Twitter strikes back
A failed attempt to measure info bubble

Kemal Erdem, Piotr Mazurek, Piotr Rarus
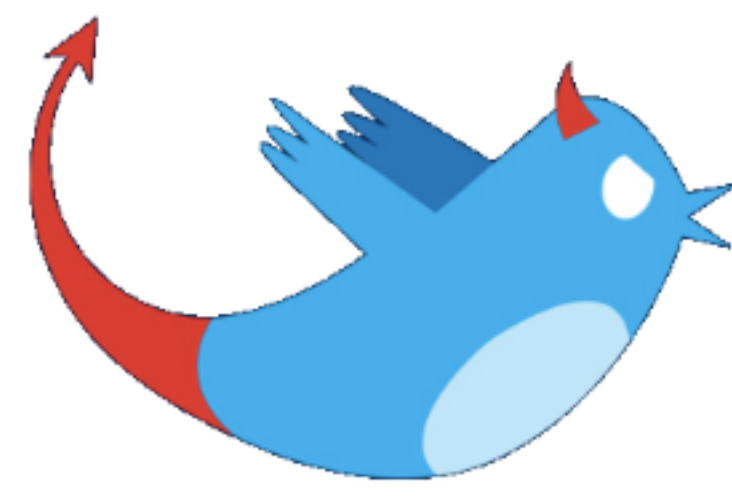
Department of Computer Science and Management
Wrocław University of Science and Technology
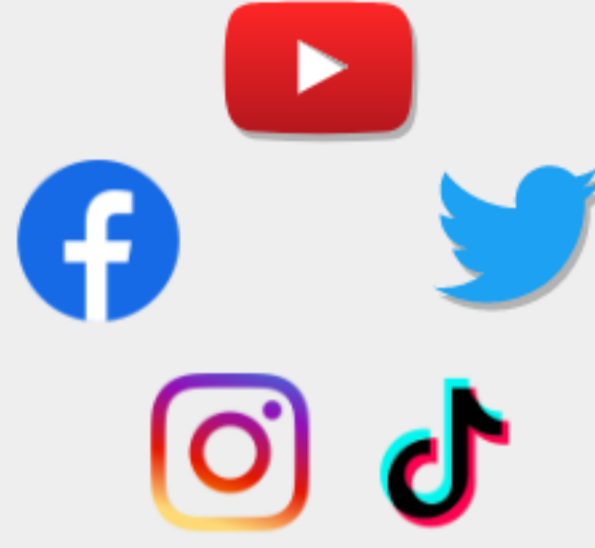Wybrzeże Stanisława Wyspiańskiego 27, 50-370 Wrocław
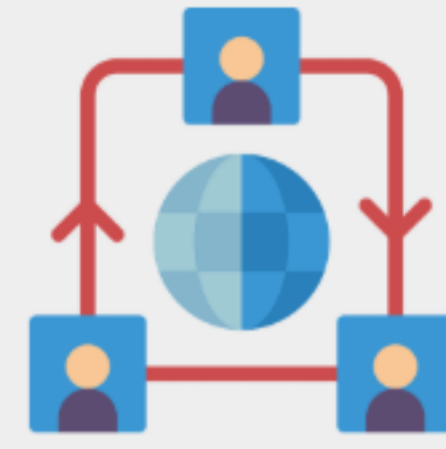
Introduction

ACCESS TO RELIABLE DATA

Social media platforms don't allow researchers to access their systems. How those systems work remains a mystery to the public.

RECOMMENDATION ENGINES

Subclass of filtering systems that are highly personalized and trying to keep the highest user engagement. Used widely by social media platforms.

INFORMATION BUBBLES

Is a state of intellectual isolation when the user stops exploring resources of the Internet and the algorithms push him to over-consume resources he already know

---

Joanna Bielecka
@JackBla10148512
Studentka Politologii UW
⊙ Warszawa   🗓 Joined November 2020
141 Following   7 Followers

Tweets   Tweets & replies   Media   Likes

Edit profile

Bots

SCRAPPING RECOMMENDATION ENGINE

We wanted to check if Twitter's recommendation engine creates information bubbles. To achieve that we've created 20 automated bots that were trying to act similar to real users. Each bot has its own bio, list of interests, and profile picture created with the help of GAN. The goal was to gather data for multiple social groups.

After collecting data we were planing to calculate entropy for each user and compare those entropies in different time periods. We've assumed that the lower entropy represents more isolated topics.

The Ban

HARMLESS RESEARCH ENDED UP WITH PERMANENT BAN

In the less than 24h all of our 20 bots were banned permanently. Bots were just observing the recommendations and didn't produce any content or conduct malicious activities.

Are you able to read it?

Twitter's T&C
🕐 23:30h
5633 words

U.S. Constitution
🕐 18:54h
4543 words

Summary

CURRENT ACCESS IS NOT GOOD ENOUGH

SOCIAL NETWORK = PUBLIC CONTROL

SCIENCE COMMUNITY SHOULD HAVE ACCESS TO SOCIAL NETWORKS' DATA

Public oversight over the actions of big tech companies is needed to ensure that no malicious action on their behalf in conducted

ACCESS BY CONTROLLED API IS NOT ENOUGH

The only way to objectively measure the information presented to users is to provide access to those data directly not by a proxy