Introduction

The Artificial Intelligence (AI) community recognises the need for more transparent and human centered systems, but at the same time, AI field suffers from a ‘diversity crisis’ in terms of AI researchers’ demographics [1]. In our project, we bring together research from social and cognitive sciences to investigate if one’s social and cultural background, which is known to constitute one’s perceptual and cognitive tendencies, also influences one’s explanation needs in human-AI interaction and if current diversity crisis can result in systems geared towards a specific user group. Our aim is to contribute to designing more inclusive and human-centered AI systems.

Current XAI

Current research in Explainable AI recognizes that:

- Users prefer different explanations because they have different goals and levels of expertise.

Our vision

We aim to research and develop XAI with an understanding, that:

- Users prefer different explanations because they have different goals and levels of expertise and because they are embedded in different socio-cultural structures that affect their cognitive style.

What is an (AI) Explanation?

1) Product
2) Process
   a) Cognitive process by Explainer (AI)
   b) Social process
   c) Cognitive process by Explainee

Cognition is Cultural

Receiving and processing an explanation is a cognitive process.

Cognitive processes are shaped, among others, by culture.

- Cultural habits and practices affect processing
- External representations affect (internal) processes
- Content of cognition affects processing

Our vision

We aim to investigate potential differences in explanation preferences between different communities of users characterised by various socio-cultural factors.

Ultimately, our goal is to democratize AI and empower users by providing explanation that align with users’ cognitive tendencies.

Conclusions

We aim to research and develop XAI with an understanding that:

- Users prefer different explanations because they have different goals and levels of expertise and because they are embedded in different socio-cultural structures that affect their cognitive style.

References


Acknowledgments

This work was supported by UKResearch and Innovation [grant number EPS0233561], in the UKRI Centre for Doctoral Training in Safe and Trusted Artificial Intelligence (www.safeandtrustedai.org).