IN4MATX 251: CSCW

Class 3:
Feeds and their Algorithms

Daniel Epstein

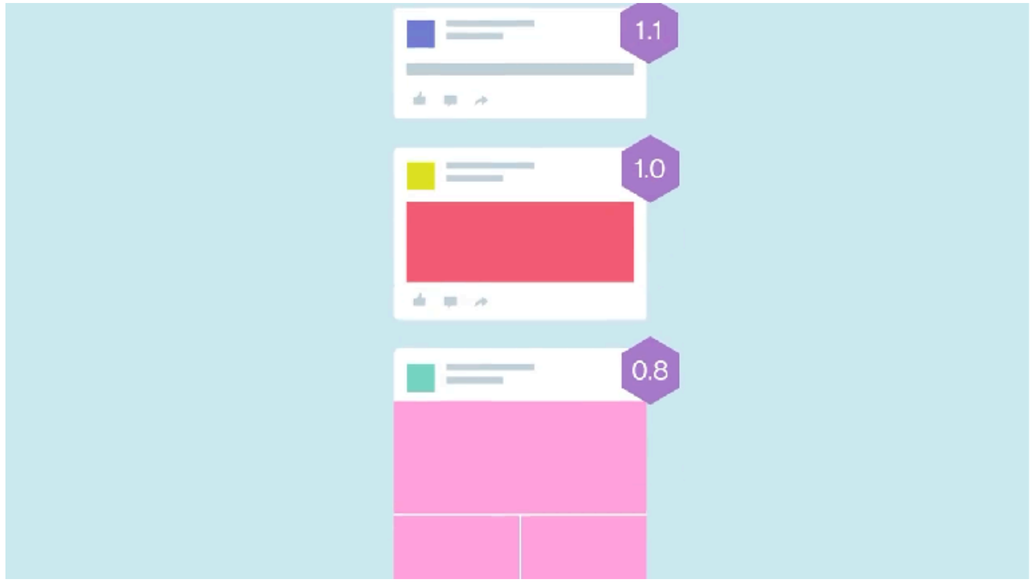
Today’s goals

By the end of today, you should be able to…

- Articulate what a feed is, and how it is represented in different social media sites
- Describe the importance of filtering and sorting algorithms to social media content represented in feeds
- Explain how a typical feed algorithm works, at a high level
- Discuss whether feeds create filter bubbles or echo chambers, particularly around online discourse

Two types of feeds

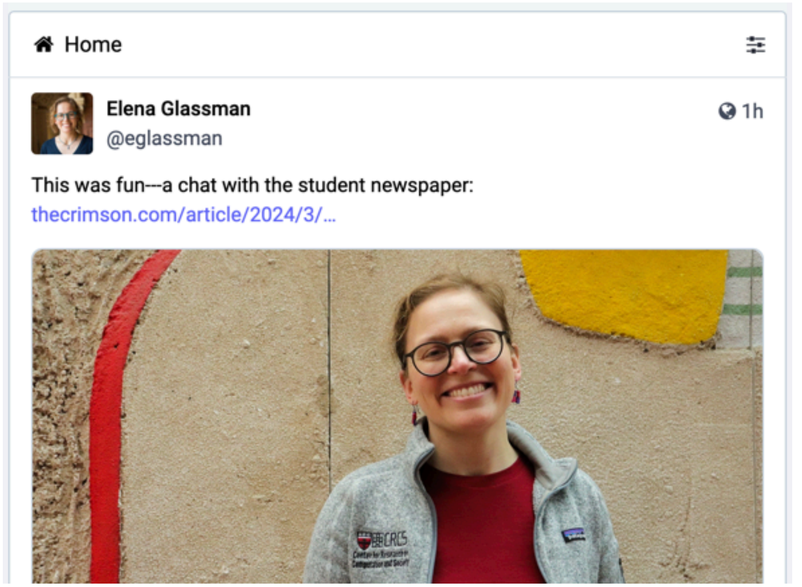
Ranking



Facebook
Instagram
Twitter/X

TikTok
Pinterest
Reddit

Chronological



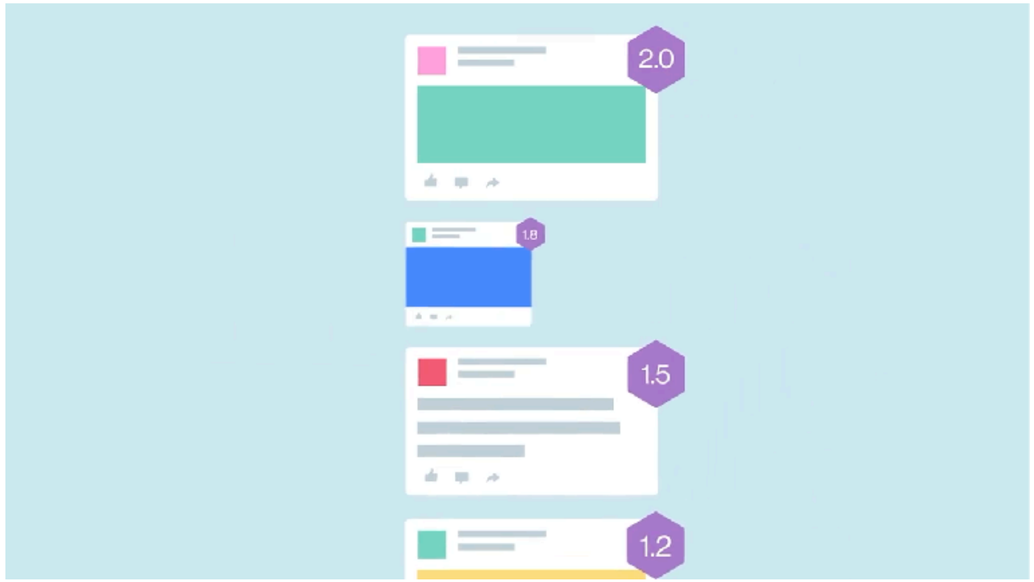
Mastodon
Email
Slack

Discord
WhatsApp
iMessage

Designing for information overload

“Ranking” approach

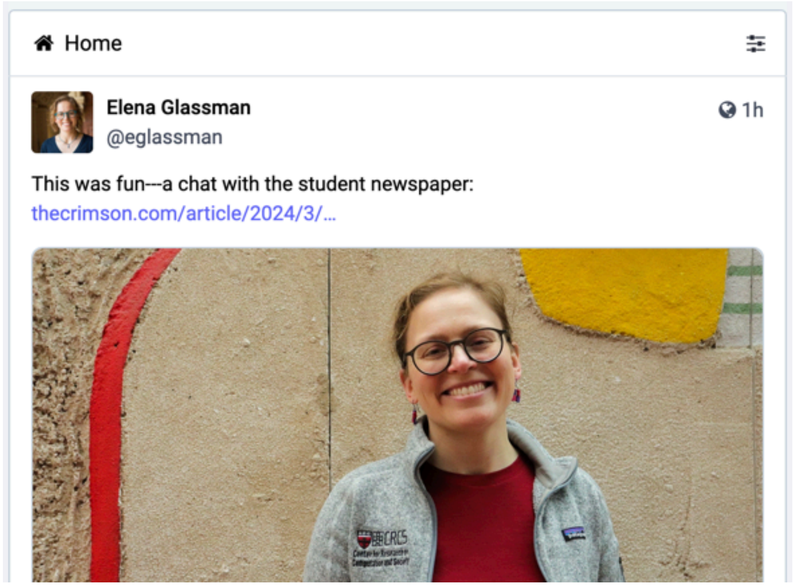
- When it’s working well, it can bring what you want to the front
- But, it’s unintuitive
- You won’t see everything, which others might not expect



Designing for information overload

“Chronological” approach

- Intuitive: follows our expected mental model
- Accounts which post more show up more
  - Burdens the user with decision-making about what to follow or not



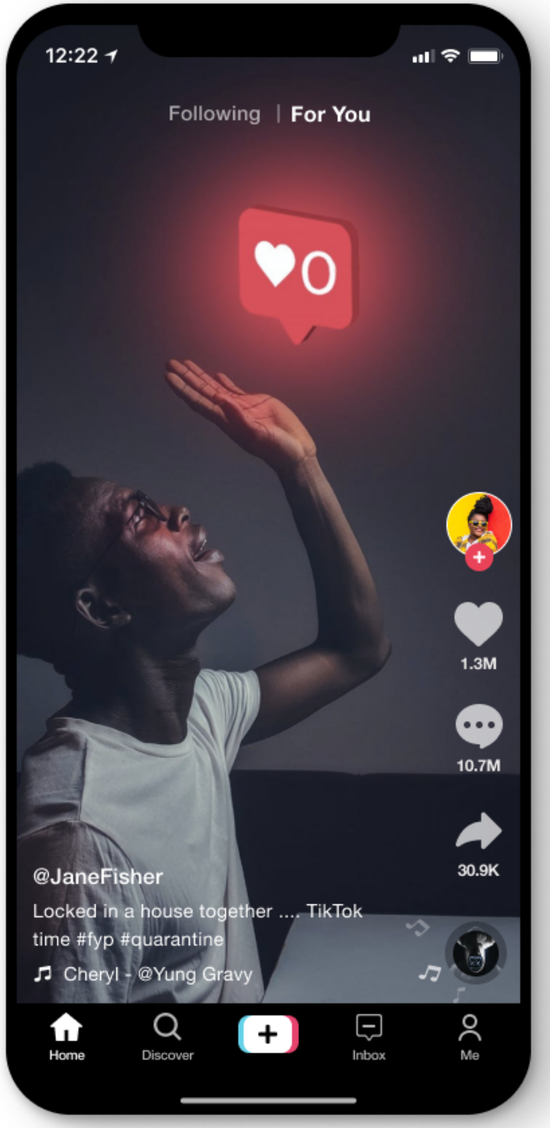
Designing for information overload

- “**Algorithms are unavoidable here.** Even sorting posts by friends in chronological order or videos by overall popularity is algorithmic; and often it is unclear there is a single, simple baseline algorithm.”
  - Dean Eckles, MIT, to the US Senate
  - Hearing on “algorithmic transparency and assessing effects of algorithmic ranking”

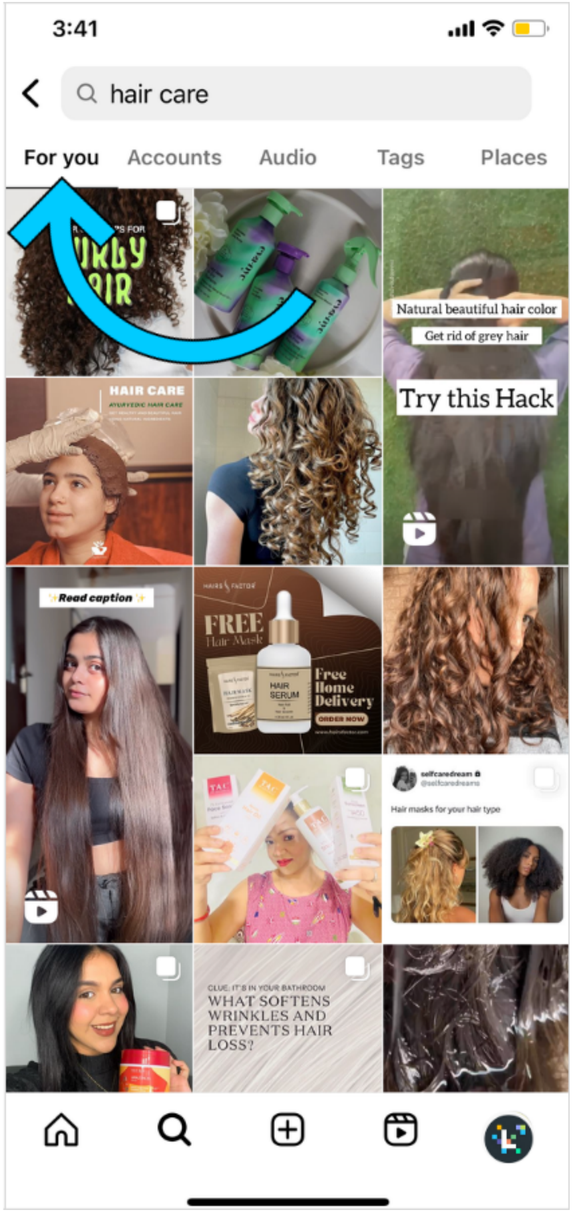
https://www.commerce.senate.gov/services/files/62102355-DC26-4909-BF90-8FB068145F18

How algorithmic feeds work

Feeds


12:22
Following | For You
1.3M
10.7M
30.9K
@JaneFisher
Locked in a house together .... TikTok time #fyp #quarantine
Cheryl - @Yung Gravy
Home
Discover
Inbox
Me


Home
What's happening?
Home shows you top Tweets first
See latest Tweets instead
You'll be switched back Home after you've been away for a while.
View content preferences
Matt Lees Retweeted
Tom Neenan @TNeenan · Mar 1
Important content.
GET THE CLEANEST THUMBS IN THE LAND
3
20
122
Jesse Cox @JesseCox · 45s
Oof...
well the stream is live, but I'm dropping frames like crazy.
We'll see how this goes.

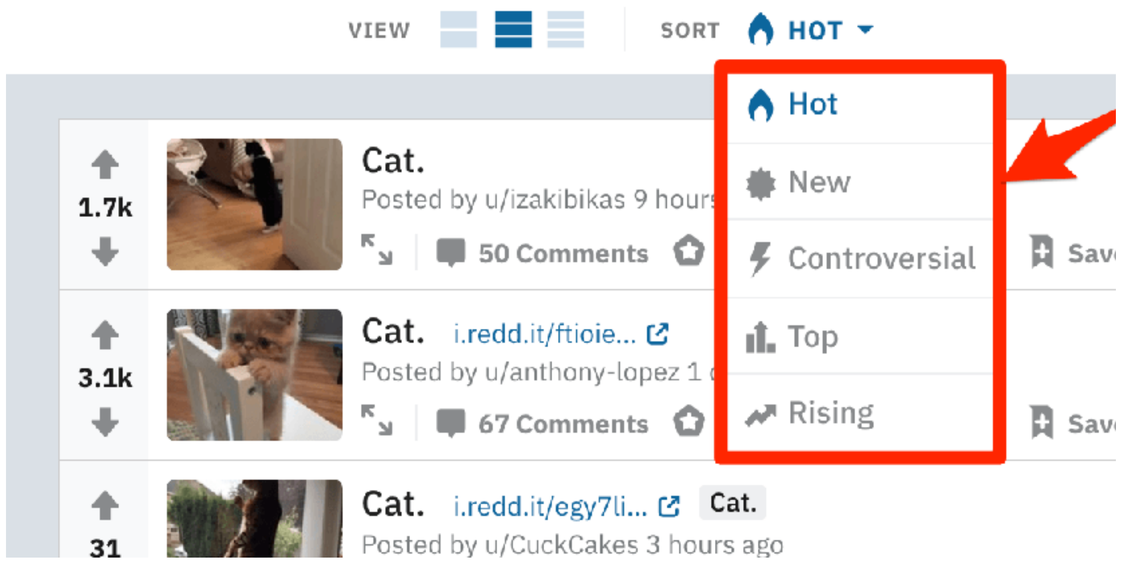

3:41
hair care
For you
Accounts
Audio
Tags
Places
Natural beautiful hair color
Get rid of grey hair
Try this Hack
HAIR CARE
Read caption
FREE
Free Home Delivery
WHAT SOFTENS WRINKLES AND PREVENTS HAIR LOSS?

Feeds

- Two approaches:
  - Global ranking
  - Personalized feed

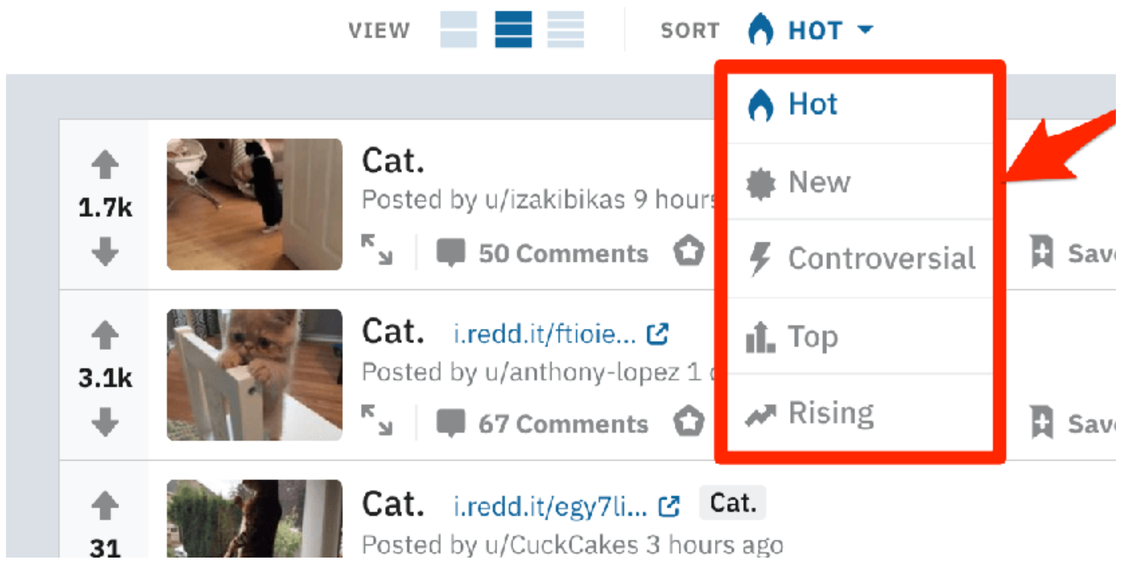
Global ranking

- Bring the most highly-rated posts to the top
- Think Reddit's "hot" ranking. Other sites like Hacker News do similar things



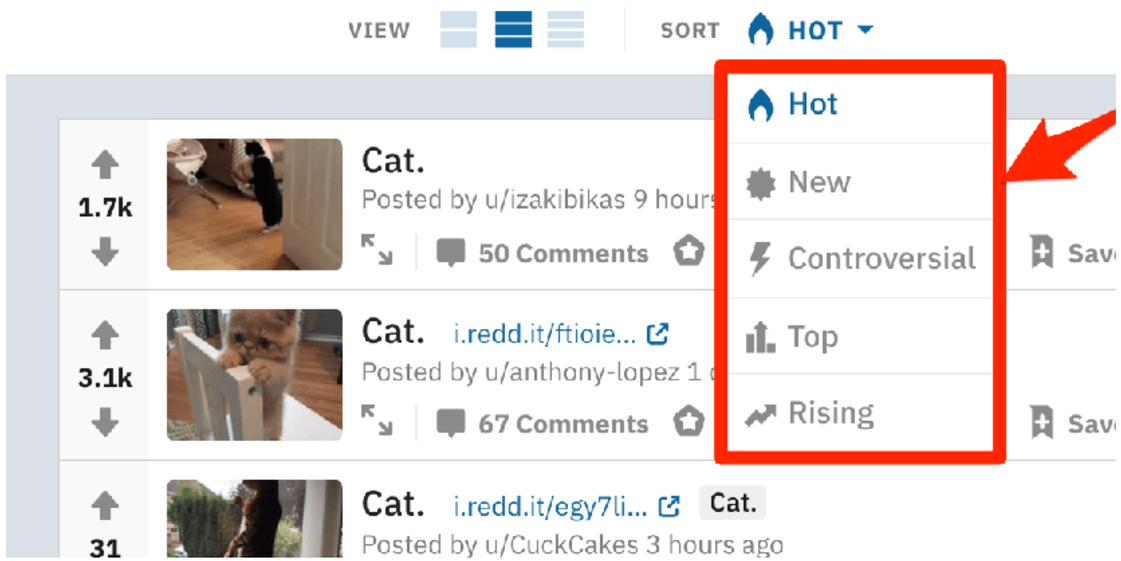
Global ranking

- First attempt: rank posts by the number of upvotes
  - E.g., a post with 1,000 upvotes should be rated more highly than a post with 100 upvotes
- What falls short here?
  - If lots of people saw the post, but a large percentage disliked it



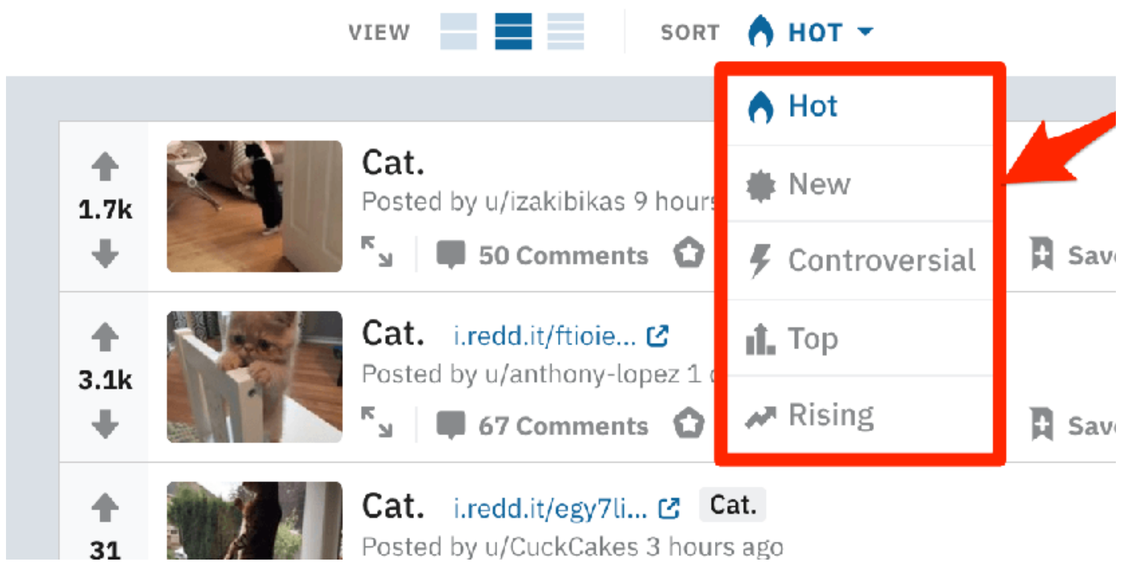
Global ranking

- Second attempt: rank posts by the number of upvotes - the number of downvotes
  - E.g., a post with 100 upvotes and 1 downvote should be rated more highly than a post with 1,000 upvotes and 1,000 downvotes
- What falls short here?
  - This ranking is static, the top posts would rarely change



Global ranking

- Final attempt: decay rankings over time
  - log(max(upvotes-downvotes, 1))
  - **Why log?**
- And, decay the log score over time
  - Platforms do this differently, Reddit linearly decays the ranking over time

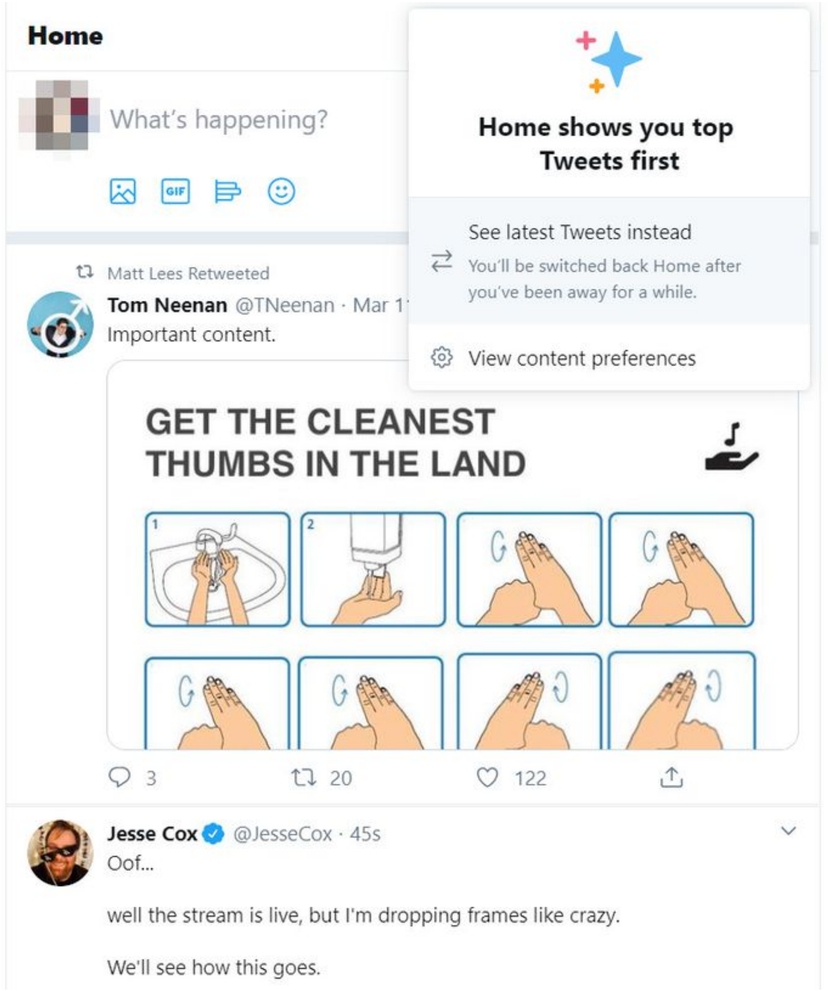


Personalized Feed

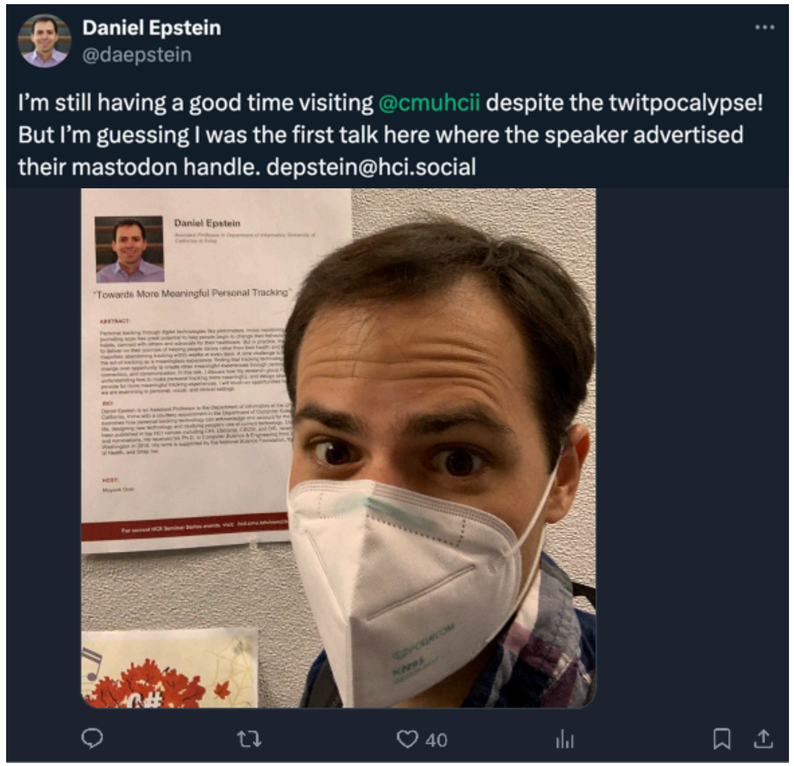
- Think TikTok's for you, Facebook or Twitter/X's feed
- Leverages some version of machine learning

Personalized Feed

- Four steps:
  - Featurize
  - Predict
  - Calculate Objective
  - Rank

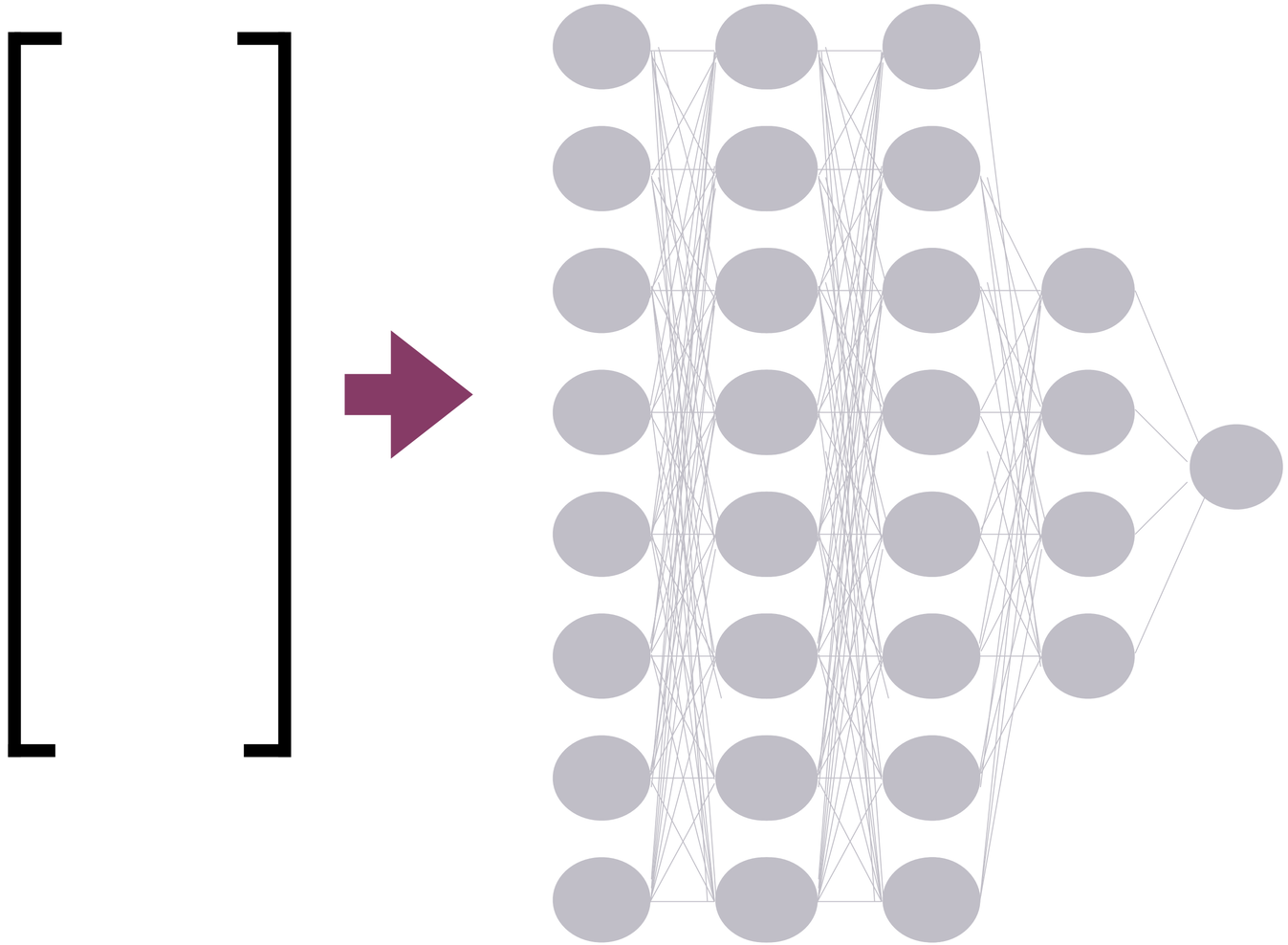


Personalized Feed: Featurize



Tie strength: 8
Contains image: yes
Image contains face: yes
Other objects: mask, paper
Positive keywords: 1
Negative words: 1
Platform: iPhone
Interactions so far: 40
Minutes since post: 52
Internet speed: 10 mbps
…

Personalized Feed: Predict



Probability of clicking “like”

P(like)

P(comment)

P(stop scrolling)

P(watch)

P(click)

P(share)

P(block)

Personalized Feed: Predict

- How do we train these deep learning algorithms?
- Training data: prior behavior on the platform
  - The system inputs when you scroll, click, etc., and uses it to predict behaviors towards future posts you haven't seen before
  - Or, others like you. If someone else clicked on a similar link, they might share other interests/qualities, and those interests can help predict your interests

Personalized Feed: Calculate objective

- So, what do we do with these predictions?

P(like) P(stop scrolling) P(click) P(block)

P(comment) P(watch) P(share)

- We define an **objective function**: a formula to combine and weight the predictions
  - How many points do you think each predicted behavior should get?

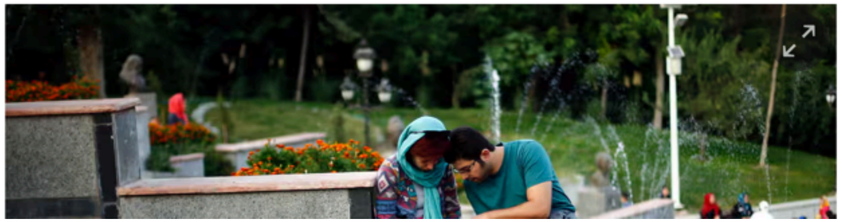
$$\sum_{p \in \text{predictions}} \text{weight}_p \cdot p$$

Quick tangent: objectives in practice

- Facebook: “Meaningful Social Interactions”
  - A weighted average of likes, reactions, reshares, and comments

Facebook overhauls News Feed in favor of 'meaningful social interactions'

Refresh of the News Feed algorithm will de-prioritize content shared by media and businesses in favor of that produced by friends and family, Zuckerberg says



Interaction type	weight
Like	1
Reaction	1.5
Reshare	1.5
Comment	15-20

https://www.theguardian.com/technology/2018/jan/11/facebook-news-feed-algorithm-overhaul-mark-zuckerberg
https://knightcolumbia.org/content/understanding-social-media-recommendation-algorithms

Quick tangent: objectives in practice

- Twitter's open-sourced algorithm:

75 points if predicted that, if you reply, the author will reply back

27 points if predicted that you'd reply

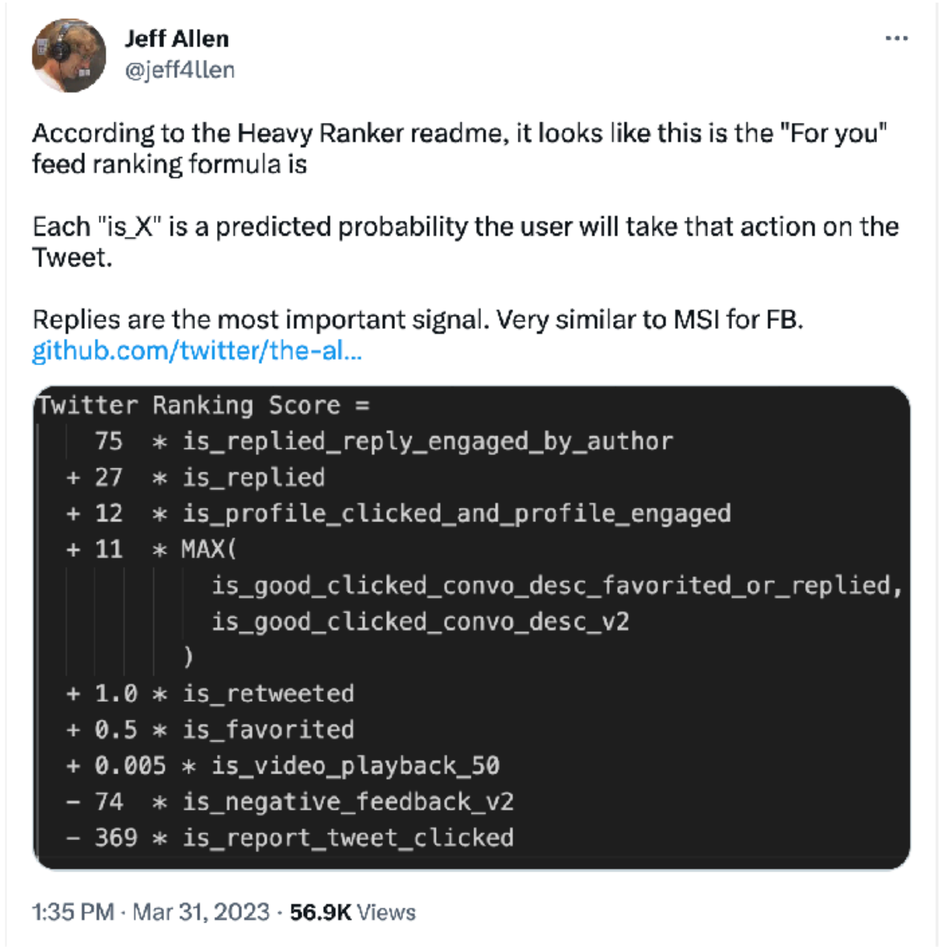
12 points if predicted that you engage with the author's Twitter profile

1 point if predicted that you'll retweet

0.5 points if predicted that you'll favorite

-74 points if predicted that you'll give negative feedback ("not interested", mute, block)

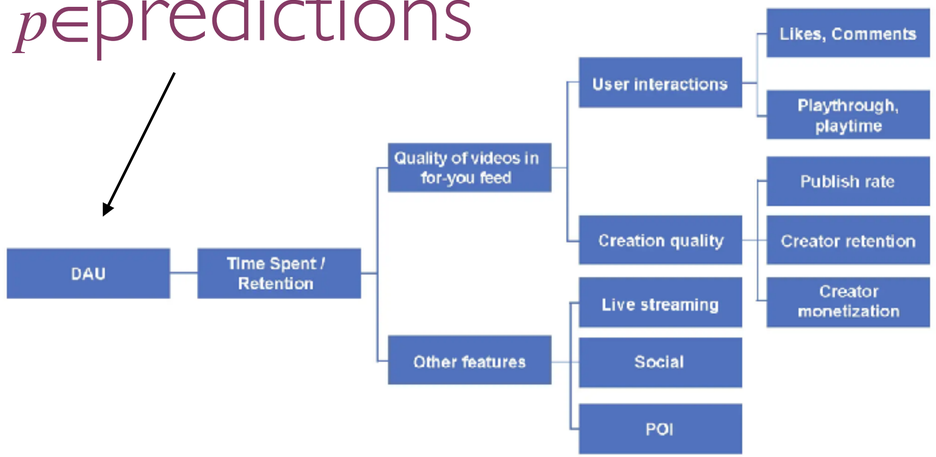
-369 points if predicted that you'll report it



Quick tangent: objectives in practice

- TikTok: some combination of viewing, liking, and commenting

$$\sum_{p \in \text{predictions}} \text{weight}_p \cdot p$$



THE MEDIA EQUATION

How TikTok Reads Your Mind

It's the most successful video app in the world. Our columnist has obtained an internal company document that offers a new level of detail about how the algorithm works.

Share full article

https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html

Discuss: Main takeaways from the Eslami et al., 2015 paper?

But, how do people think feeds work?

But, how do people <u>think</u> feeds work?

- In 2015: unaware
- Participants got upset that posts by close friends and family members weren’t shown
- People attributed decisions to friends excluding them rather than the algorithm
- But, people were mostly satisfied with their feeds

Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., ... & Sandvig, C. (2015, April). " I always assumed that I wasn't really that close to [her]" Reasoning about Invisible Algorithms in News Feeds. In Proceedings of the 33rd annual ACM conference on human factors in computing systems (pp. 153-162).

Discuss: If we replicated the study today, how would the results be the same? Different?

Discuss: Main takeaways from the Karizat et al., 2021 paper?

But, how do people think feeds work?

- 2021: Folk theories, particularly as it pertains to identity
- People think the algorithms know about the identities of people posting, and suppress content from minority groups
- So, they push back, individually and collectively

Karizat, N., Delmonaco, D., Eslami, M., & Andalibi, N. (2021). Algorithmic folk theories and identity: How TikTok users co-produce Knowledge of identity and engage in algorithmic resistance. Proceedings of the ACM on human-computer interaction, 5(CSCW2), 1-44.

But, how do people <u>think</u> feeds work?

- But more, co-production is happening. People's experiences with the algorithm, and the content shared, do shape how they perceive their identities
- Belief becomes reality

Karizat, N., Delmonaco, D., Eslami, M., & Andalibi, N. (2021). Algorithmic folk theories and identity: How TikTok users co-produce Knowledge of identity and engage in algorithmic resistance. Proceedings of the ACM on human-computer interaction, 5(CSCW2), 1-44.

Discuss: What is the Identity Strainer Theory? What is Algorithmic Priviledge?

Does all of this mean that feeds are echo chambers?

Feeds as filter bubbles

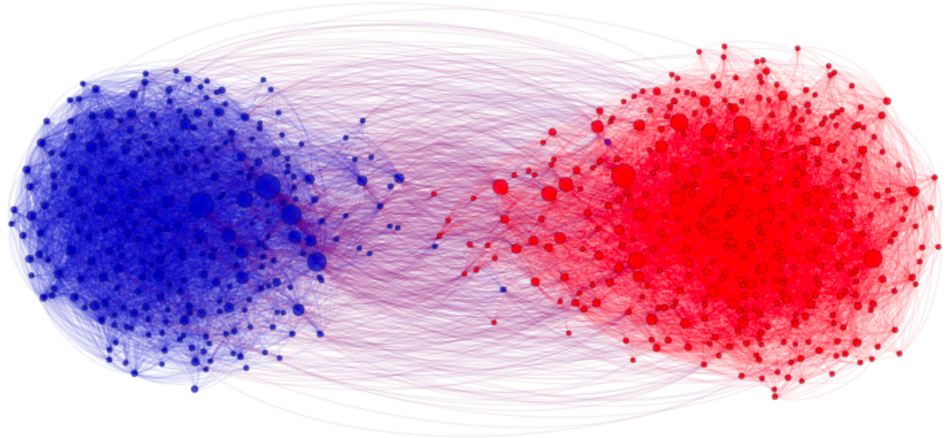
- Filter bubbles occur when people are only shown content that they like
  - This is a natural outcome when your platform optimizes for engagement
- Example: YouTube channels that are slightly less mainstream become recommendation gateways to more and more radical channels

Figure 6: Recommendation graph of YouTube channels.

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A., & Meira Jr, W. (2020, January). Auditing radicalization pathways on YouTube. In Proceedings of the 2020 conference on fairness, accountability, and transparency (pp. 131-141).

Feeds as echo chambers

- If *your* feed only shows you the things that *you* want to see, and *my* feed only shows me the things that *I* want to see…
  - Won't that result in an echo chamber, where we only hear people who share our opinions?
  - Won't this further polarize our society?

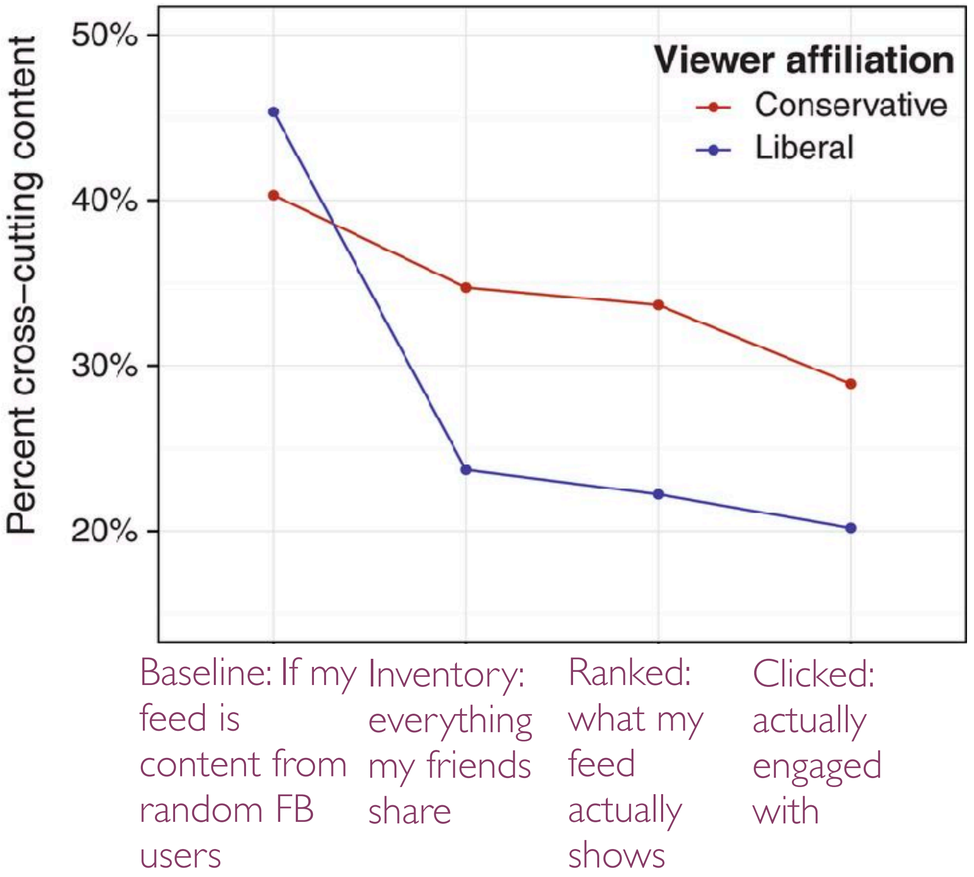


Adamic, L. A., & Glance, N. (2005, August). The political blogosphere and the 2004 US election: divided they blog. In Proceedings of the 3rd international workshop on Link discovery (pp. 36-43).

Feeds as echo chambers

It's more complicated

- From Facebook researchers, a study of log data to understand political news in people's feeds
- The biggest drop is *homophily*: we tend to only friend people who share our views
- The feed itself has a much more minor impact



Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. Science, 348(6239), 1130-1132.

Feeds as echo chambers

It’s more complicated

- People who use social media are exposed to more cross-cutting ideological news than those who don’t
- Having people subscribe to news sources that don’t align with their views have extremely minimal influence on their perspective on those views
- Instead, we might have it backward.
  - It’s not that we’re more polarized because social media only exposes us to *similar* viewpoints, but rather because social media exposes us to a *wider variety* of people

Fletcher, R., & Nielsen, R. K. (2017). Are news audiences increasingly fragmented? A cross-national comparative analysis of cross-platform news audience fragmentation and duplication. Journal of communication, 67(4), 476-498.
Levy, R. E. (2021). Social media, news consumption, and polarization: Evidence from a field experiment. American economic review, 111(3), 831-870.
Törnberg, P. (2022). How digital media drive affective polarization through partisan sorting. Proceedings of the National Academy of Sciences, 119(42), e2207159119.

Summary

- One strategy for managing information overload in social media feeds is to filter what content they show to people
  - Global rankings aggregate up/downvotes, and trail off over time
  - Personalized rankings predict on-platform behaviors, assign weights to each predicted behavior to determine a score, and rank according to that score
- A lot has been said about feeds creating filter bubbles and echo chambers. While there are clearly negative outcomes to feeds, the story is more complicated, and our understanding of the topic is still forming

Today’s goals

By the end of today, you should be able to…

- Articulate what a feed is, and how it is represented in different social media sites
- Describe the importance of filtering and sorting algorithms to social media content represented in feeds
- Explain how a typical feed algorithm works, at a high level
- Discuss whether feeds create filter bubbles or echo chambers, particularly around online discourse

IN4MATX 251: CSCW

Class 3:
Feeds and their Algorithms

Daniel Epstein