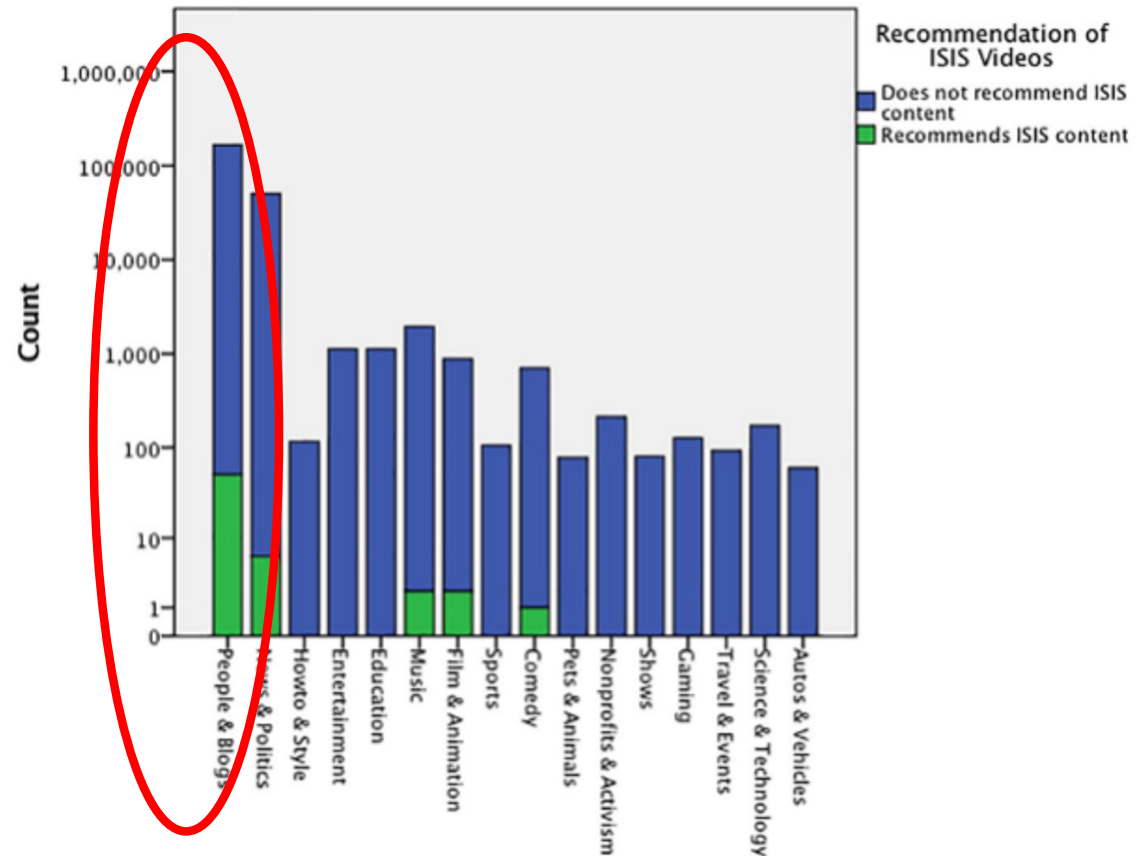
Identified how much ISIS content recommended to users



Some Cues are Easily Identifiable

Most recommending videos:

- Found in “People and Blogs” or “News & Politics”
- Were recent
- Were newscasts
- Were in English

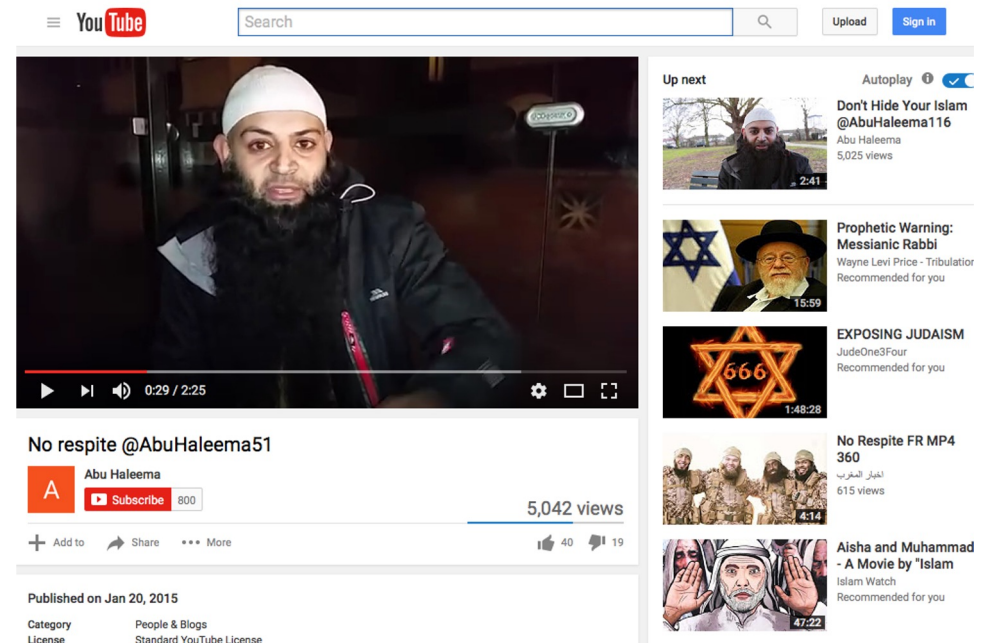




Some Cues are Easily Identifiable

Most videos recommending ISIS content on YouTube had easily identifiable :

- Radical keywords
- Words/phrases from ISIS videos (e.g., 'no respite')



Did YouTube
improve?





YouTube's accountability efforts



In 2017, YouTube and other technology companies launched a major offensive:

Global Internet Forum to Counter Terrorism (GIFCT)

- “...to prevent terrorists and violent extremists from exploiting digital platforms,” (GIFCT, 2019)

Collective mechanism to increase platform accountability

- Combating extreme content online via human moderation and machine-based methods



YouTube Moderated Far More After Pressured



- Calls to Action led to moderation (e.g., GIFCT)
- Much less extremist content was available
- Through human moderation and machine-based detection, **YouTube** sought to be **more accountable** of the **content** it was **hosting**



BUT: Best in class platform accountability groups being disbanded

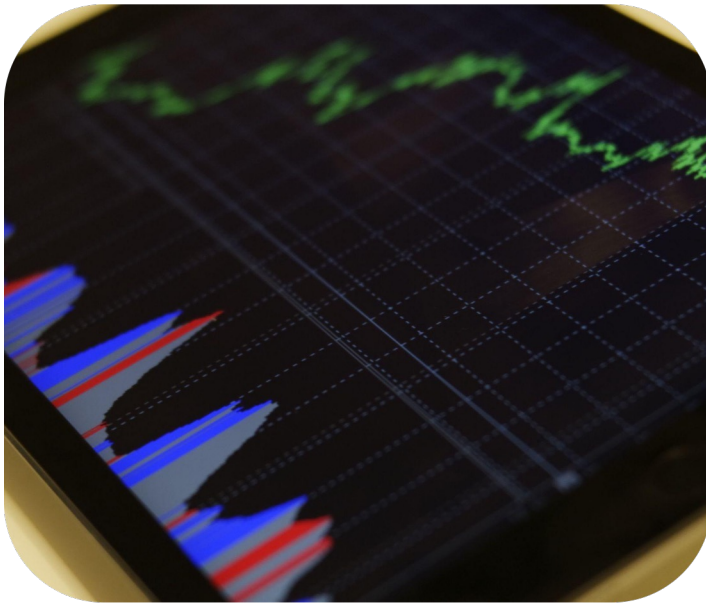
- Musk Fired Twitter's 'Ethical AI' Team
- Therefore, we can't leave moderation to tech platforms

<https://www.wired.com/story/twitter-ethicaal-ai-team/>





Accountability: Risks



- Making platforms accountable may lead to:
 - More black boxing
 - Moves from public to private (e.g., Twitter)
 - More products being end-to-end encrypted



Takeaways



- Pressure on platforms can bring change
- Relatively small changes can eliminate many threats
- Creative reverse engineering makes is a good platform accountability method
- Platform accountability should be longitudinal (including historical)



The University of Texas at Austin
School of Journalism
Moody College of Communication



WHAT STARTS HERE CHANGES THE WORLD

Special Thanks:

Kami Vinton, M.S.

Doctoral Candidate

School of Journalism and Media

The University of Texas at Austin

GRA, The Computational Media Lab

River Terrell

Senior, Honors Biology

Science Communication

The University of Texas at Austin

RA, The Computational Media Lab



The University of Texas at Austin
School of Journalism
Moody College of Communication



WHAT STARTS HERE CHANGES THE WORLD



Dhiraj Murthy, Ph.D.

Professor, The University of Texas at Austin

Moody College

School of Journalism and Media

Founder, The Computational Media Lab

<https://www.computationalmedialab.com>

dhiraj.murthy@austin.utexas.edu