'That's Not Me': Surprising Algorithmic Inferences

Samantha Hautea, Anjali Munasinghe, & Emilee Rader
Behavior, Information, & Technology Lab

MICHIGAN STATE
U N I V E R S I T Y

Study Aims

Identifying surprise can be a way of revealing existing expectations. Individuals are surprised **by what they don't expect**.

Inferences are a type of information that is **algorithmically derived** from user data.

Platforms like Google and Facebook provide pages where users can see these inferences:
- What connections do users perceive between their online behavior and assigned inferences?
- What inferences surprise users?
- What does this reveal about user perceptions and their privacy expectations?

Methods

Survey
Male: 28
Female: 66
Other: 1
Average Age: 39
Facebook: 44
Google: 51

Interview
Male: 9
Female: 12
Other: 0
Average Age: 42
Facebook: 8
Google: 13

Inductive Qualitative Analysis and Iterative Coding

Figure 1: Data collection and analysis steps.

- Survey respondents (n=95) downloaded their Facebook/Google inferences page (randomly assigned) then rated how relevant the inferences were to themselves
- Interviewees (n=21) selected from respondents were asked to discuss a report generated from their inferences and their survey responses

Findings

After thematic coding and analysis, we identified 4 recurring themes around surprising inferences:

Irrelevant Inferences

"I feel like there's things on here that I'm like, 'Oh, yeah. That's clearly not me.'"
P21, W, Google, age 40

Close friends of people with birthdays in a week

Life, Physical and Social Sciences

Architecture and Engineering

Figure 2: Samples of inferences provided by Facebook

Outdated Inferences

"There's just things that were hard for me to answer whether they really say who I am ... in the past they have been some- times, but maybe not today."
P07, M, Google, age 71

Related to Friends & Family

"Some of these could be things that don't really pertain to me just because maybe there's something that my fiance asked me to look up on my computer."
P04, W, Google, age 32

Parenting

Residential Sales

Strategy Games

Figure 3: Samples of inferences provided by Google

No Connection to Online Activity

"I have an adult son who lives at home but ... I'm not looking up parental tips, I'm not looking up stuff about babies or children."
P11, M, Facebook, age 63

Conclusion

Participants mostly expected inferences to be based on their online actions tracked over time.

Other sources of inferences were harder to grasp

Participants differed in their beliefs (and expectations) about what is possible with their data.

This view of inferences limits an individual's ability to recognize and prevent when privacy violations occur.