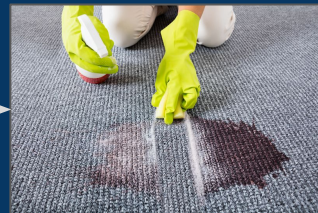
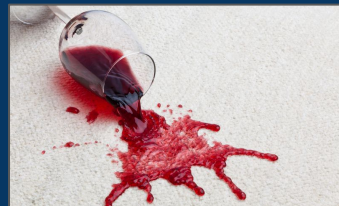




SCHOOL OF INFORMATION CENTER FOR
SOCIAL MEDIA RESPONSIBILITY
UNIVERSITY OF MICHIGAN



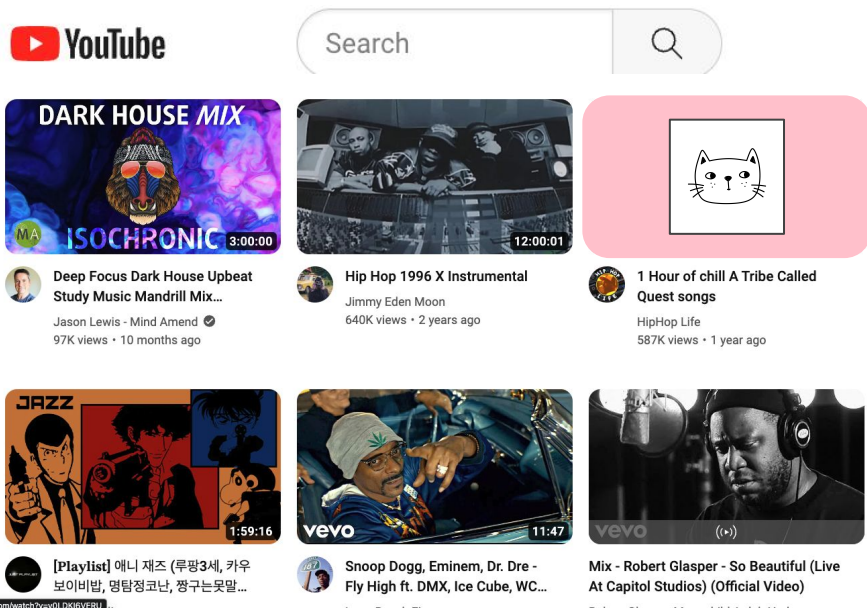
How to Train Your YouTube Recommender to Avoid Unwanted Videos

Alex Liu, Siqi Wu, Paul Resnick

ICWSM 2024
Buffalo, NY, USA



(1) watch a few cat videos



(1) watch a few cat videos

Search

DARK HOUSE MIX
ISOCHRONIC 3:00:00

Deep Focus Dark House Upbeat Study Music Mandrill Mix...
Jason Lewis · Mind Amend ✓
97K views · 10 months ago

Hip Hop 1996 X Instrumental
Jimmy Eden Moon
640K views · 2 years ago

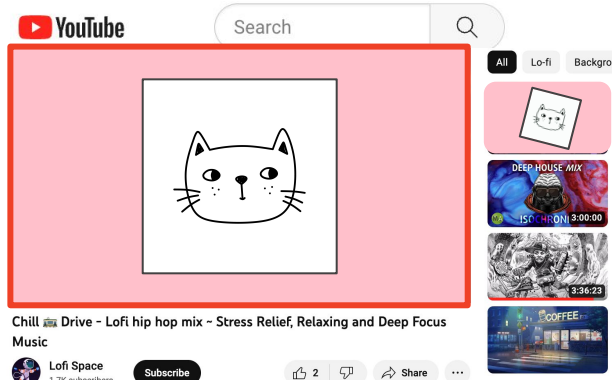
1 Hour of chill A Tribe Called Quest songs
HipHop Life
587K views · 1 year ago

[Playlist] 애니 재즈 (루팡3세, 카우 보이비밥, 명탐정코난, 핑구는웃말...
bombeats
1:59:16

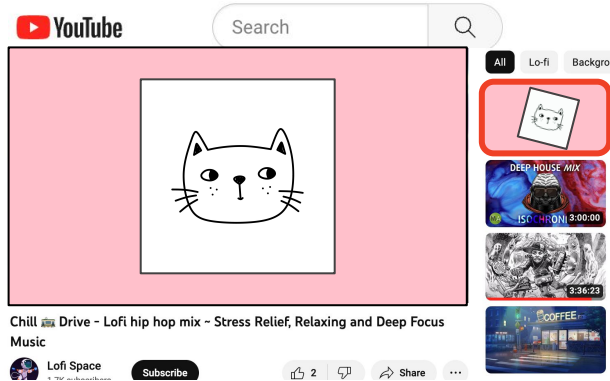
Snoop Dogg, Eminem, Dr. Dre - Fly High ft. DMX, Ice Cube, WC...
Snoop Dogg
11:47

Mix - Robert Glasper - So Beautiful (Live At Capitol Studios) (Official Video)
Robert Glasper
11:47

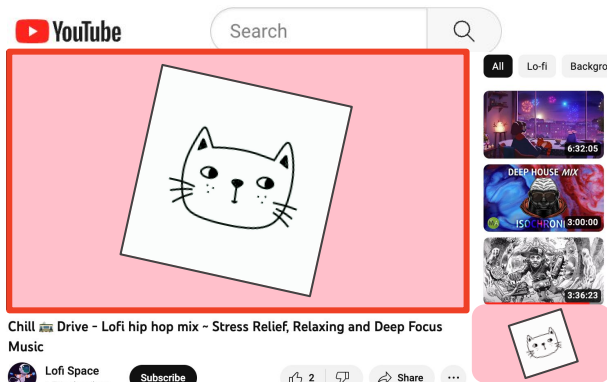
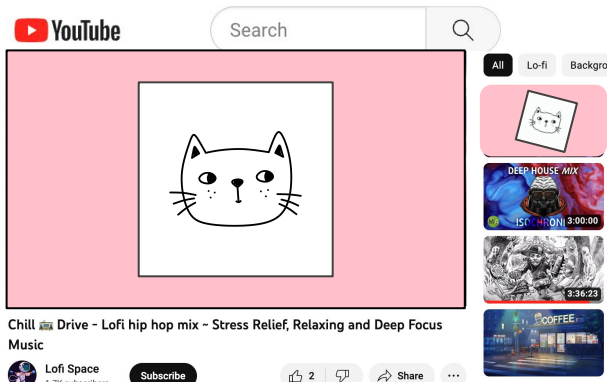
(1) watch a few cat videos



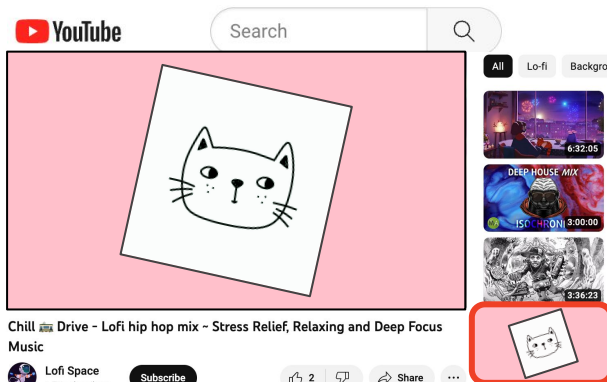
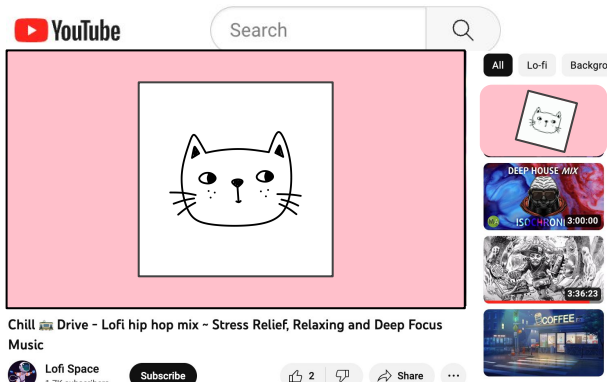
(1) watch a few cat videos



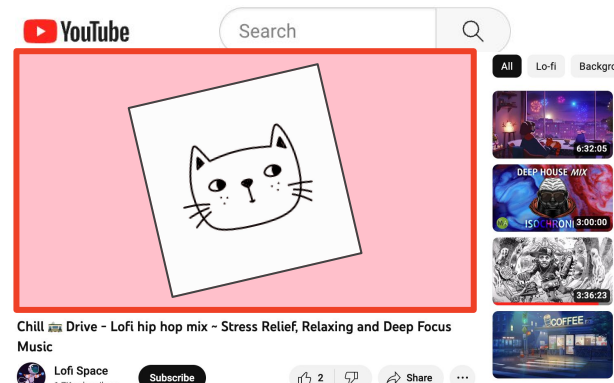
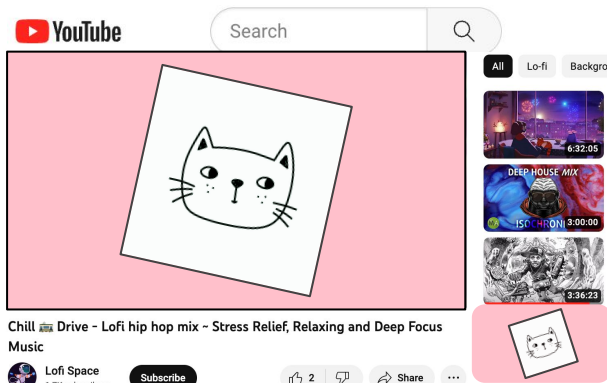
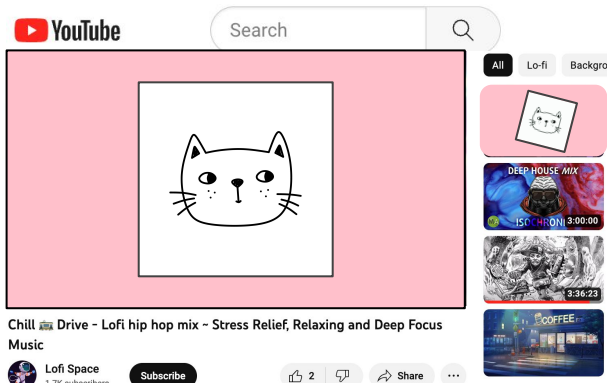
(1) watch a few cat videos



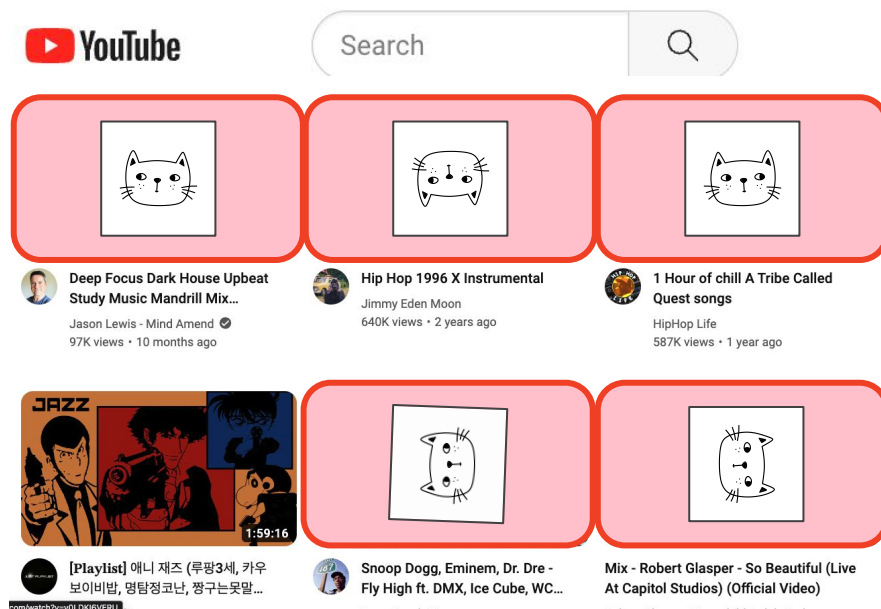
(1) watch a few cat videos



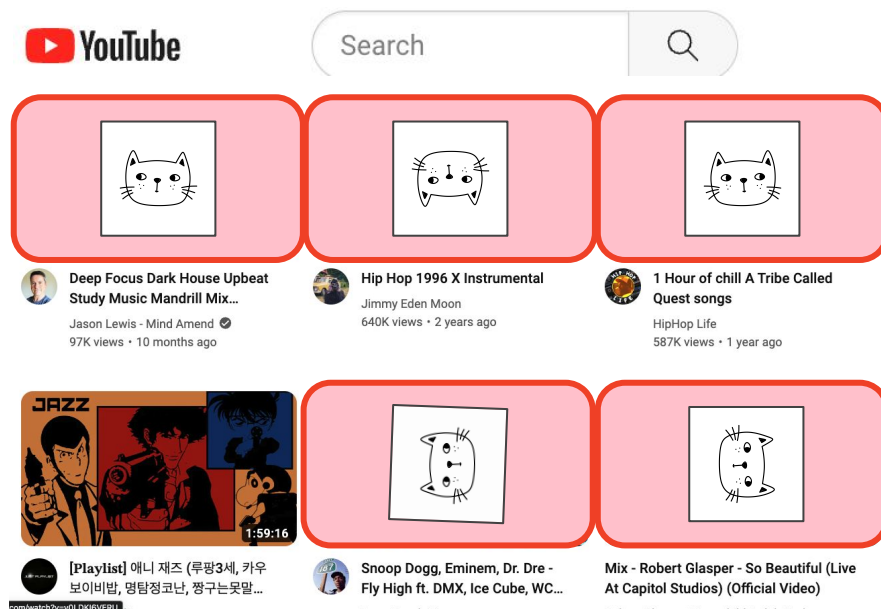
(1) watch a few cat videos



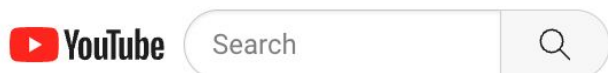
(1) watch a few cat videos (2) too many cat recommendations!




(1) watch a few cat videos (2) too many cat recommendations! (3) how to remove?




(1) watch a few cat videos (2) too many cat recommendations! (3) how to remove?

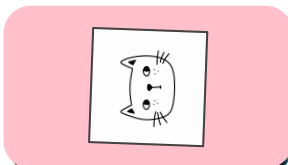



 **Deep Focus Dark House Upbeat Study Music Mandrill Mix...**
Jason Lewis - Mind Amend ✓
97K views · 10 months ago

 **Hip Hop 1996 X Instrumental**
Jimmy Eden Moon
640K views · 2 years ago

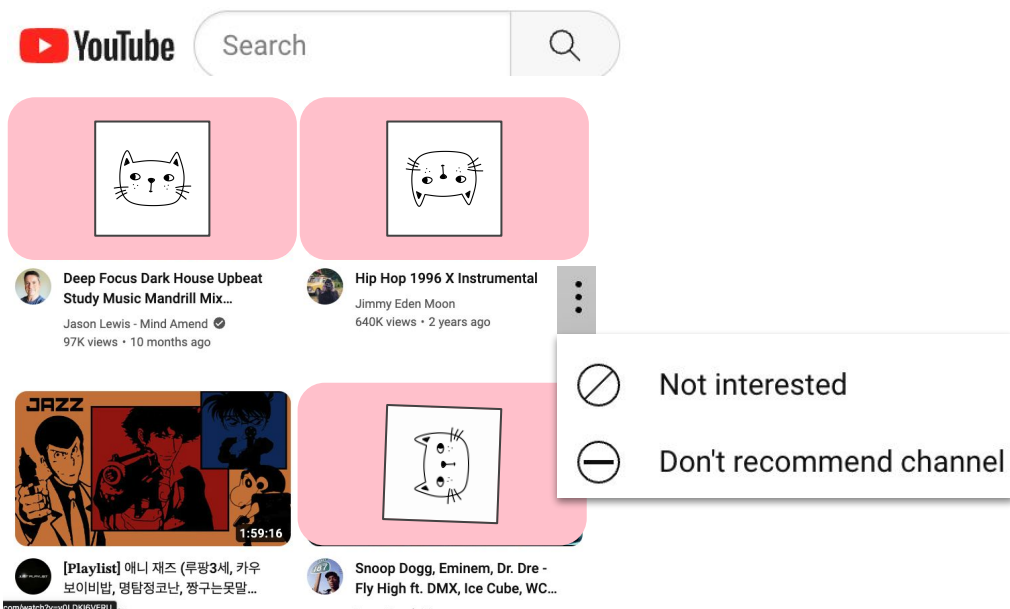


 **[Playlist] 애니 재즈 (루팡3세, 카우 보이비밥, 명탐정코난, 정구는웃말...**
bom(yeah3pveo)ilKXKv2R1...



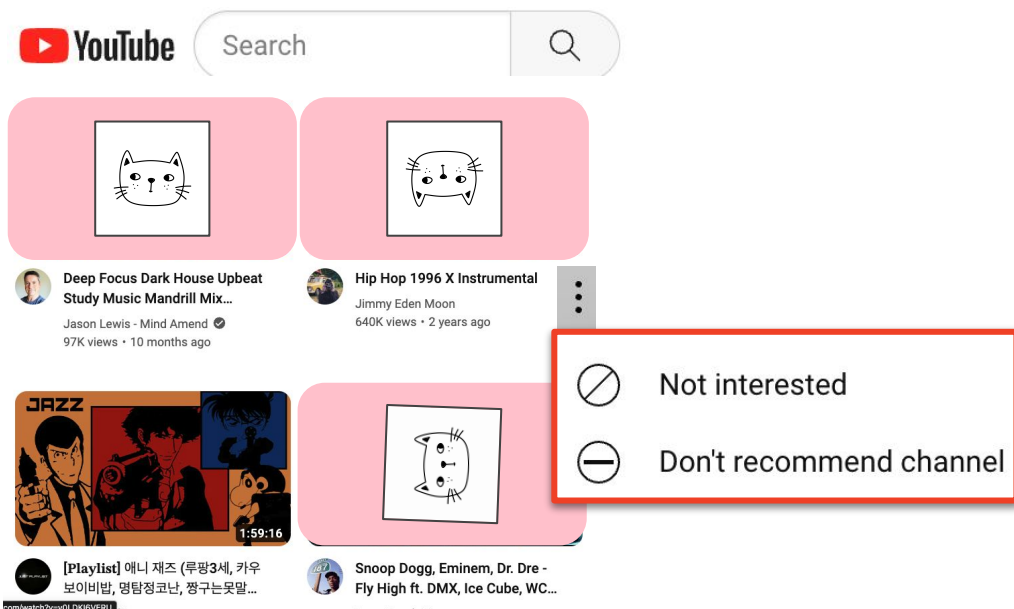
 **Snoop Dogg, Eminem, Dr. Dre - Fly High ft. DMX, Ice Cube, WC...**

(1) watch a few cat videos (2) too many cat recommendations! (3) how to remove?



The screenshot shows the YouTube homepage interface. At the top is the YouTube logo and a search bar. Below are video recommendations. Two recommendations feature a cat face icon: 'Deep Focus Dark House Upbeat Study Music Mandrill Mix...' by Jason Lewis - Mind Amend (97K views, 10 months ago) and 'Hip Hop 1996 X Instrumental' by Jimmy Eden Moon (640K views, 2 years ago). A third recommendation is a 'JAZZ' playlist. A fourth recommendation is 'Snoop Dogg, Eminem, Dr. Dre - Fly High ft. DMX, Ice Cube, WC...'. A three-dot menu is visible next to the 'Hip Hop 1996 X Instrumental' video, with a dropdown menu showing two options: 'Not interested' (with a crossed-out circle icon) and 'Don't recommend channel' (with a minus sign icon).

(1) watch a few cat videos (2) too many cat recommendations! (3) how to remove?



The screenshot shows the YouTube homepage interface. At the top is the YouTube logo and a search bar. Below are several video recommendations. Two of the thumbnails are cat faces. A red-bordered box highlights a feedback menu with two options: 'Not interested' (represented by a circle with a diagonal line) and 'Don't recommend channel' (represented by a circle with a horizontal line).

YouTube Search

Deep Focus Dark House Upbeat Study Music Mandrill Mix...
Jason Lewis - Mind Amend ✓
97K views • 10 months ago

Hip Hop 1996 X Instrumental
Jimmy Eden Moon
640K views • 2 years ago

JAZZ
1:59:16

[Playlist] 애니 재즈 (루팡3세, 카우 보이비밥, 명탐정코난, 정구는웃말...
bom(yeah3pew011KXV2R1...

Snoop Dogg, Eminem, Dr. Dre - Fly High ft. DMX, Ice Cube, WC...

Not interested

Don't recommend channel

(1) watch a few cat videos (2) too many cat recommendations! (3) how to remove?

Watch history



Not interested



Don't recommend channel



(1) watch a few cat videos (2) too many cat recommendations! (3) how to remove?

Watch history



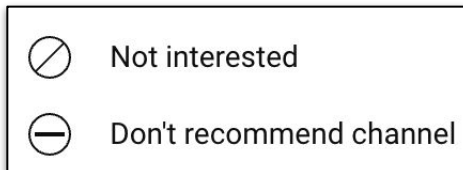
Not interested



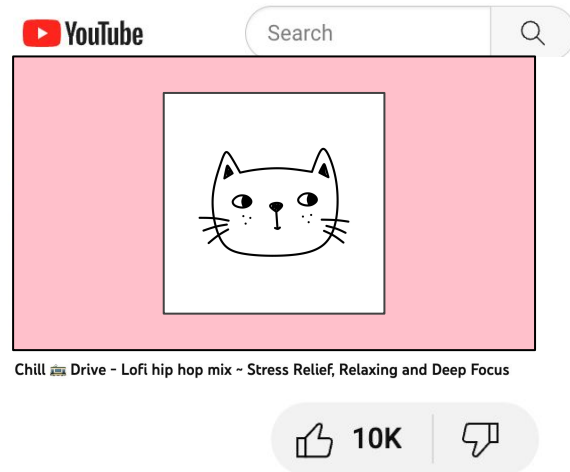
Don't recommend channel



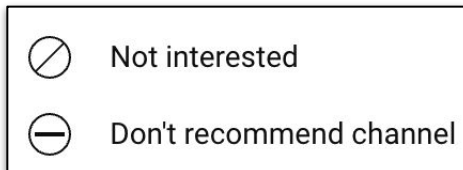
(1) watch a few cat videos (2) too many cat recommendations! (3) **how to remove?**



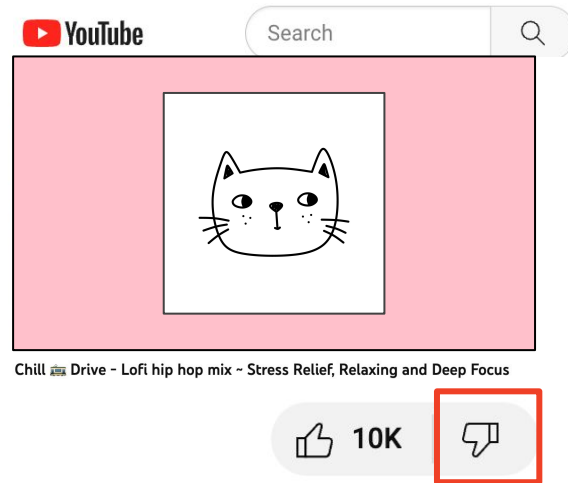
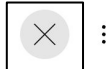
Watch history



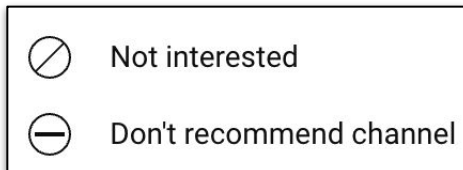
(1) watch a few cat videos (2) too many cat recommendations! (3) how to remove?



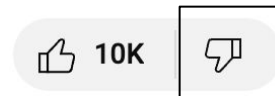
Watch history



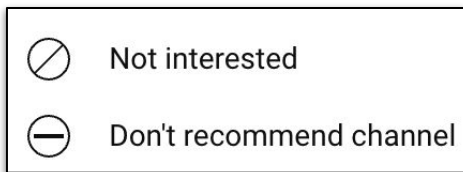
(1) watch a few cat videos (2) too many cat recommendations! **(3) how to remove?**



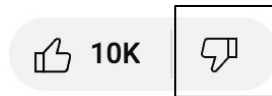
Watch history



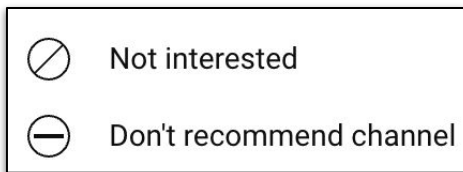
- YouTube makes unwanted recommendations
 - Societally-harmful content (Hosseinmardi et al. 2021, 2024; Lewis 2018; Ribeiro et al. 2020).
 - YouTube Regrets (Mozilla Foundation, 2021).



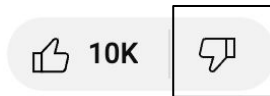
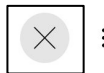
Watch history



- YouTube makes unwanted recommendations
- Many platform features could help remove them
 - → **But, do they work?**



Watch history

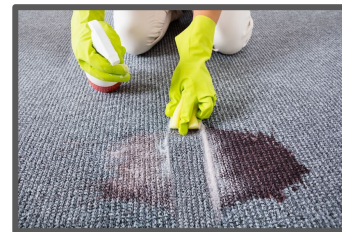


1. How big are YouTube **stains***?
 - a. What are the differences in topics?



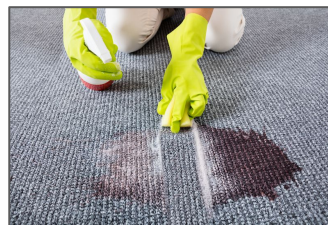
*stain = % recommendations of a topic, after watching many of their videos

1. How big are YouTube **stains***?
2. Can platform features **remove stain***?
 - a. Which ones work best?



*stain = % recommendations of a topic, after watching many of their videos

1. How big are YouTube **stains***?
2. Can platform features **remove stain***?
3. How do users interact with these features?
 - a. How many are **aware** of them?
 - b. How many **use** them?
 - c. How **effective** do they find them?



*stain = % recommendations of a topic, after watching many of their videos

1. Introduction
2. **Sock puppet audit**
 - a. **Methods**
 - b. Results
3. User survey
4. Conclusion

(1) Stain phase



(1) Stain phase



(2) Scrub phase



(1) Stain phase



(2) Scrub phase



One of:

Alt-Right

Antitheism

Political Left

Political Right

One of:

☐ Don't recommend channel

☐ Not interested

:

(1) Stain phase

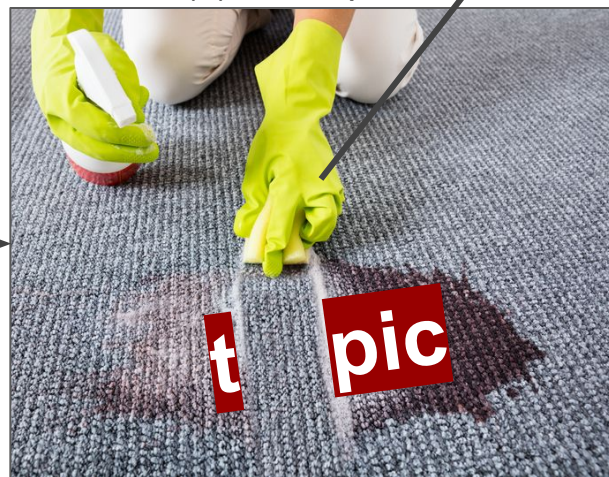


One of:

Alt-Right**Antitheism**

Political Left**Political Right**

(2) Scrub phase



Data collection

(1) Stain phase

P1

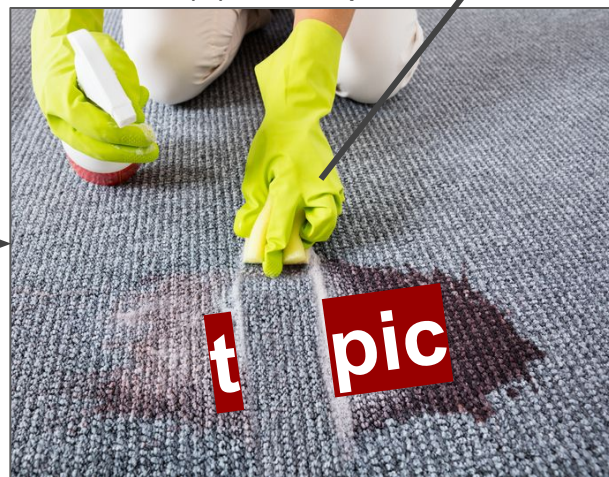


One of:

- Alt-Right
- Antitheism
- Political Left
- Political Right

P2

(2) Scrub phase



P3

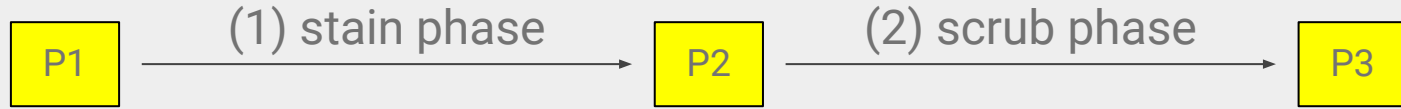
One of:

One of:

- Don't recommend channel
- Not interested
- "Watch neutral"

Topic: Cats


Scrub:  Not interested



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

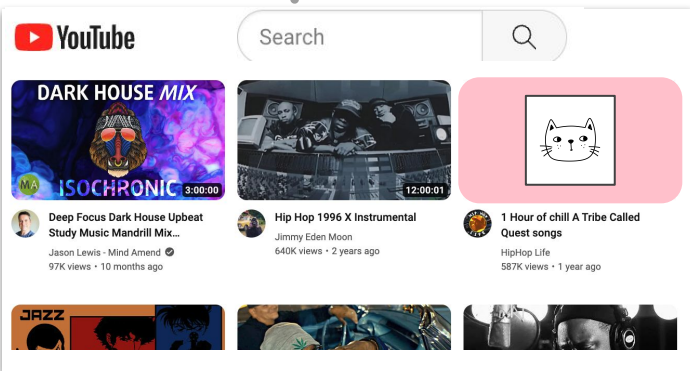
P1

(1) stain phase

P2

(2) scrub phase


P3



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

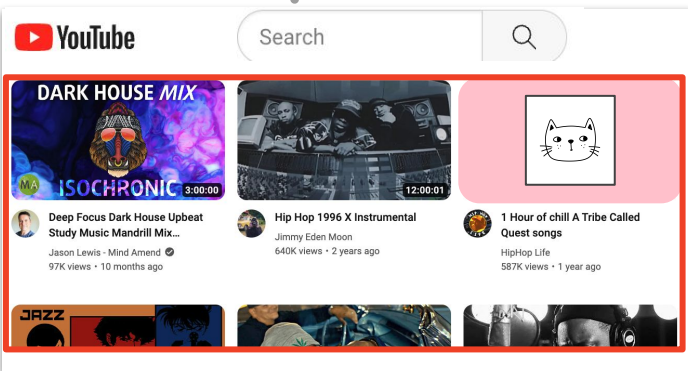
P1

(1) stain phase

P2

(2) scrub phase

P3

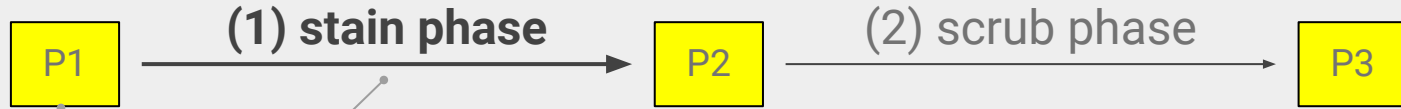


Sock puppet- Example

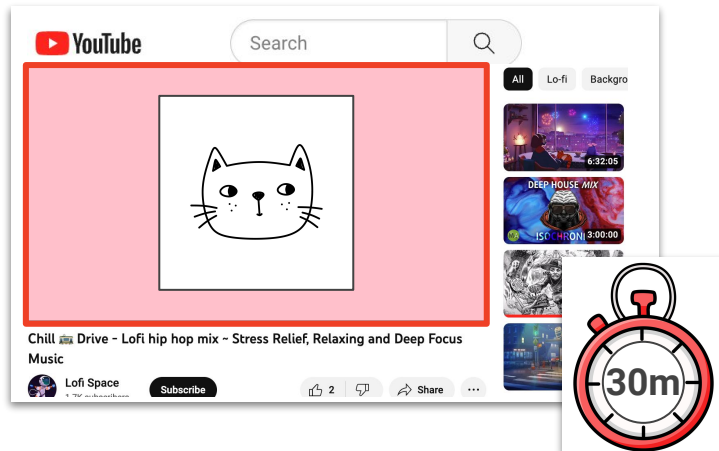


Topic: Cats

Scrub:  Not interested




collect homepage
recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

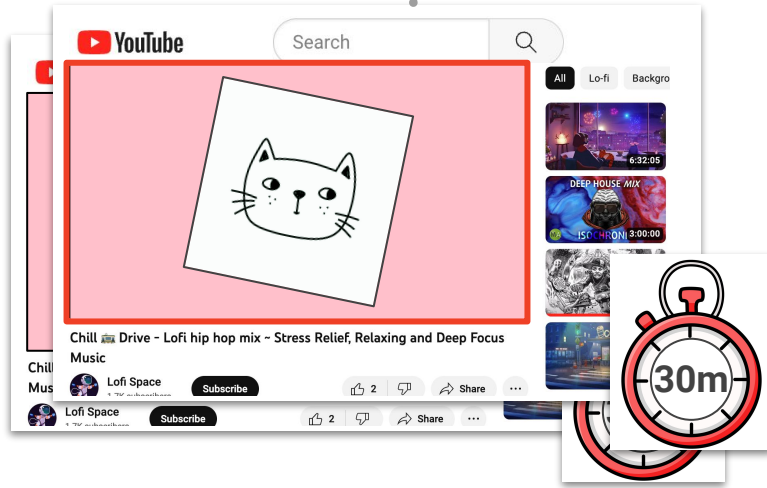
(1) stain phase

P2

(2) scrub phase

P3

collect homepage
recommendations

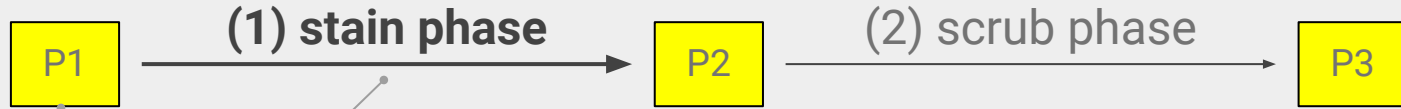


Sock puppet- Example

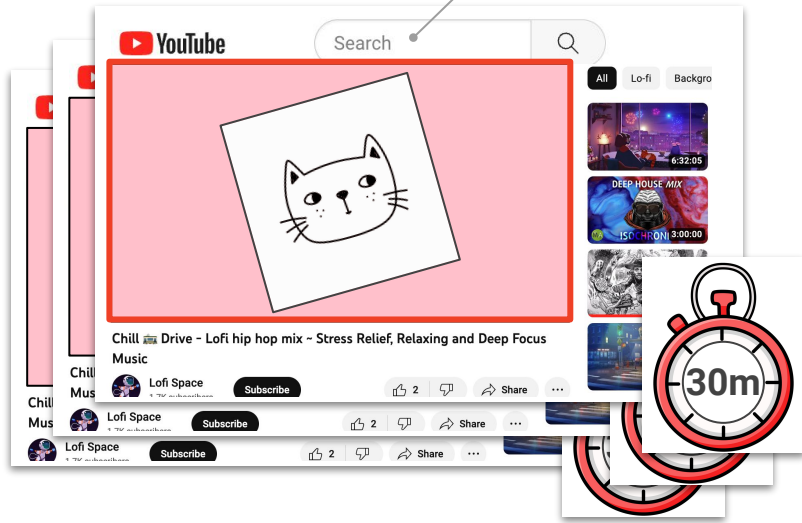


Topic: Cats

Scrub:  Not interested




collect homepage
recommendations

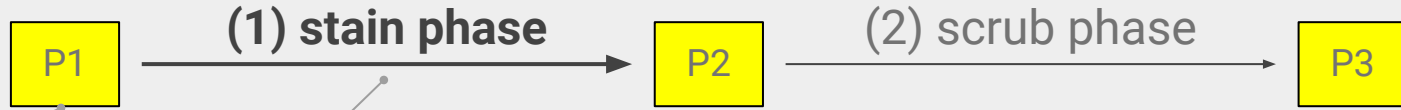


Sock puppet- Example

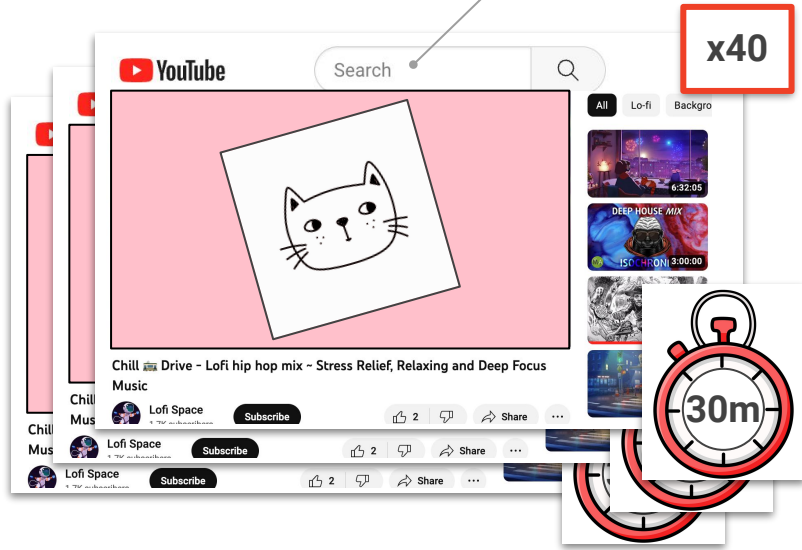


Topic: Cats

Scrub:  Not interested




collect homepage
recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

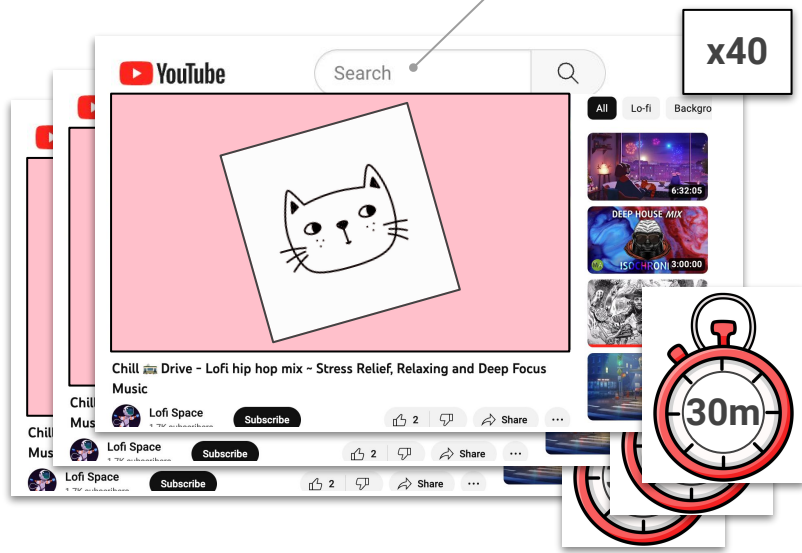
P2

(2) scrub phase

P3

collect homepage
recommendations


collect homepage
recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

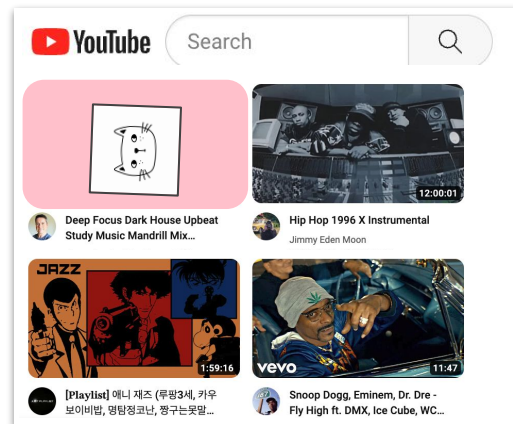
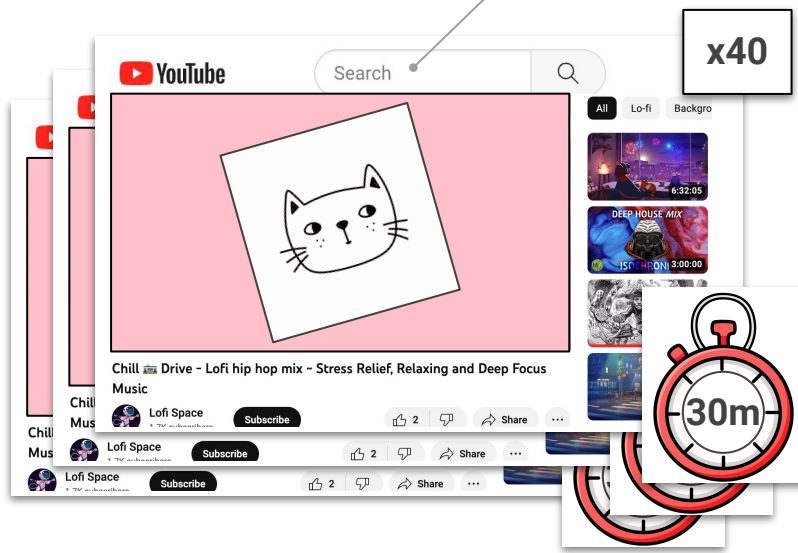
P2

(2) scrub phase

P3

collect homepage
recommendations


collect homepage
recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

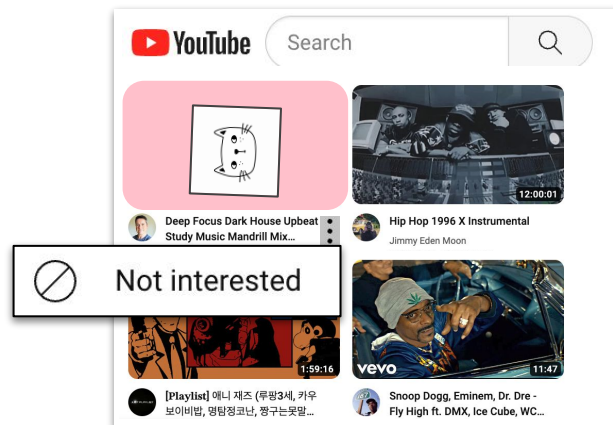
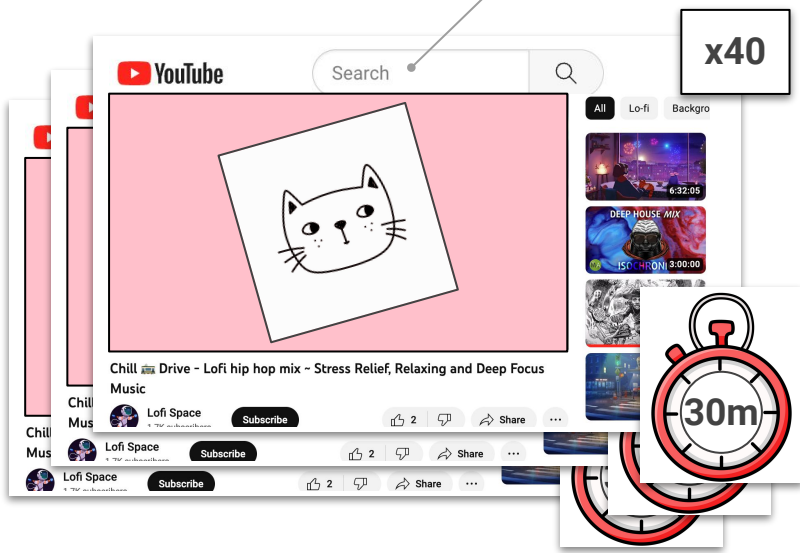
P2

(2) scrub phase

P3

collect homepage
recommendations

collect homepage
recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

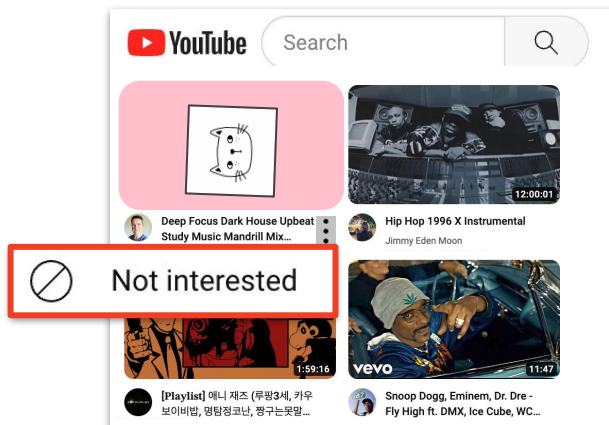
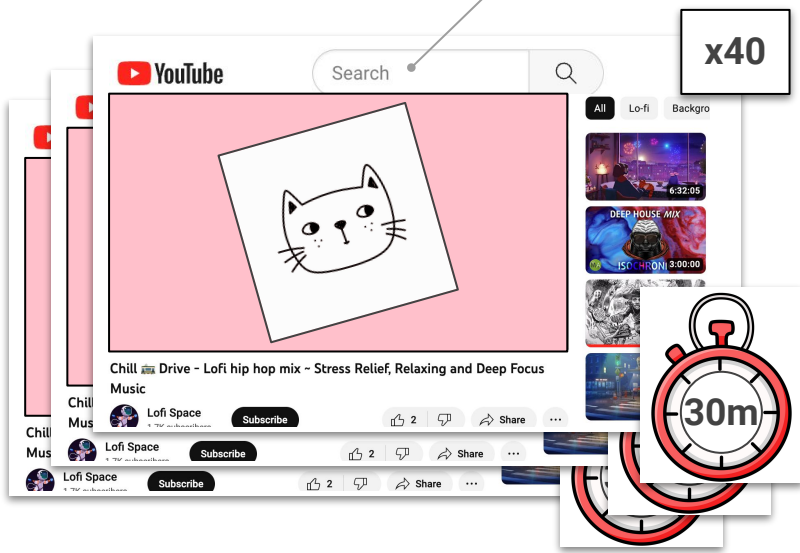
P2

(2) scrub phase

P3

collect homepage
recommendations


collect homepage
recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

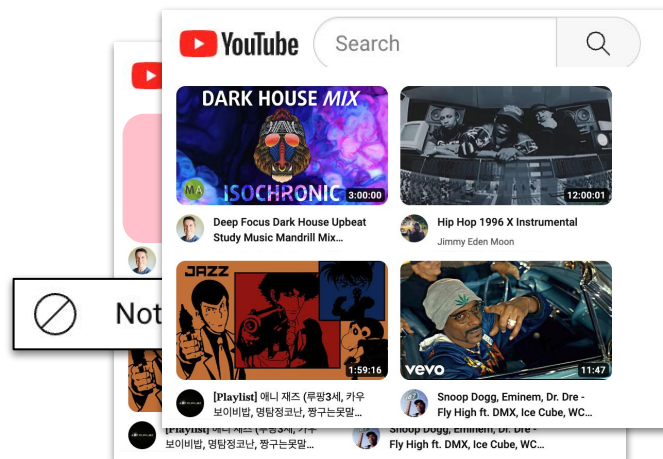
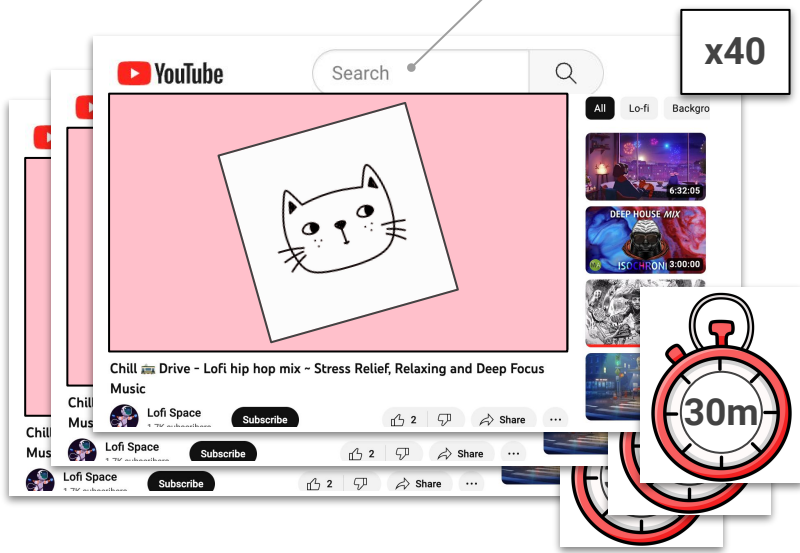
P2

(2) scrub phase

P3

collect homepage
recommendations


collect homepage
recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

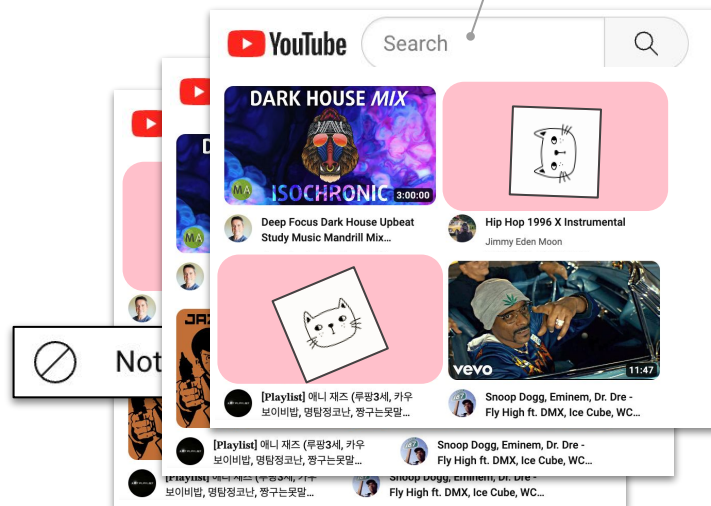
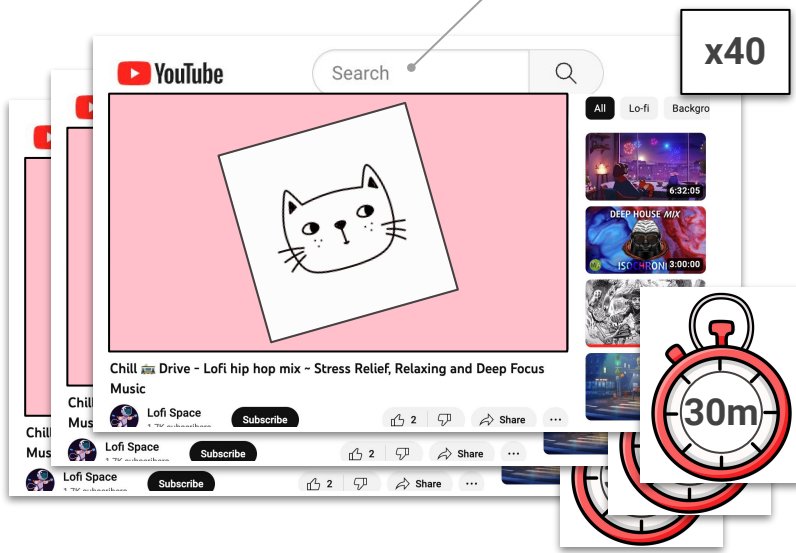
P2

(2) scrub phase

P3

collect homepage recommendations


collect homepage recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

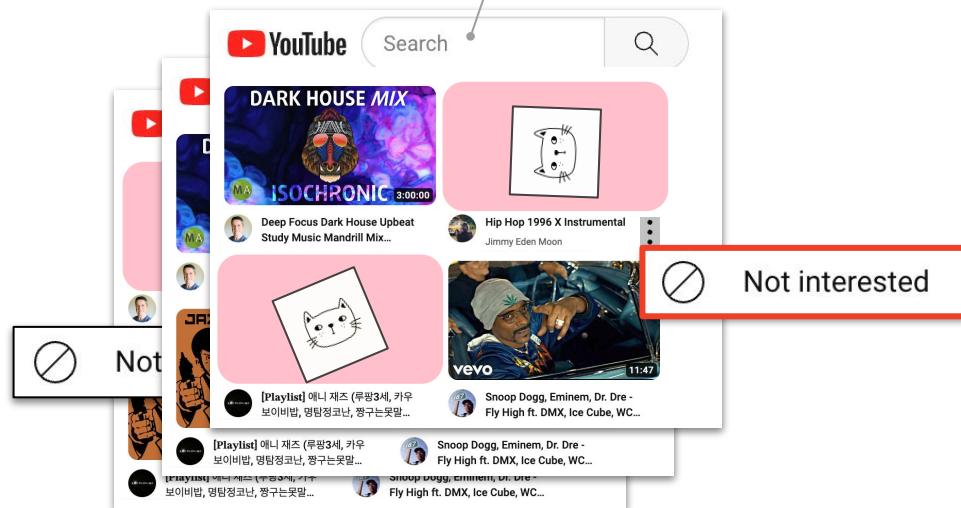
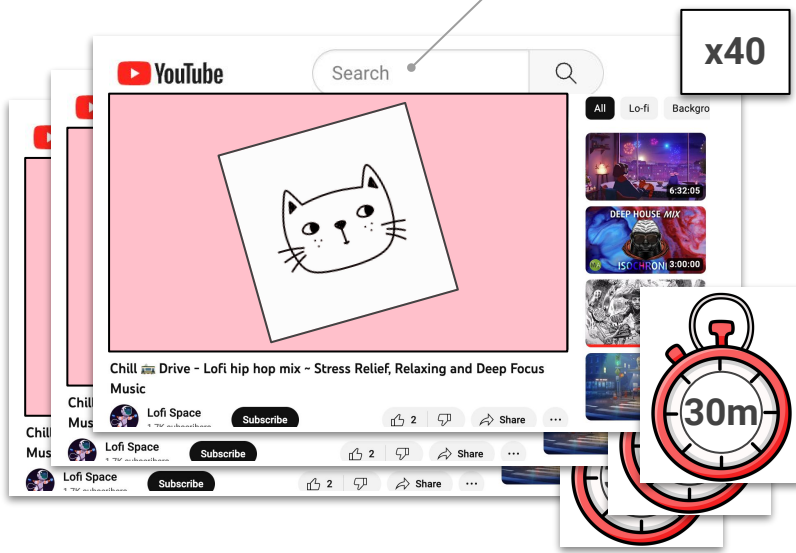
P2

(2) scrub phase

P3

collect homepage recommendations


collect homepage recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

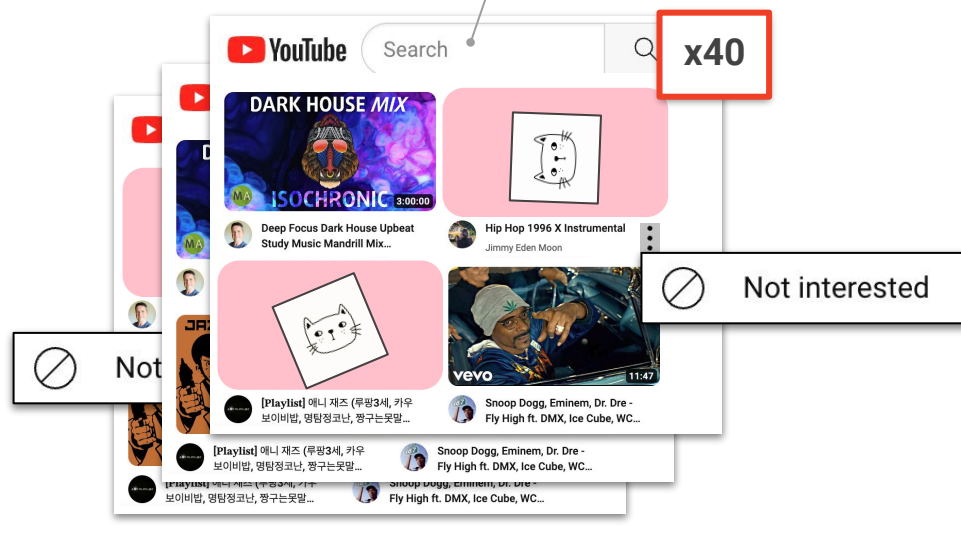
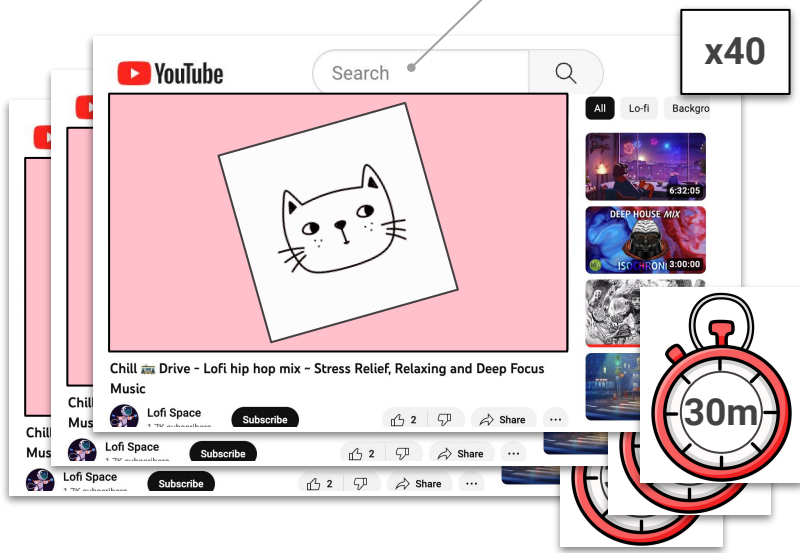
P2

(2) scrub phase

P3

collect homepage recommendations


collect homepage recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

P2

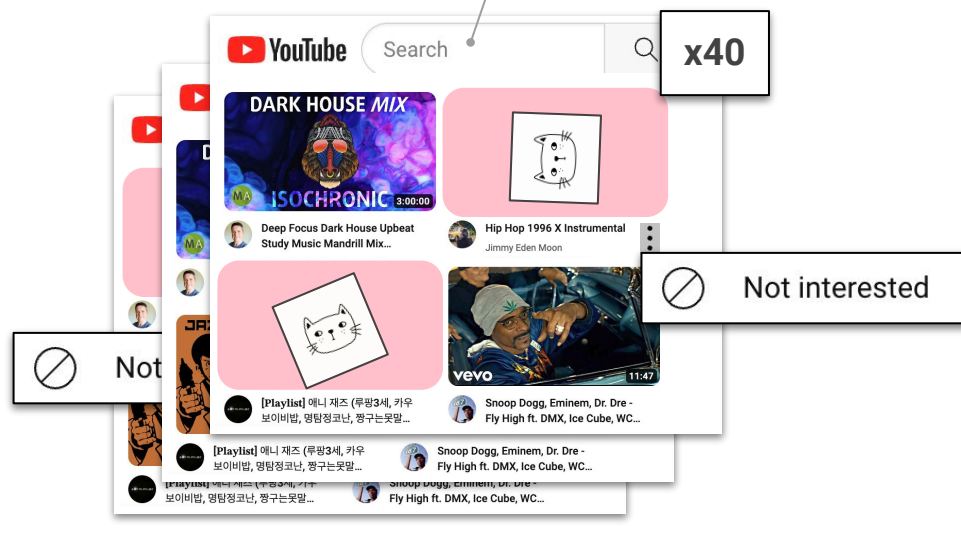
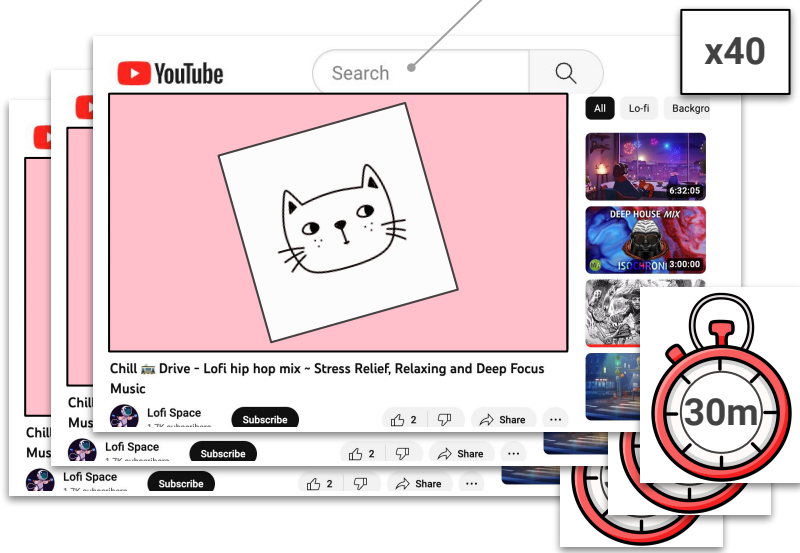
(2) scrub phase

P3

collect homepage
recommendations

collect homepage
recommendations

collect homepage
recommendations



Sock puppet- Example



Topic: Cats

Scrub:  Not interested

P1

(1) stain phase

P2

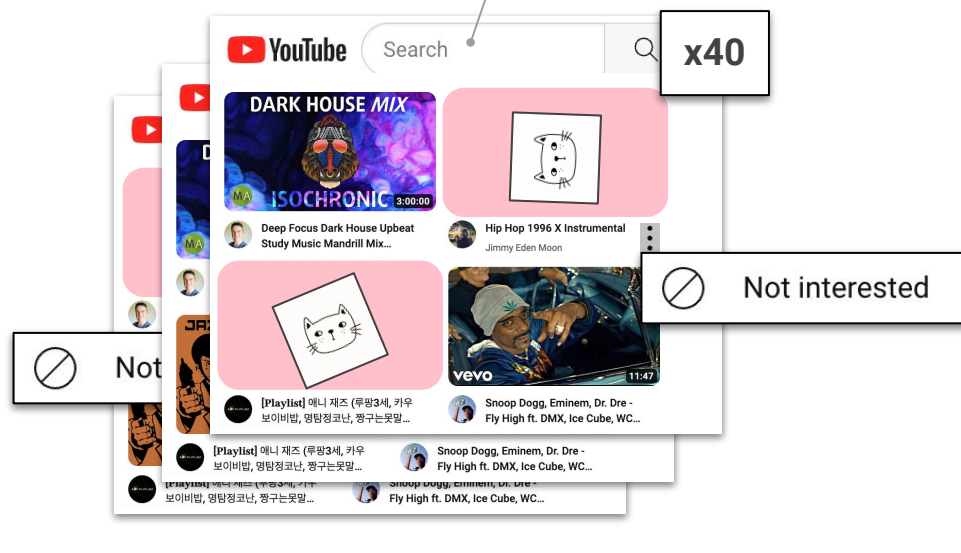
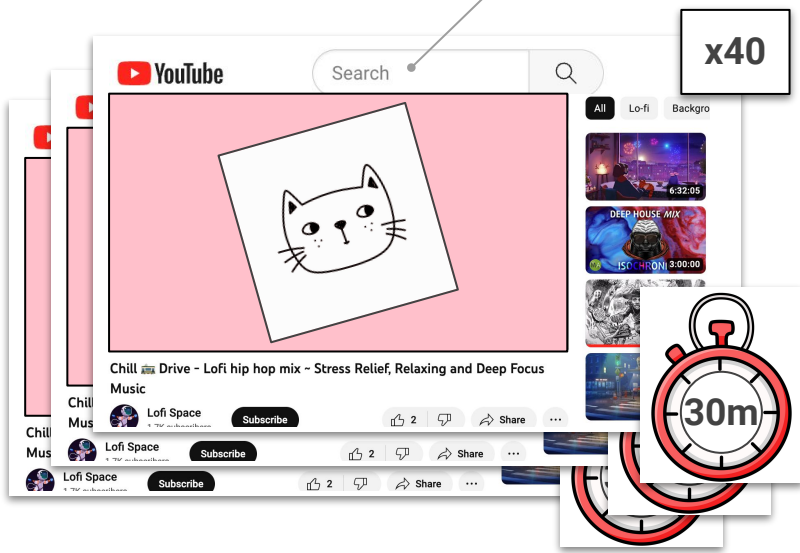
(2) scrub phase

P3

collect homepage
recommendations

collect homepage
recommendations

collect homepage
recommendations



Schedule of sock puppets

- For each topic, test 7 strategies, 5 times → 35 bots in parallel

Manual labeling of recommendations

- On-topic / Off-topic
- 0.66-0.95 Cohen's Kappa for different topics

1. Introduction
2. **Sock puppet audit**
 - a. Methods
 - b. Results**
3. User survey
4. Conclusion

- Stain happens on the homepage
 - but never reaches over half



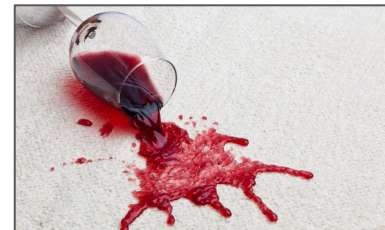
- Stain happens on the homepage
 - but never reaches over half



<i>Alt-Right</i>	
P1	P2
<hr/>	
2%	$\xrightarrow{*}$ 20%

$\boxed{\xrightarrow{*}}$ = significant difference (MWU test)

- Stain happens on the homepage
 - but never reaches over half




<i>Alt-Right</i>		<i>Antitheist</i>		<i>Political Left</i>		<i>Political Right</i>	
P1	P2	P1	P2	P1	P2	P1	P2
2%	* \rightarrow 20%	11%	* \rightarrow 37%	6%	* \rightarrow 28%	9%	* \rightarrow 29%



= significant difference (MWU test)

-  Not interested works best




 = significant difference (MWU test)

-  Not interested works best






	<i>Alt-right</i>	
	P2	P3
<i>Watch neutral</i>	12%	8%

 = significant difference (MWU test)

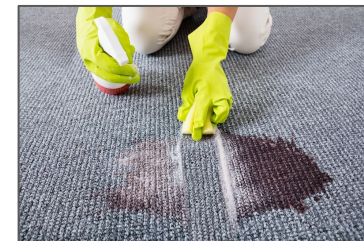
-  Not interested works best




	<i>Alt-right</i>		<i>Antitheist</i>		<i>Political Left</i>		<i>Political Right</i>		Avg. relative change P2 to P3
	P2	P3	P2	P3	P2	P3	P2	P3	
<i>Watch neutral</i>	12%	8%	34%	22%	32% 	18%	26% 	16%	-38%

 = significant difference (MWU test)

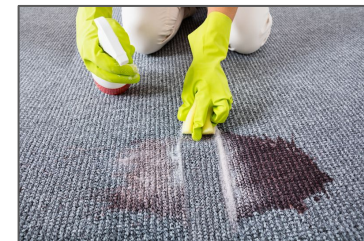
-  Not interested works best




	<i>Alt-right</i>		<i>Antitheist</i>		<i>Political Left</i>		<i>Political Right</i>		Avg. relative change P2 to P3
	P2	P3	P2	P3	P2	P3	P2	P3	
<i>Watch neutral</i>	12%	8%	34%	22%	32%	* → 18%	26%	* → 16%	-38%
<i>Delete</i>	20%	* → 0%	32%	24%	30%	* → 0%	24%	* → 4%	-77%
<i>Dislike</i>	20%	4%	30%	* → 10%	24%	28%	32%	* → 16%	-45%
<i>Not interested</i>	18%	* → 0%	42%	* → 2%	28%	* → 10%	26%	* → 2%	-88%
<i>No channel</i>	28%	* → 10%	44%	* → 22%	30%	* → 16%	36%	* → 24%	-49%
<i>Dislike rec.</i>	20%	* → 8%	42%	* → 14%	32%	24%	26%	34%	-30%

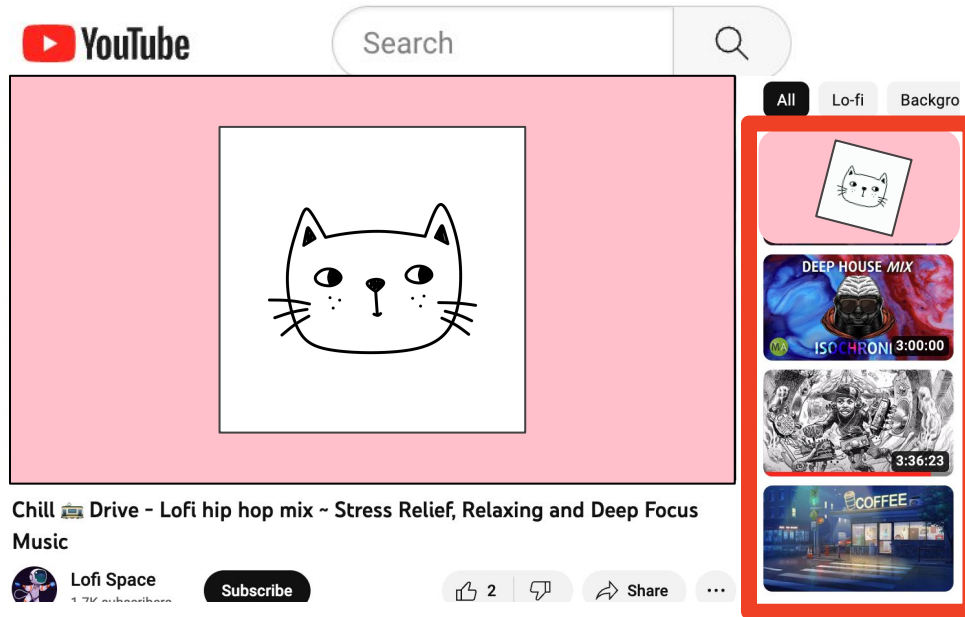
 = significant difference (MWU test)

-  Not interested works best

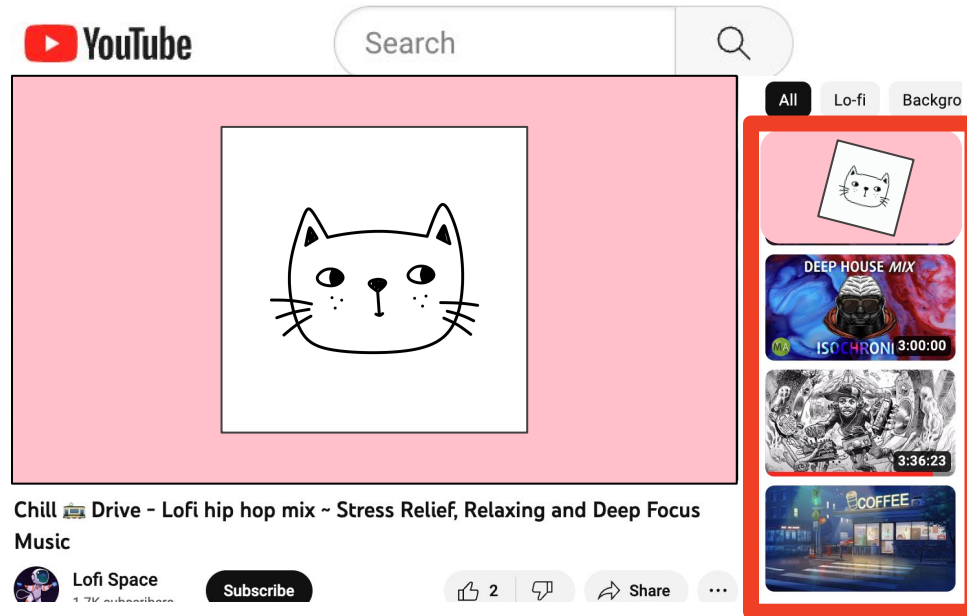


	<i>Alt-right</i>		<i>Antitheist</i>		<i>Political Left</i>		<i>Political Right</i>		Avg. relative change P2 to P3
	P2	P3	P2	P3	P2	P3	P2	P3	
<i>Watch neutral</i>	12%	8%	34%	22%	32% → 18%		26% → 16%		-38%
<i>Delete</i>	20% → 0%		32%	24%	30% → 0%		24% → 4%		-77%
<i>Dislike</i>	20%	4%	30% → 10%		24%	28%	32% → 16%		-45%
<i>Not interested</i>	18% → 0%		42% → 2%		28% → 10%		26% → 2%		-88%
<i>No channel</i>	28% → 10%		44% → 22%		30% → 16%		36% → 24%		-49%
<i>Dislike rec.</i>	20% → 8%		42% → 14%		32%	24%	26%	34%	-30%

 = significant difference (MWU test)



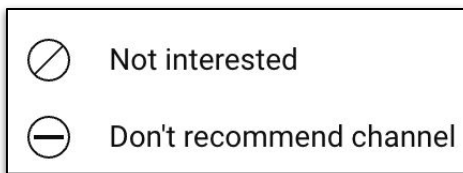
- Both the stain phase and scrub phase were largely *ineffective*



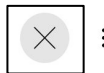
1. Introduction
2. Sock puppet audit
- 3. User survey**
4. Conclusion

For every button:

1. Were you **aware** this button existed?
2. Have you **used** it to remove unwanted recommendations (given aware?)
3. How **effective** was it (given used)?

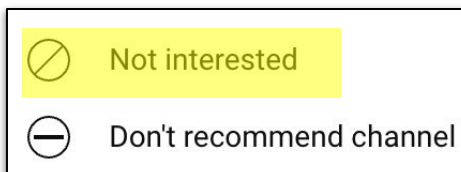


Watch history

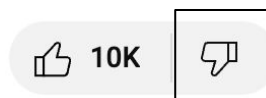


 Not interested :

1. Were you **aware** this button existed? → **56%** \pm 7%
2. Have you **used** it (given aware)? → **83%** \pm 8%
3. How **effective** was it (given used)? → **3.4/5** \pm 0.3



Watch history



Results

- Partial stain occurs
- Many YouTube users (44%) are unaware of the most effective recommendation-removal strategy (☐ Not interested)

Results

- Partial stain occurs
- Many YouTube users (44%) are unaware of the most effective recommendation-removal strategy (☐ Not interested)

Implications

- Users: Use the ☐ Not interested button
- Platform: Make effective features more prominent

Results

- Partial stain occurs
- Many YouTube users (44%) are unaware of the most effective recommendation-removal strategy (☐ Not interested)

Implications

- Users: Use the ☐ Not interested button
- Platform: Make effective features more prominent

... Thank you!

Works cited and sources

- Hosseinmardi, H.; Ghasemian, A.; Clauset, A.; Mobius, M.; Rothschild, D. M.; and Watts, D. J. 2021. Examining the consumption of radical content on YouTube. PNAS.
- Hosseinmardi, H.; Ghasemian, A.; Rivera-Lanas, M.; Horta Ribeiro, M.; West, R.; and Watts, D. J. 2024. Causally estimating the effect of YouTube's recommender system using counterfactual bots. PNAS.
- Lewis, R. 2018. Alternative influence: Broadcasting the reactionary right on YouTube. Data & Society.
- McCrosky, Jesse, and Brandi Geurkink. "YouTube Regrets." Mozilla Foundation. 2021.
- Ribeiro, M. H.; Ottoni, R.; West, R.; Almeida, V. A. F.; and Meira, W. 2020. Auditing radicalization pathways on YouTube. In FAccT.
- YouTube screenshots: Retrieved from YouTube.com. 2024.
- Cat cartoon