Behind the Algorithm

Jakob Svensson
Ecrea, Lugano, Nov 1st
Algorithms are on the Agenda
... especially when they fail ...

Facebook manually controlling their algorithms
When firing its trending team ➔ weird outcomes

Is Google racist?

(Tufekci, 2015)
(Thielman, 2016)
(Allen, 2016)
Algorithms are on the Agenda

Is Amazon homophobic? (Striphas, 2015)
Is Google Play homophobic? (Ananny, 2011)

Microsoft’s chat bot Tay (Neff & Nagy, 2016)
Algorithms are on the Agenda

Gender biases in image search algorithms (Kay et al, 2015)

Google image search: gender stereotypes

- Google image search for “C.E.O.” produced 11 percent women, even though 27 percent of United States chief executives are women.
Algorithms are on the Agenda

Is Flickr racist? (Haern, 2015)
Blacks are not recognized as humans in face-recognition (Sandvig et al, 2016)
Algorithms replace editors (DeVito, 2016) to journalists (van Dalen, 2012; Bozdag, 2013).

Steiner (2012) ⇒ that algorithms control financial markets, the music that reaches our ears, and even how we choose a partner.

Algorithms are responsible for selecting the information that reaches us (Gillespie, 2014), which has consequences for the shaping of our social and economic life (Kitchin, 2017).
Algorithmically generated *news feeds* influence the issues on our agenda and how these issues are framed (Just & Latzer, 2016), which in turn influences our decisions, preferences and even election results (Tufekci, 2015).

Cambridge Analytica using for political purposes personal data harvested about Facebook users.

cf. US – targeting users in Cuba on their profiled stance on the Revolution.
“Big data” → more than its size this is data that can be searched, aggregated and triangulated (by algorithms) with other sets of data (Shorey & Howard, 2016: 5033).

Datafication = transforming social action into online quantified data (van Dijck, 2014)
Dataism = widespread belief in the objective quantification and potential tracking of all kinds of human behavior and sociality through online media technologies” (ibid.)
Data-essentialism? → data being the essence of basically everything

Connected to increasing storage facilities (in terms of both processing and price)
Algorithms as socio-material processes

1) input, (the designing/programming based around problems that need to be solved), which

2) leads to the formulation of one (or several) calculations (sort, filter, rank, profile users, weigh → see Bozdag, 2013)

3) calculations that then result in some kind of outcome (output)
the Algorithmic Process / Situation

Input → Calculation → Output

big data context

behind the screen

in front of the screen
A “much-needed” sociological approach to research on algorithms. The focus is on the humans behind them.

Being engineered by humans, they embody social norms, values, imaginations, perceptions, rules, processes and are encoded with human intentions that may or may not be fulfilled.

Pivotal if we intend to have an informed discussion of power, and shifting relations of power, in contemporary digitalized societies.
The Humans Behind

Silicon Valley is “incredibly white and male” (Yarow, 2015)

Minority groups = 26 % of population, but only 5 % of tech
Women = 51 % and less than 20 % of tech (James West - Ted Talk 2015)

Face-recognition algorithms would probably recognize black people if black people were involved in their design/ training data. Image-search algorithms would probably be less gender biased if more women were involved
Research question: What logic, or combination of logics, informs the practices of designing and programming algorithms?

A study software engineers and their intentions, imaginations/ perceptions, rules, ideals, different cultures and how this feeds into their programming and designing of algorithms.

1) An interview study targeting software engineers, algorithm programmers and designers at in particular social media and search engine organizations

2) A study of a news organization. The study will take place at a leading Scandinavian daily and study different actors’ work with their webpage and the ranking/ placing of news.
Web–editors replaced by an algorithm 2015

The Problem? = Profitability

Mechanical work – pushing articles up and down
Work-saving “do more stuff without employing more people” → automation

“either journalists pull and drop stuff on the portal and spend a lot of time doing such work, or, we decide that we will try to automate this”

Controversial → “as a web editor, you did a job and spent a lot of time doing it and someone said that we should do that automatically. That was a challenge” Then there were commercial challenges. Do we get enough pageviews and advertising revenue in this way? Can we sell subscriptions? Will it feel like the XXX?”

→ Editor-led Algorithm
Actors behind the News-ranking Algorithm

Actors
Programmers, UX designers, Data-analytics, Media Group, the Brand, Profit (subscription vs. the ad department), Editors, Social media editor and Journalists

Clashes between tech. & journalists
→ even though tech always tried to get editor approval (editor-led algorithm)
→ putting the two sites next to each other to convince

“There was a feeling that they (journalists) underestimated how difficult it would be to build this site, some did not understand the technical bit” (programmer)

“There are no long perspectives, you know what's going to happen next week, but then it ends, then you have no idea. From an editorial perspective, it is perfect, but from a product development perspective and technology perspective, it's a bit more difficult because one wants a long-term perspective” (tech development)

→ Algorithm “fika”
Time (in terms of latest and longevity), News value (1-5), Subscription conversion, Popularity (clicks) - and these parameters can be tweaked

*News room Screens*
- Subscription conversion
- Most read (clicked) last minute
- Most read last hour
- Social Media
Logics behind the News-ranking Algorithm

**News logic** → based on news values (breaking)
-vs.- **Social media logic** → trending (often based on your friends - personalizing)

“do not share breaking news or ongoing news on social media” (social media ed.)

**Democratic / public service/sphere logic** → higher purpose
-vs.- **Commercial logic / media group logic**

the problem was that the news paper was not profitable – synergy effects, direct traffic between the media groups different journalistic products
-vs.- **Personalization logic** → filter bubbles, give readers an overview, be different from social media

Seems to be a conflict between ad and tech actors vs. editors, journalists and subscription “We in this media industry do not keep up (with social media), the norm for what people expect is not established on our platforms” (tech development)
Logics behind the News-ranking Algorithm

**Ad logic** → click, popularity selling readers views
vs. **subscription logic** → locked content that should be worth paying for
→ “the oracle” – when to lock free content

**Brand logic** → “could see who was the editor before – now it is a more coherent product”
vs. **Personalization logic** → be different from social media

**Publisher logic** → “the algorithms cannot become like magic, we need to have control, especially in a newspaper with a publisher” (editor) → Editor-led algorithm

“Tame the algorithm” “massage the algorithm” (journalists)
“We get continuous feedback in the form of data, we do data-driven development” (tech development)
“things and stuff (saker och ting) shall be put into ones and zeros” (data analytics)
Thank you for listening!

jakob.svensson@mau.se