Ethics and NLP
Some announcements

Sebastian Padó
26.4.2021
Seminar organization

- Looks like we have about 24 participants
  - = 12 sessions
- We need an extra session!
  - July 26, 2021 (Monday after term ends, usual slot)
- We have a lot of interest for privacy, fact checking
  - I will look for additional papers
- No love for topics
  - 4 (bias in NLP for social sciences)
  - 21 (multimodal information)
  - Isn’t anyone interested in this? Drop me an email ;-)
Prelim assignment

- Session of May 3rd
  - Jiwon Kim: Topic 1 – what is bias?
  - Sakshi Shukla: Topic 2 – debiasing
- Remember to incorporate **both** the context and the technical paper!
- Of course, you don’t have to submit a week in advance ;-
- Remaining assignments to follow in the next days
Collaborative note taking

• Does this URL work? (was in last week’s slides)
  • https://cryptpad.fr/whiteboard/#/2/whiteboard/edit/DxBrFc4yzoVCGMfBr6mo9bla/
Questions for Zittrain

1. What is the one topic or statement in the talk that struck you most?

2. What is agenda setting?

3. Zittrain defines AI as 'arcane pervasive tightly coupled adaptive autonomous systems'. Please unpack that definition.

4. Zittrain talks about many problems that involve feedback loops (e.g. voter badge presentation in FB, rider rating in Uber). Are these really AI problems?

5. Two important concepts appears to be fairness and responsibility. Please define them and outline whether you believe companies need to be fair and responsible.
Questions for Zittrain

6. Do you buy the dichotomy between "tools" and "friends" that Zittrain claims, and the technological movement towards "friends"? What consequences does that have?

7. The topic of human involvement comes up several times. Can human involvement solve the problems that were sketched? (Which? How?)

8. At 1:57, Zittrain talks about 'overfitting'. Do you think that this the right term? Why / why not?

9. Why do justifications matter so much?

10. What do you think are responsibilities of Data Scientists?