Algorithms’ Impact on Disinfo

Guillaume Chaslot
University Paris Est / Advisor at Center For Humane Technology

@gchaslot
YouTube’s AI impact

- 1 billions hours watched per day
- 70% of content comes from AI recommendations
- => humans spent > 100,000,000 years on content promoted by AI
AI’s Incentives

- Implicit: Watch Time
- Semi-Explicit: Engagement
- Explicit: YouTube surveys
Problems with Engagement

"Best" (sic) mention:

Meilleure mention a obtenu 210 engagements

Max Keiser, tweet poet.
@maxkeiser - 2 mai

@ghaslot That’s my network asshole. Be careful. You have been warned.
Policy Wave

Zuckerberg 2018:

Natural Engagement Pattern

Approaching the Line

ENGAGEMENT

ALLOWED CONTENT

POLICY LINE

PROHIBITED CONTENT
Borderline Content

Zuckerberg 2018:

Adjusted to Discourage Borderline Content

Approaching the Line

Allowed Content

Prohibited Content
Surfing Borderline Wave
Als create vicious circles

- Users will notice that some content is amplified more
- They will create more of it
- The amount of content make it appear “true” or “popular”
Al's create vicious circles

- Users will notice that some content is amplified more
- They will create more of it
- The amount of content make it appear “true” or “popular”
AlgoTransparency

- Scrap thousands of YouTube channels
- Follow Recommendations
- Count number of channels recommending each videos
AlgoTransparency.org

4834 recommended videos by YouTube on 04/03/2019

- **GRAHAM HANCOCK - AMERICA BEFORE: THE KEY TO EARTH'S LOST CIVILIZATION - Part 1/2 | London Real**
  - 282,437 views
  - 6,492 likes
  - 355 dislikes
  - Number of channels recommending it: 99

- **They Discovered Something at This Place That Could Change Everything!**
  - Matrix Wisdom
  - 37,832 views
  - 1,021 likes
  - 156 dislikes
  - Number of channels recommending it: 88

- **New Linda Moulton Howe Something HUGE Has Been Discovered in Antarctica**
  - UAMN TV
  - 9,646 views
  - 8,414 likes
  - 573 dislikes
  - Number of channels recommending it: 84
France, 2019 elections:
Freedom of Speech is Drowning

- Renee DiResta: Free Speech Is Not the Same As Free Reach
- Algorithmic Amplification is drowning Free Speech
We are drowning in disinformation / abuse

The best AIs of the world are making things worse

6 months: Facebook and YouTube pledge to demote “harmful” content, but algorithms still amplify disinfo
What is needed

- Transparency
- Platform “fixes”
- Design algorithm to be robust to abuse