

# Algorithmic curation and transparency of news algorithms

# Algorithmic curation

- The automatic selection, organizing and presenting of news
- Obvious examples are the social media sites like Twitter (most of the trending topics are news) and Facebook, but also news sites like Yahoo news and curation apps like Google News & Weather.

# Do people understand that their news are curated by an algorithm?

- Often they don't. (How often? It's a little unclear.)
- Their understanding (and misunderstanding) of how the algorithm behaves has an effect on their actions.
- There are independent curation tools for Facebook.

# Algorithm as a Black Box

- If internal working of an algorithm is unknown it appears as a black box. There is an input, something happens in the black box and an output comes out.
- Algorithm can be reverse engineered if both input and output are fully known. If either input or output is not fully known, reverse engineering will be much more difficult, and maybe impossible at least in any exact way.
- Exact reproduction of an algorithm is not necessary to understand it. If two algorithms produce the same output with the same input they can be considered equivalent, regardless of their internal structure.

# News algorithms are often secret

- News organizations are often private firms. Keeping the algorithm hidden offers a competitive advantage.
- Keeping the exact algorithm secret makes gaming and manipulating the system more difficult.
- There could be some embarrassing things hidden in the algorithms...

# Reverse engineering

- Analysis is easier if both input and output of the algorithm are known. This kind of situation might be for example in some forms of financial journalism.
- Analysis is easier in cases where analysis of natural language is not necessary. Good example of analyzing movie rating algorithms: <https://fivethirtyeight.com/features/fandango-movies-ratings/>

# Analyzing texts

- Statistical methods
- Text mining
- Opinion mining / sentiment analysis
- Etc...

# Why would you want to reverse engineer a news algorithm

- Competitor
- Investigative journalism
- Target of the news to manipulate the system



# Transparency

- How much transparency there should be in news algorithms?
- How much secrecy?
- Making the algorithm hard to reverse engineer and manipulation resistant