Data for Diversity-Aware Technology: Ethical Considerations

Insights from the project WeNet – The Internet of Us

Author: Laura Schelenz, International Center for Ethics in the Sciences and Humanities, Tübingen, Germany

Data for Diversity-Aware Technology

Ethical Challenges

Operationalization of diversity

Diversity-*awareness* in tech development teams

Collection of large amounts of sensitive data

Pattern inference through machine learning

Protection of privacy of data subjects, data minimization

Representation of minorities in the dataset

© Johannes Ortner via Flickr

WENET
INTERNET OF US

WHAT IS DIVERSITY-AWARE TECHNOLOGY?

Diversity-aware technology…

- leverages the diversity of technology users to their advantage
  - Diversity helps achieve a «good» outcome = diversity as **instrumental** value
- can help reduce computer-mediated bias against certain social groups
  - Diversity realizes the goal of inclusion and representation of minorities = diversity as instrumental/**intrinsic** value
- mirrors the diversity of its users → how can we **operationalize** diversity?

What kind of diversity?

- WeNet understands diversity in terms of social practices

  = routine behavior

  e.g. cooking, riding a bike

- Large amounts of sensitive data: eating habits, transportation, shopping, use of university buildings, student performance

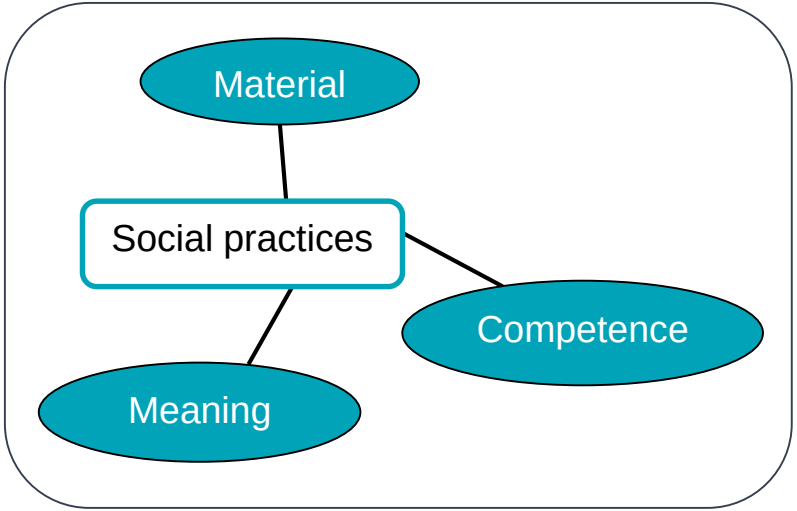
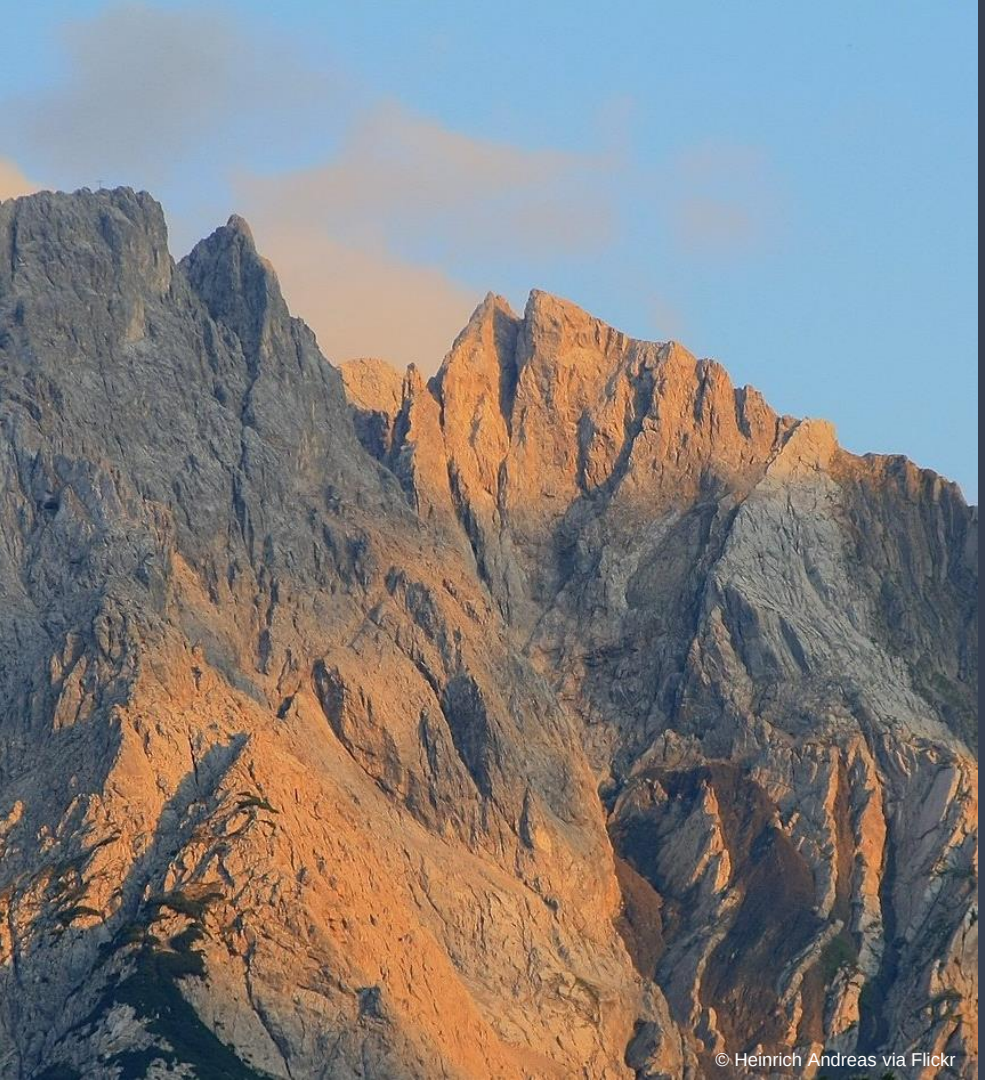


Figure 1: Operationalization of social practices in WeNet



WENET
INTERNET OF US

ETHICAL CHALLENGES

Ethical challenges

- the need to **collect massive amounts of sensitive data**

→data minimization, data protection, and privacy rights

- how to **account for minorities in the dataset**

→implicit bias and constraints of category-building

- how to **account for minorities in computer models**

→machine learning and statistics

Ethical challenge #dataprotection

Diversity-aware technology poses risks to data subjects

Misuse, loss, hacking of information

particularly risky if sensitive data is involved

Easy identification of data subjects

the rarer and more “dispersed” the data points, the easier it will be to trace the information back to the data subjects

Discrimination of marginalized groups

the more data that is “out in the open”, the more data can be used against a person

Easy to circumvent Article 5, data minimization

risk in claiming the need to collect vast amounts of data for diversity-aware technology

Ethical challenge #equalrepresentation

Implicit bias = scripts that prevent us from recognizing “full” diversity

Implicit bias

stems from “schemas” that we internalize at a young age and that are activated subconsciously; schemas help interpret a situation (cf. Sally Haslanger)

Prejudice
Racism/Sexism
→ Structural Injustice

Example from WeNet: social practice of “working”

contract-based employment vs. reproductive labor

- social practices are coded to ideas of gender
- implicit bias in the operationalization of diversity may result in the marginalization of female users

George (user)

- student assistant (5hrs a week) at a local NGO
- has a one-year contract
- wants to gain experience and extra money

Aisha (user)

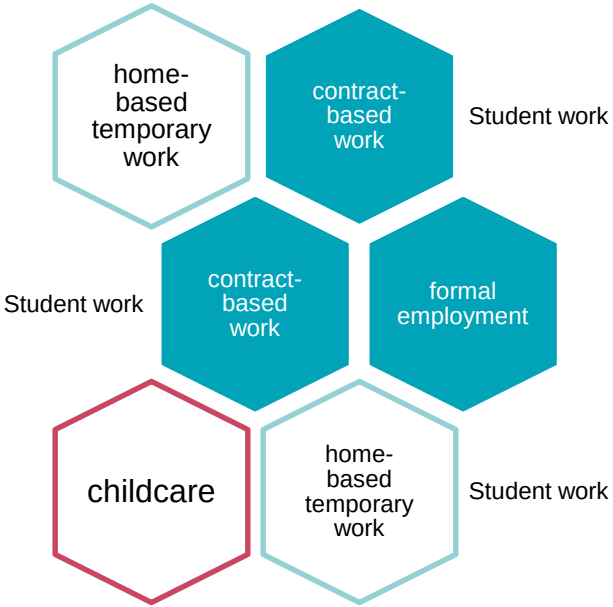
- takes care of her grandma every weekday between 7 and 9 pm (when her mother has to leave for her night shift)
- annoyed because cannot meet friends

Ethical challenge #algorithmicjustice

Computer models built from data sets must be diversity-aware

Machine learning and pattern recognition

- Diversity represented in the dataset may be further reduced by machine learning methods
- Algorithms may be optimized for the majority of the population but not minorities




© Damien Roué via Flickr

WENET
INTERNET OF US

RECOMMENDATIONS

Recommendations

- ✓ Diversity-aware technology needs interdisciplinary cooperation
- ✓ Develop diversity-aware technology that leverages diversity for a “good” outcome *and* ensures non-discrimination
- ✓ Protect data subjects’ privacy, explore innovative solutions that help represent diversity by collecting *less* data
- ✓ Develop a data collection plan that explicitly seeks to reduce bias in the dataset; answer the question “How do we account for minorities in the dataset in a way that properly represents them?”
- ✓ Test how the computer models fare with regard to fairness; answer the question “How do our models affect minorities and is there disparate treatment resulting from our technology?”
- ✓ Increase diversity-*awareness* in tech development teams: provide training to enhance sensitivity to questions of gender, race, and class discrimination

GET IN TOUCH

Website
www.internetofus.eu
www.izew.uni-tuebingen.de

Email
laura.schelenz@uni-tuebingen.de

Twitter
@WeNetProject
@LauraSchelenz

WeNet project is funded by the EU's Horizon2020 programme under Grant Agreement number 823783.

WENET
INTERNET OF US

THANK YOU!