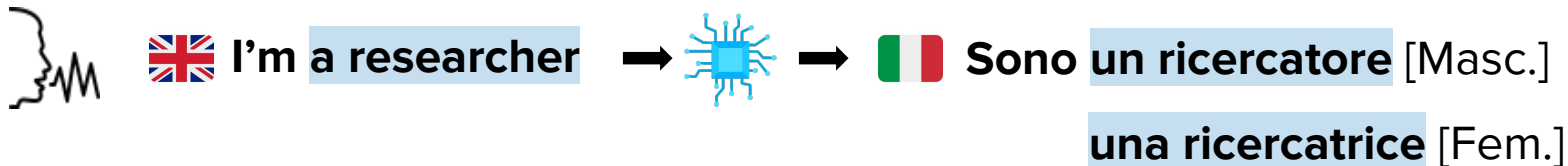


AI Ethics Today

# **Gender bias in ST**

# Speech translation of gender



## What is the desired model behavior?

### Current state

- Models default to masculine or produce stereotypical gender translations
- Rely on speaker's vocal characteristics

### Possible interventions

- Generate gender-neutral solutions (**Sono un\* ricercat\***)
- Ask users to provide gender metadata

# Equity

- *Max advantage to the least advantaged*
- *Inclusiveness: no one left behind*
  - ◆ No vocal traits -> avoid gender recognition errors in people with laryngeal disorders or transgender
  - ◆ Ask the user for the preferred gender (when possible)
  - ◆ Gender-neutral outputs

# Social well-being

- *Maximize the overall happiness*
  - ◆ Use all the possible features -> improve the performance
  - ◆ Choose the gender of the majority if not clear (eg: women? cisgender people?)
  - ◆ Ask the user for the preferred gender (when possible)

# Autonomy

- *Right to self-determination*
- *Individual's right to make informed decisions*
  - ◆ Let the user specify their gender, and the gender of all the other speakers
  - ◆ Warn the user that this means sharing sensitive data
  - ◆ Not mandatory, but the software won't work properly otherwise

# Privacy

- *Control over access to one's personal information*
- *Individual's right to choose what information to share*
  - ◆ The user shares sensitive information (gender) with the company
  - ◆ In some countries being transgender is still illegal
  - ◆ As data tools improve, more and more can be done with less information

# It is a matter of choice

There is **no** possibility to maximize all the principles at once.

1. **[DEFAULT]** Models produce default to masculine or stereotypical gender translations (Current State)
2. **[VOCAL]** Rely on speaker's vocal characteristics (Current State)
3. **[NEUTRAL]** Generate gender-neutral solutions (*Possible interventions*)
4. **[USER]** Ask users to provide gender metadata (*Possible interventions*)

# Final Considerations

	[DEFAULT]	[VOCAL]	[NEUTRAL]	[USER]
• Equity	✗	✗	✓	✓
• Social well-being	✗	✓	✗	✓
• Autonomy	✗	✗	✗	✓
• Privacy	✓	✓	✓	✗