TOWARDS FAIRNESS IN AI FOR PWD: A RESEARCH ROADMAP

Group Members: Helena Donaldson, Weston Seybold, Havalock Yin

INTRODUCTION OF PAPER

Overall Idea:

- AI systems can enhance the lives of people with disabilities (PWD) but may also introduce bias or exclusion.
- Current AI technologies often fail to work equitably for PWD due to underrepresentation in training data and evaluation.

Research Focus:

- Identifies gaps in AI fairness for PWD across various AI technologies: computer vision, speech recognition, text processing, and more.
- Highlights risks like exclusion, misrepresentation, and safety concerns for PWD in AI systems.

THE PAPER'S ROADMAP

Roadmap for Fairness:

- **Goal 1**: Identify fairness and inclusion issues in AI for PWD.
- **Goal 2**: Test failure scenarios and bias mitigation techniques.
- **Goal 3**: Create inclusive datasets.
- **Goal 4**: Develop new models and techniques to improve AI fairness for PWD.

THE MOTIVATION OF THE PAPER

Ensures AI benefits are accessible to everyone, regardless of ability.

Promotes ethical AI design that prioritizes inclusion.

COMPUTER VISION

PAPER DISCUSSION

- Many Computer Vision systems are not trained with diverse data, leading to biases against PWD.
- Improvements in data inclusion and testing are critical to make these systems work fairly for everyone.

COMPUTER VISION – RESEARCH EXPANSION

Creative Expression: Personalization of AI Tools, Collaboration, Not Substitution, Promoting Diverse Datasets

Human Contact: Dependency on AI for Social and Emotional Support, Social Integration Challenges, Dependency on AI for Social and Emotional Support

Accountability: AI System Failures, Bias and Discrimination, Lack of Transparency in AI Decisions

SPEECH SYSTEMS

PAPER DISCUSSION

- Speech systems often **lack inclusivity** for PWD, with biases leading to poor recognition and interpretation.
- The development of **personalized models** and inclusive training datasets is essential to improving fairness in these systems.

SPEECH SYSTEMS – RESEARCH EXPANSION

Creative Expression: Loss of Human Articulation and Expression, Loss of Ideas and Emotion

Human Contact: Difficulty fostering relationships, Human Reliance on AI Speech Systems

Accountability: Biases in speech and ideas present, Biases in culture and articulation, accountability for speech systems

TEXT PROCESSING

PAPER DISCUSSION

- Text Processing systems include biases and face challenges when used by those with cognitive or intellectual disabilities
- Systems could have a positive impacts on reading, writing, and communications for PWD
  - These systems may also be detrimental if they produce inaccuracies and limit voice/emotion

TEXT PROCESSING-RESEARCH EXPANSION

Creative Expression: Systems limit individual voice/emotion present in writings and other generated texts

Human Contact: Greater communication for those with cognitive/intellectual disabilities

Accountability: Responsibility for biases and inaccuracies in text processing systems may be unclear

INTEGRATIVE AI

PAPER DISCUSSION

- Information Retrieval (IR) AI, both content-based and behavior-based, could have bias on the contents being provided to PWD. The fact that IR tools are often related to commercial activities such as ad posting also increases the chance of discrimination towards PWD
- Conversational Agents could amplify existing biased against PWD through returning stereotyped content in conversations, resulting poor user experience. Moreover, conversational agents are ought to have expressive medium to provide supportive conversations, such as sign languages (for people who are deaf) or pictures and/or icons (for people with aphasia or autism).

INTEGRATIVE AI – RESEARCH EXPANSION

Creative Expression: Formation of so called the golden chamber. Limiting the information a PWD could receive due to bias.

Human Contact: May affect PWD due to the bias information provided. Or could improve the social contacts of PWD with supportive conversations with expressive media

Accountability: Insuring IR tools are performing correctly event when related to commercial activities. Prevent conversational agents from malicious inputs and follow the rules of accessibilities.

OTHER AI TECHNIQUES

PAPER DISCUSSION

Outlier Detection Algorithms

- Lack of representation of PWD may lead to **flagging**
- Example: Atypical performance times

Use of non-representative training data sets

Aggregate metrics can fail to capture PWD needs

- Used to evaluate effectiveness of AI systems
- Hides how performance varies across a group

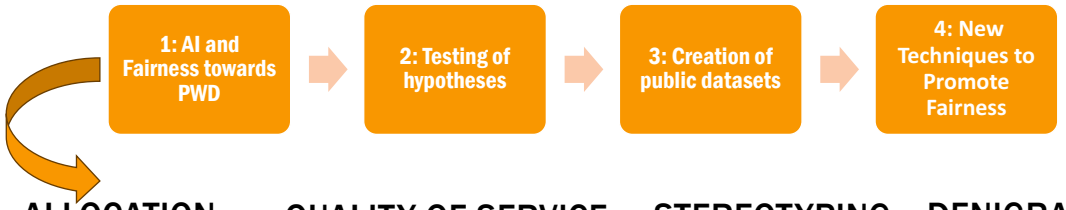
OTHER AI TECHNIQUES: RESEARCH EXPANSION

Creative Expression: Impact of non-representative data sets, poor metric use to evaluate AGI systems

Human Contact: Impact of non-representative data sets, poor metric use to evaluate AGI systems

Accountability: Ensuring accountability for proper collection and use of representative data sets

DISCUSSION



ALLOCATION

Example:

Using an incorrect prediction of emotional state for PWD in hiring

QUALITY OF SERVICE

Example:

Voice-activated speakers that can't recognize speech from PWD

STEREOTYPING

Example:

IR Systems returning biased content regarding PWD

DENIGRATION

Example:

Inaccurately flagging PWD input as outliers

DISCUSSION

CHALLENGE TO CONSIDER:

Is it more beneficial to create general models fair across demographic groups or creating personalized models for use by certain demographic groups?

EXAMPLE

- Automatic Speech Recognition tailored for the Deaf community.

DISCUSSION: RESEARCH EXPANSION

Creative Expression: Greater PWD creative output addressing stereotyping stories

Human Contact: Issues regarding emotional communication

Accountability: Measures of accountability for instances of quality of service disparities

LIMITATIONS OF PAPER

1. Technological developments examined in the paper doesn't explore impact of newer technologies (paper was written in 2019)
2. Lack of focus on eventual impact of AI fairness issues with PWD (effects discussed are more immediate than long-term)
3. Doesn't address difficulty of creating inclusive datasets with PWD

CITATIONS

Guo, A., Kamar, E., Vaughan, J. W., Wallach, H., & Morris, M. R. (2020). Toward fairness in AI for people with disabilities SBG@ a research roadmap. *ACM SIGACCESS accessibility and computing*, (125), 1-1.