Studying Al in its Social Context

HUMAINT Winter school on Fairness, Accountability & Transparency in Artificial Intelligence

Professor Gina Neff @ginasue



Society fears the AI of the future, but we are not ready for the AI of today

Al as fact, future, fantasy



Al as fact, future





The challenge for researchers: Al is becoming the infrastructure for society; but how do we study infrastructure?

Theoretical Toolkit

Theories of Technology Construction

1. Science & Technologies Studies approaches to socio-technical networks eg Latour 2005; Law 2009; Mol 2010

2. Cultural Historical Activity Theory

eg. Kaptelinin & Nardi 2006; Engeström 2018; Engeström 2015

3. Affordance theories of social affordances of technologies; design-in-use, social construction of technologies e.g. Nagy & Neff 2015; Davis & Chouinard 2015; Davis 2020

4. **Socio-materiality** e.g., Orlikowski, W. J. (2007); Leonardi, P. M., Nardi, B. A., & Kallinikos, J. (2012); Nicolini, D., Gherardi, S., & Yanow, D. (2003).

5. Sociotechnical imaginaries e.g. Jasanoff & Kim 2015

Normatively: Fixing Al's social challenges

1. Amplification of small differences through technically inflexible—but seemingly responsive—systems.

2. Limits of technological solutions to problems of social bias.

3. Growing gulf between people who design and deploy versus those who use and are affected by AI systems.

4. Need for new data and inputs comes at costs to existing workplaces and routines.

5. Limits to data-driven solutions to social and political problems.

Methods Toolkit

SURVEY: Do selfie-related behaviors influence selfobjectification and appearance concerns among adolescents?

The longitudinal and reciprocal relationships between selfie-related behaviors and self-objectification and appearance concerns among adolescents new media & society 1–22 © The Author(s) 2019 Article reuse guidelines: sagepub.com/journals-permissions DOI: 10.1177/1461444819894346 journals.sagepub.com/home

Wang, Y., Xie, X., Fardouly, J., Vartanian, L. R., & Lei, L. (2019). The longitudinal and reciprocal relationships between selfie-related behaviors and self-objectification and appearance concerns among adolescents. New Media & Society. https://doi.org/10.1177/1461444819894346

Article

INTERVIEW, SITE OBSERVATIONS: How does Amazon's labour system work with material and cultural infrastructures? What is the relationship between algorithms, management and technical and political rationalities?

Article

Machinic dispossession and augmented despotism: Digital work in an Amazon warehouse new media & society 1–17 © The Author(s) 2019 Article reuse guidelines: sagepub.com/journals-permissions DOI: 10.1177/1461444819891613 journals.sagepub.com/home/nms SAGE

Delfanti, A. (2019). Machinic dispossession and augmented despotism: Digital work in an Amazon warehouse. *New Media & Society*. <u>https://doi.org/10.1177/1461444819891613</u>

EXPERIMENT: Do people find AI trustworthy in social settings? Shown AirBnB profiles said to be written by humans and AI which do people find trustworthy under what conditions?

Maurice Jakesch, Megan French, Xiao Ma, Jeffrey T. Hancock, and Mor Naaman. 2019. AI-Mediated Communication: How the Perception that Profile Text was Written by AI Affects Trustworthiness. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (*CHI '19*). Association for Computing Machinery, New York, NY, USA, Paper 239, 1–13. DOI:https://doi.org/10.1145/3290605.3300469 The Challenge: Empirical evidence of emerging technology can be tough; what is often needed is scoping and theorising

Solutions?

Proximal studies Edge cases Emergent user perspective Connection to expert/insider position, "studying up" The Challenge: "Studying up" in technology presents political, ethical and practical challenges

Source: A. Tanweer, Tanweer, A. (2018). Data science of the social: How the practice is responding to ethical crisis and spreading across sectors (Doctoral dissertation). https://digital.lib.washington.edu/researchworks/handle/1773/43343

Source: A. Tanweer, Tanweer, A. (2018). Data science of the social: How the practice is responding to ethical crisis and spreading across sectors (Doctoral dissertation). https://digital.lib.washington.edu/researchworks/handle/1773/43343

Studying AI in Social Context

Research needed

- **Case studies:** Map, track, compare and measure changes in people's practices with AI adoption across multiple settings;
- **Identify** key elements of social infrastructure that serve as levers for responsible use, intervention and accountability in Al systems;
- **Measure** how people actually respond in AI systems applying existing social norms, conventions and heuristics;
- Identify areas of how domestic life shapes and is shaped by introduction of AI;
- **Compare** the impact of differences in the unique local, cultural, and regulatory contexts on practices in AI systems;
- Map changing roles for women's authority, agency and voice in systems built for AI and in emerging forms of organizations.

Al Studies in Practice

Participatory Design Workshop with 40 experts generated 5 types of problems in AI integration

- Biased data → biased outputs
- 2. 'Growing pains'
- 3. Mismatched Expectations
- 4. 'Wrong' Perceptions
- 5. 'Bad' Applications

Gina Neff & Jevin West, Facebook SocSciFoo, 2019; https://callingbullshit.org/

Today's AI rollout is the world's largest ever social experiment

Safety Suis

POTENTIAL HARMS FROM ALGORITHMIC DECISION-MAKING

Chart Contents Courtesy of Megan Smith, Former CTO of the United States

Facial Recognition Software Gender Shades found it 'works' better for men with light skin

Error rates for women with dark skin ranged from 20%-34%

– Joy Buolamwini & Timnit Gebru 2018 Gender Shades

Demo: Assessing Bias in Machine Translation

Ő egy gyerekvigyázó. Ő egy pék.

Ő egy ápoló. Ő egy munkatárs.

Ő egy személyi asszisztens. Ő egy pénzügyi asszisztens. Ő egy vállalati kapcsolattartó asszisztens. Ő egy toborzási munkatárs.

Ő egy tanár. Ő egy általános iskolai tanár. Ő egy középiskolai tanár.

Ő egy tudós. Ő egy matematikus. Ő egy informatikus. Ő egy RFID fejlesztő mérnök.

Ő egy jogász. Ő egy bankár. Ő egy pénzember.

Ő egy zenész. Ő egy zeneszerző.

Ő egy konyhafőnök. Ő egy séf. Ő egy szakács.

Normative tools: What should be done

1) What and whose goals are being achieved or promised through

- 2) What structured performance using
- 3) What division of labour
- 4) Under whose control and
- 5) At whose expense?

My own Guiding principles for Data-driven Innovation

Expanding knowledge and ethical capacity **outside** of science and technology is critical for society.

Participation *with* engineering, data science and AI communities in policy, commercial and research spheres.

Support expanding social capacity for the responsible use of AI **on the ground** integrated into every day life.

@ginasue@OllOxfordGina.Neff@oii.ox.ac.uk

A 4

B4

.

A 3

83

1) What and whose goals are being achieved or promised through

- 2) What structured performance using
- 3) What division of labour
- 4) Under whose control and
- 5) At whose expense?

| What do I need to know? | Why do I need to know this? | What kind of data will answer the questions? | Where can I find the data? | Whom do I contact for access? | Timelines for acquisition |
|---|---|--|--|--|--|
| What do students plan to do after high school? | To assess the degree to which coherent post–high school career planning affects high school completion | Student survey; follow-up survey of students attending college and getting jobs | Counseling offices; Tribal Social Services office; Dept. of Probation; Alumni Association | Homeroom teachers; school personnel; parents; former students; community social service workers | Student survey, first week in May Follow-up survey, summer and fall |
| What do teachers think about their students' capabilities? | To assess teacher expectations of student success | Teacher survey; teacher interviews | _ | Building principals; individual classroom teachers | Teacher interviews, November (subgroup) Teacher survey, April (all teachers) |
| What do teachers know about the home culture of their students? | To assess teachers' cultural awareness | Teacher interviews; teacher survey; logs of participation in staff development activities | Individual teachers' classrooms and records | Building principals; individual classroom teachers; assistant superintendent for staff development | Teacher interviews, November (subgroup) Teacher survey, April (all teachers) |
| What do teachers do to integrate knowledge of the student's home culture community into their teaching? | To assess the degree of discontinuity between school culture and home culture | Teachers' lesson plans; classroom observations; logs of participation in staff development activities | Individual teachers' classrooms and records | Building principals; individual classroom teachers; assistant superintendent for staff development | Lesson plans, Dec.–June Observations, Sept. 1–May 30 Staff development, June logs |

SOURCE: This figure was published in *Ethnography and Qualitative Design in Educational Research*, 2nd ed. by M. D. LeCompte & J. Preissle, with R. Tesch. Copyright 1993 by Academic Press.

1) What do I need to know 2) Why do I need to know this 3) What kind of data will answer these questions? 4) Where can I find these data? 5) What are access and timeline considerations

Topics for Discussion

Discussion 1

- 6-7 groups
- Each group 1 topic
- Use Framework 1 and / or 2 for generating research ideas for key research questions
- Think about positivist/post-positivist; interepretive; transforming; pragmatist approaches.
- Write a research question for a doctorate
- Report out topic and question

Broad Topics

- 1. Autonomous vehicles
- 2. Use of AI in hospitals or in medicine
- 3. Smart replies in email (e.g., Gmail automated response options)
- 4. Smart home devices (e.g., Amazon Ring)
- 5. Al-enhanced systems used by workers (e.g., Boeing 737 Max)
- 6. Al recommendation systems (ie FaceBook newsfeed)
- 7. Al in financial sector products

Discussion 2

- Same groups
- Same topics
- **Pick at least two different methods** for answering the question from session 1.
- What are advantages and disadvantages of these methods?
- Be ready to report out.