ITUKALEIDOSCOPE

SANTA FE 2018

Machine learning for a 5G future

A Gendered Perspective on Artificial Intelligence

Smriti Parsheera
National Institute of Public Finance & Policy
smriti.parsheera@gmail.com

26-28 November
Santa Fe, Argentina

Overview

- The **origins of AI** and its **gender dimensions**

- The **questions**:

  - **Who** undertakes the process of knowledge-making in AI?
  - **What** research agendas are being set?
  - **How** are the research goals being realised?
  - Do these decisions reflect the **diverse experience and needs** of human society?

- A proposal to re-envision AI from a gendered perspective -- to **eliminate rather than reinforce** human biases

What is AI? - Many different perspectives

- **1950** - **Turing Test**

- **1955** - **Dartmouth proposal** described itself as:
  - a "*2 month,* **10 man study** *of artificial intelligence*"
  - a "*proposal to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves*"

- **Stanford's 100 year study** – *AI is what AI researchers do*

- **Several questions and criticisms**

The "context" of AI research

"*They were interested in intelligence, and they needed somewhere to start. So they looked around at who the smartest people were, and they were themselves, of course. They were all essentially mathematicians by training, and mathematicians do two things - they prove theorems and play chess. And they said, hey, if it proves a theorem or plays chess, it must be smart*."

– Bob Wilnesky, AI researcher

The gender question: fair representation?

- The **gender imbalance** in STEM research:
  - Women constitute **less than 29 percent** of scientific researchers globally
  - Developing countries reflect lower figures.
  - IEEE's AI Hall of Fame - did not include a single woman
  - ITU just elected the first woman in its top management

- The **effects of under-representation**:
  - **Under-representation of ideas** in setting AI agendas
  - Systematic **discrimination** in the work environment - salaries, promotions and harassment
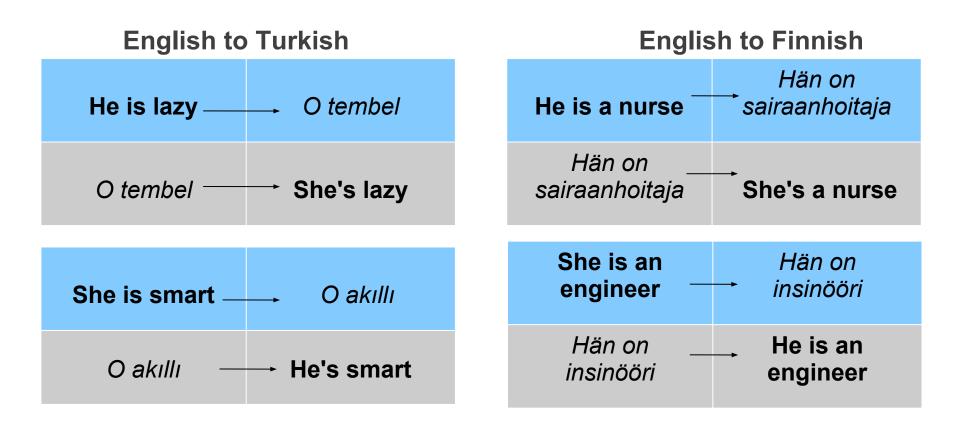 - The **leaky pipe** problem of STEM

Constraints of training datasets

- Training data sets reflect the **perceptions and realities of the real world**

- Can end up strengthening and reinforcing society's **existing biases**

- **Some illustrations:**

    - **Man: Women :: Computer programmer: Homemaker** (*Bolukbasi et al, 2016*)

    - **Facial recognition tools** - 20-34 % error rates for darker skinned females (*Buolamwini, 2018*)

    - Why do most **voice assistants** have female sounding voices?

Illustration -- Google Translate

English to Turkish

He is lazy ⟶	*O tembel*
O tembel ⟶	**She's lazy**

She is smart ⟶	*O akıllı*
O akıllı ⟶	**He's smart**

English to Finnish

He is a nurse ⟶	*Hän on sairaanhoitaja*
Hän on sairaanhoitaja ⟶	**She's a nurse**

She is an engineer ⟶	*Hän on insinööri*
Hän on insinööri ⟶	**He is an engineer**

Re-envisioning AI from a gendered perspective

- **Representation of women in AI research**

  - specialized programmes and increased intake in educational institutions
  - funding support
  - mentorship initiatives
  - equal opportunities in the job market

- **Towards more inclusive research agendas**

  - participation of businesses, governments and the public
  - financial support, ethical frameworks, open data resources
  - rooted in the social, cultural and institutional context

The three-step process

STEP I

Embed "**fairness by design**" in AI frameworks (*Abbasi et al, 2018*)

- **Cross-disciplinary teams** of data scientists and social scientists;

- Identifying and addressing the biases of **human annotators**;

- Building f**airness measures in the assessment metrics** of the program;

- Critical **mass of training samples** to meet fairness measures; and

- Adopting **debiasing techniques**.

The three-step process

STEP 2

- Research and development in **fairness enhancing tools**

- Mechanisms for **mainstreaming of solutions** and making them easily accessible to the research community

- **Example:** Google's What-If tool

STEP 3

- Measures to **fix the underlying data sets** that inform AI research

- **Example**: feminization and neutralisation of language

Conclusion

- **Gender** of AI researchers has played a **key role** in determining the contours of the discipline

- The problem of AI is **not just about the representation of women** – it is also about the **gender distortions in training data sets**

- Conscious and unconscious **human biases** seep into AI artifacts and need to be addressed through **systematic interventions - 3 step process**

- This debate is of equal relevance to **other dimensions** like education, race, class, religion and nationality and the **intersectionality** of these factors

ITUKALEIDOSCOPE
SANTA FE 2018

Thank you