

Your Echos are Heard

Tracking, Profiling, and Ad Targeting in the Amazon Smart Speaker Ecosystem

Umar Iqbal, Pouneh Nikkah Bahrami, Rahmadi Trimananda, Hao Cui, Alexander Gamero-Garrido, Daniel Dubois, David Choffnes, Athina Markopoulou, Franziska Roesner, Zubair Shafiq



Overview

- Smart speaker interactions can be used to infer sensitive information about users
- There is limited transparency in data collection and usage practices of smart speaker platforms
- We build an auditing framework to bring transparency in smart speaker platforms
- We test on Amazon smart speaker ecosystem and find that Amazon and third parties collect smart speaker interaction data
- We also find that Amazon processes voice data to infer user interests and uses it to serve targeted ads on-platform as well as off-platform

Privacy Issues

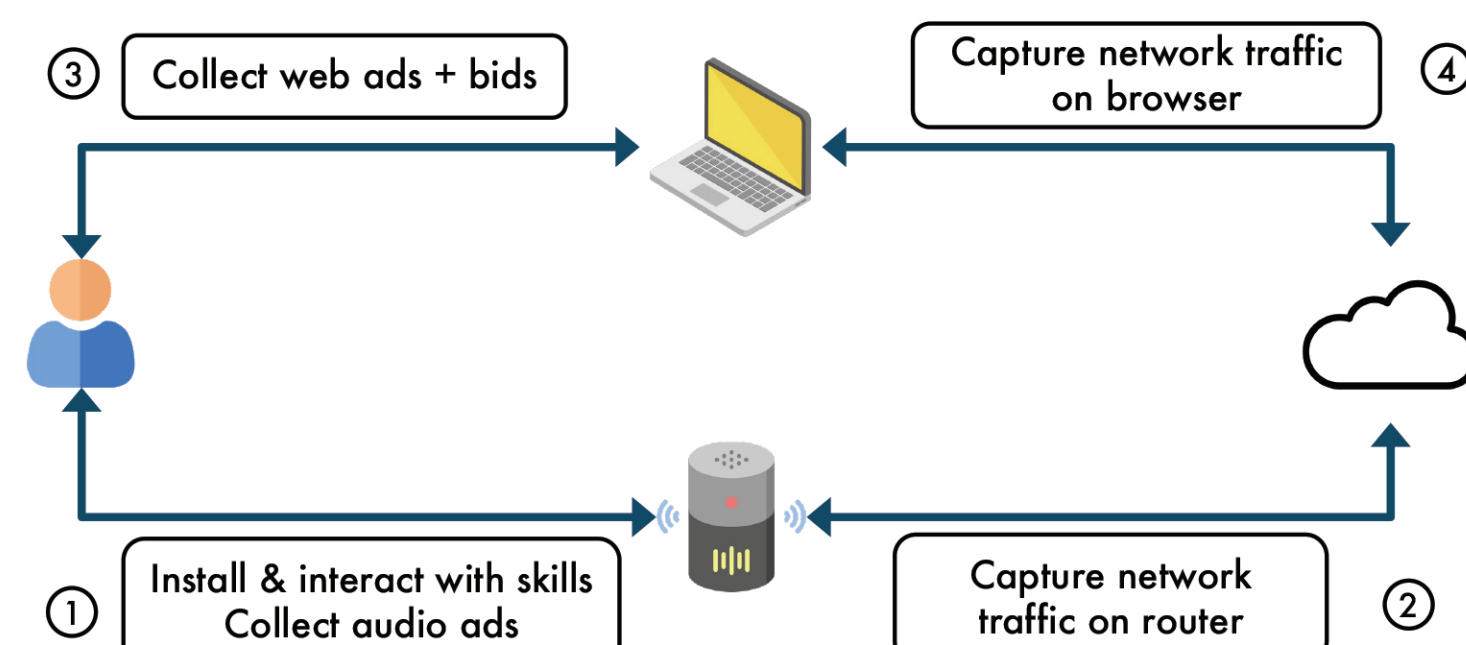
- Voice input can leak sensitive information
 - Voice recordings:** Tone can leak personality traits
 - Transcripts:** Content can leak personal information
 - Metadata:** Interaction logs can leak personal information
- Amazon has patented several privacy-infringing practices to monetize smart speaker interactions
 - Advertising products based on inferences from physical & emotional characteristics of users' voices
 - Targeting cough-drop ads at users with colds

Auditing is hard!

- Problem:** Limited transparency from Amazon
 - Users may not trust the platform
- Goal:** Independent & repeatable audits on unmodified off-the-shelf devices
- Challenge:** Smart speakers do not provide any interfaces to analyze data

Auditing Framework

- Key idea:** Leak data and measure its collection, usage, and sharing
- Leaking data
 - Install skills from different categories
 - Interact with skills by uttering sample utterances provided by skills
- Measuring collection
 - Intercept encrypted network traffic on a custom router
 - Capture unencrypted network traffic by instrumenting Alexa Voice Service SDK
- Inferring usage: Monitor targeted ads to infer data collection and usage
 - Relevance of **ad content** to leaked data
 - High bids** from advertisers when data is leaked



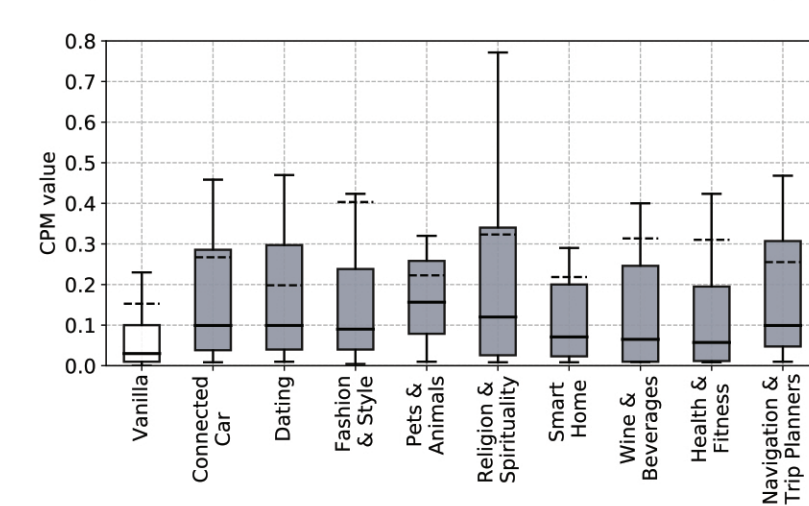
Data is Used For Targeted Advertising

- Amazon processes Echo interactions to infer interest segments
 - First step in targeted ad delivery
- Smart Home:** Electronics, Home, Kitchen
- Fashion & Style:** Beauty, Personal Care, Fashion

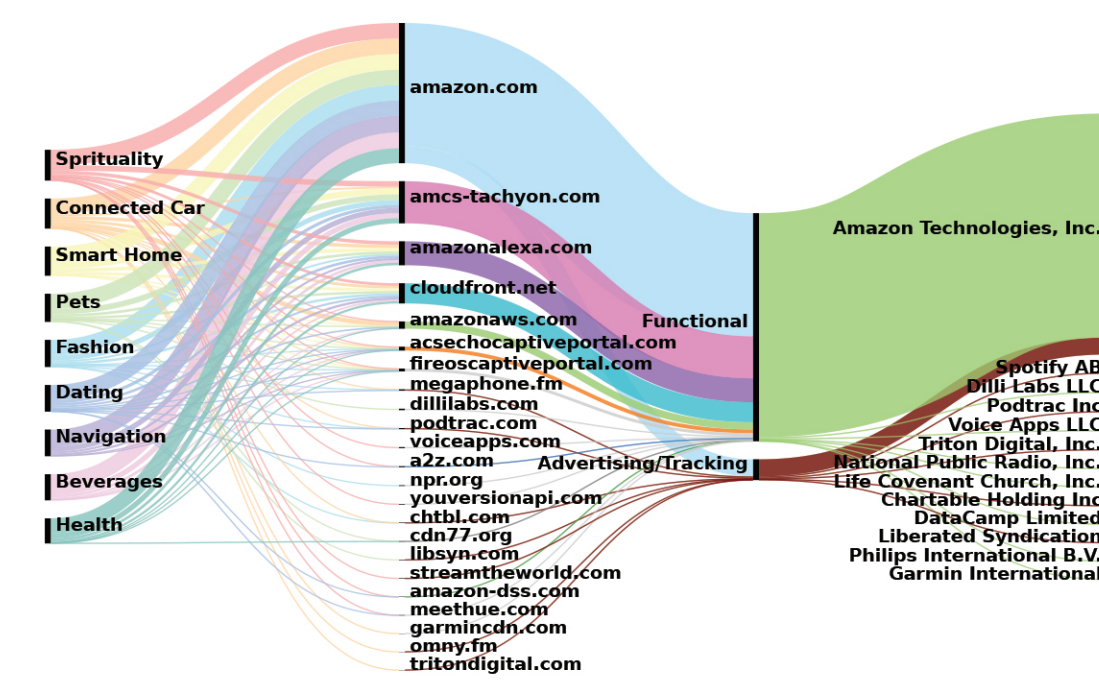


Fashion & Style persona: At Ross find the latest spring trends on the brands you love dresses, tops, accessories, shoes and more. All for less. Get your new spring looks and save 20 to 60% off department store prices. No kidding. Across yes for less

Statistically significant higher bids
30X to 27X higher for some personas



Amazon & Third-Parties Collect User Data



- All skills (have to) share data with Amazon
- Amazon mediates communication b/w skills & users
- 8.3% of traffic goes to advertisers/trackers

Some Disclosures Are Potentially Misleading

- Privacy policy acknowledges collection of Echo interaction data
- No explicit acknowledgement or denial about data usage in privacy policy, Alexa Privacy Hub, and Alexa Device FAQs
- Public statement to NYT: "(We) do not use voice recordings to target ads."
- Response to our research: "record of (transaction) may inform relevant ads"
- People (~63%) don't expect Amazon to use voice recordings or derived information to target ads
- People are uncomfortable with the use of voice recordings (~72%) and derived information (~65%) for targeted ads

"Even when companies say they don't use your (specific) voice, they use ALL your data - including demographics and log in and time questions asked - so they might as well use your actual voice recording!"

Impact

This block contains a collage of news articles and legal documents:

- The Verge:** Researchers find Amazon uses Alexa voice data to target you with ads
- AP:** Amazon keeps growing, and so does its cache of data on you
- la Repubblica:** "Ecco come Alexa rivela le vostre conversazioni ad Amazon, che le usa per calibrare la pubblicità"
- AD-TECH:** Ad Targeting in Amazon Smart Speakers. While we give attention to textual data on the web, many do not know the unique power of echo interactions with smart devices for ad targeting. Today,
- Study:** How Amazon uses Echo smart speaker conversations to target ads
- PRIVACY CON:** UNITED STATES DISTRICT COURT WESTERN DISTRICT OF WASHINGTON. Plaintiffs vs. AMAZON.COM, INC., a Delaware corporation, and AMAZON.COM SERVICES, LLC, a Washington limited liability company. CLASS ACTION COMPLAINT. JURY TRIAL DEMANDED.
- BEUC:** The European Consumer Organisation (BEUC)
- AXIOS:** Lawsuit claims Amazon using Alexa to target ads at customers

Key Takeaways

- Amazon & third-parties (including ad/tracking services) collect Echo interaction data
- Amazon infers user interests from Echo interactions and uses them to target ads
- Interest inference and their usage seems to be inconsistent with Amazon's public statements
- Users are uncomfortable about ad targeting based on Echo interactions
- More information on the website: alexaechos.com
- Funding
 - NSF - CNS-1956393, CNS-1955227, CNS-2103439, CNS-2114230, CNS-1909020
 - CRA CIFellows 2021 under NSF CNS-2127309
 - NEU Future Faculty Fellowship 2021
 - Consumer Reports Digital Lab
- Contact
 - Umar Iqbal
 - Web: umariqbal.com
 - Twitter: @umaarr6