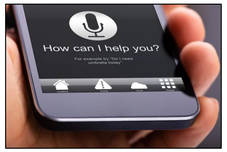
Northeastern University
Khoury College of Computer Sciences

VOICE ASSISTANT USER PROFILING

SHRIYA DHAUNDIYAL

INTRODUCTION

- Voice Assistants are becoming increasingly pervasive in our everyday lives. While convenient, these systems also raise important privacy concerns.
- One of them is understanding to what extent they use or share information from past user interactions.



RESEARCH QUESTIONS

- If voice assistant interactions lead to user profiling.
- When and how these profiling activities happen?
- What are the risks associated with such profiling?
- If the said activities have been disclosed on the privacy policy of the devices.

FOCUS TECHNOLOGY

Google Assistant

RESEARCH METHODOLOGY

- Research into disclosed privacy policies of Amazon, Google and Apple and their VA technologies.
- Searched for disclosure of profiling tags/categories on these platforms.
- Analysis of personal data on these platforms to look for profiling and association.
- Conduct experiments with VAs to generate a list of interest questions.
- Experimented with fake profiles to see trends for profiling.
- Classify target/sensitive tags to build interest personas.
- Tags are compounded to infer interests for the users.

CREATING INTEREST PERSONAS

LUXURY PERSONA

- Age : 42 years
- Gender : Female (Input NEUTRAL)
- Income : High Income
- Marital Status : Single
- Demographics : United States

AFTER VA INTERACTION

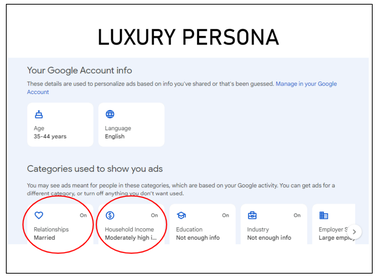
- Gender : Male
- Marital Status : Engaged
- Income Tag : High – Upper High Income
- Home Ownership : Home-Owner

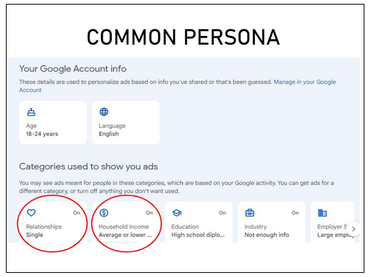
COMMON PERSONA

- Age : 22 years
- Gender : Male (Input NEUTRAL)
- Income : Middle Income
- Marital Status : Single
- Demographics : Asia

AFTER VA INTERACTION

- Gender : Female
- Marital Status : In a relationship
- Income Tag : Average – Lower Income
- Home Ownership : Renters





CONFIRMATION OF PROFILING

Observed differences in recommendation of products by VA upon neutral interaction.

YouTube advertisements targeted for each persona based on their interests.

YouTube

Changed recommendation feed on Google News and YouTube based on personas.

CONCLUSIONS

- Voice assistant interactions leads to user profiling.
- Profiling tags can be changed based on the type of searches – algorithm can be made to believe in a fake persona.
- Policy disclosure is vague and finding profiling information is hard or available only upon request.

ACKNOWLEDGEMENTS

- I would like to thank Professor David Choffnes, Dr. Daniel Dubois and Tianrui Hu for their constant support and guidance with the research project.
- Special thanks to all members of the Mon(IoT)r Research Group.