"WHAT I WON'T BUILD" IN NLP
BY RACHEL TATMAN @ ALP 2020 CONF.

WHAT'S MOST URGENT "DOGMA" IN NLP?

① Used to be: "Language technologies must work equally well for everyone."

② Now is: "Language technologists [NLP researchers & developers] must proactively mitigate harm from our work."

   ↳ eg minimizing differences in accuracy between subgroups

   ↳ eg "should this be built at all?" — We have limited resources & need to use them wisely. We don't need to immediately research everything just because it's there.

Just like particle physics

KEY QUESTION

What do we mean by "responsible"?

"Do you consider yourself responsible for the usages imagined from the applications / algorithms you create?" (Ethical grounds)

STANDARD ETHICS THAT WE NLP RESEARCHERS SHOULD ABIDE BY (ACL ACM CODE OF ETHICS)

politics screws all of this up? and it should, sometimes.

① contribute to society & to human well-being, acknowledging that all people are stakeholders in computing

② Avoid harm

③ Be honest & trustworthy

④ Be fair & take action not to discriminate

⑤ Respect the work req'd to produce new ideas, inventions, creative works, & computing artifacts

whose?

⑥ Respect privacy

→

⑦ Honor confidentiality.

WHAT SHE WON'T BUILD

→ deceptive tech that tries to pass itself as a human (without telling the human it's interacting with)

ties in nicely to linguistics ch. of textbk.

→ social category detectors →using sociolinguistic identity to predict your race, gender...

→ surveillance technology

QUESTIONS TO ASK

① Who benefits from this system existing?
— harmed?

② Can users choose NOT to interact with this system?

③ Does this system re-inforce systemic inequalities?

④ Does this genuinely better the world?