



Tracking, Profiling, and Ad Targeting in the Alexa Echo Smart Speaker Ecosystem

Umar Iqbal

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

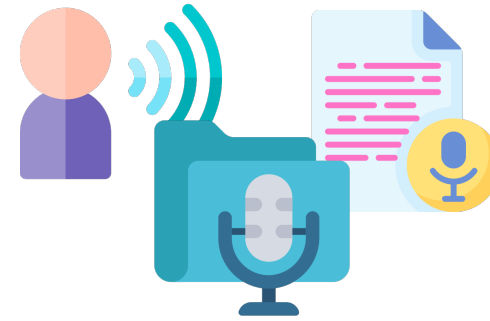
Internet Measurement Conference 2023



Smart speaker usage & data

95 Million

people use smart
speakers in the US^[1]



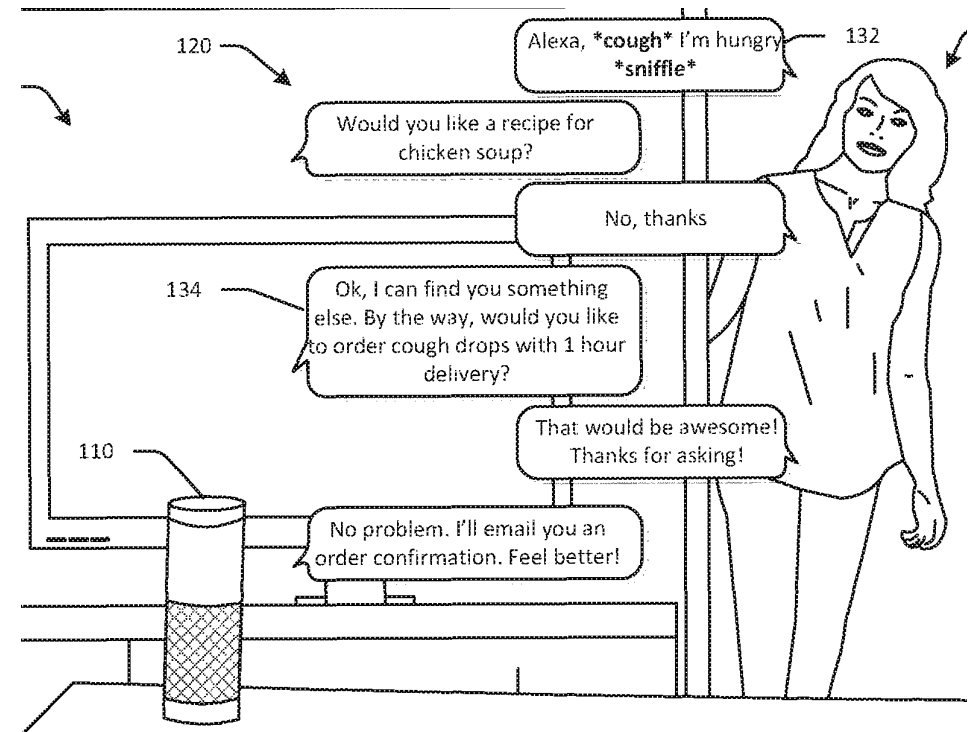
Information in recordings,
transcripts, & metadata

Platform practices

Dangerous Skills Got Certified: Measuring the Trustworthiness of Skill Certification in Voice Personal Assistant Platforms – Cheng et al. CCS '20

SkillDetective: Automated Policy-Violation Detection of Voice Assistant Applications in the Wild – Young et al. USENIX Security '22

When Speakers Are All Ears: Characterizing Misactivations of IoT Smart Speakers – Dubois et al. PETS '20



Patents to infer info. for advertising¹

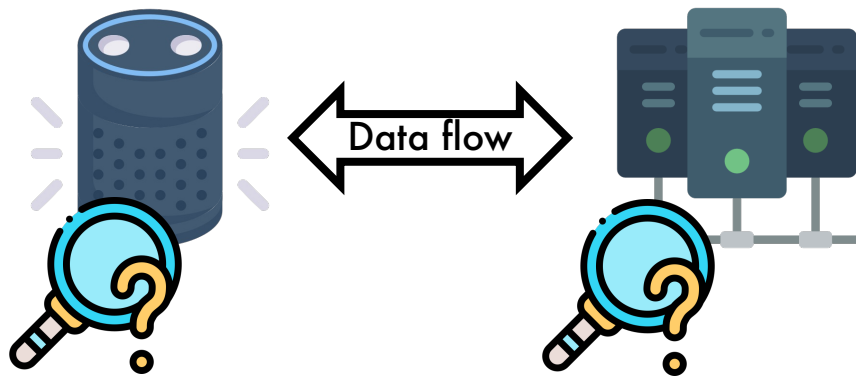
[1] <https://patents.google.com/patent/US10096319B1/en>

Need for transparency



Data collection and usage by platforms and skills

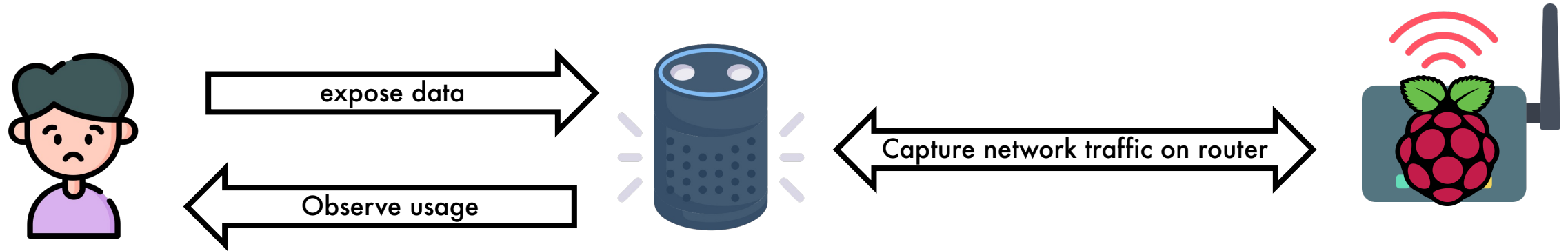
Who collects user data & how is it used?



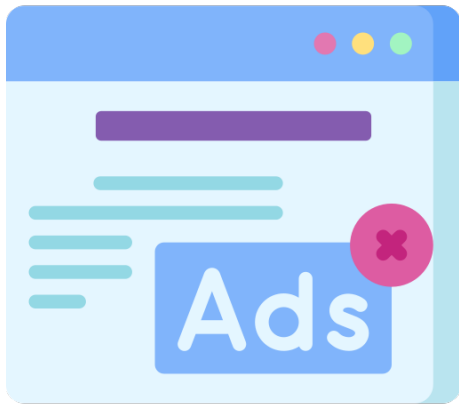
No interfaces to observe data processing

Auditing framework

Key idea: Expose data and look for collection & usage



Measure usage in advertising



Relevance to the exposed content

e.g., clothing ads after interacting with clothing skills

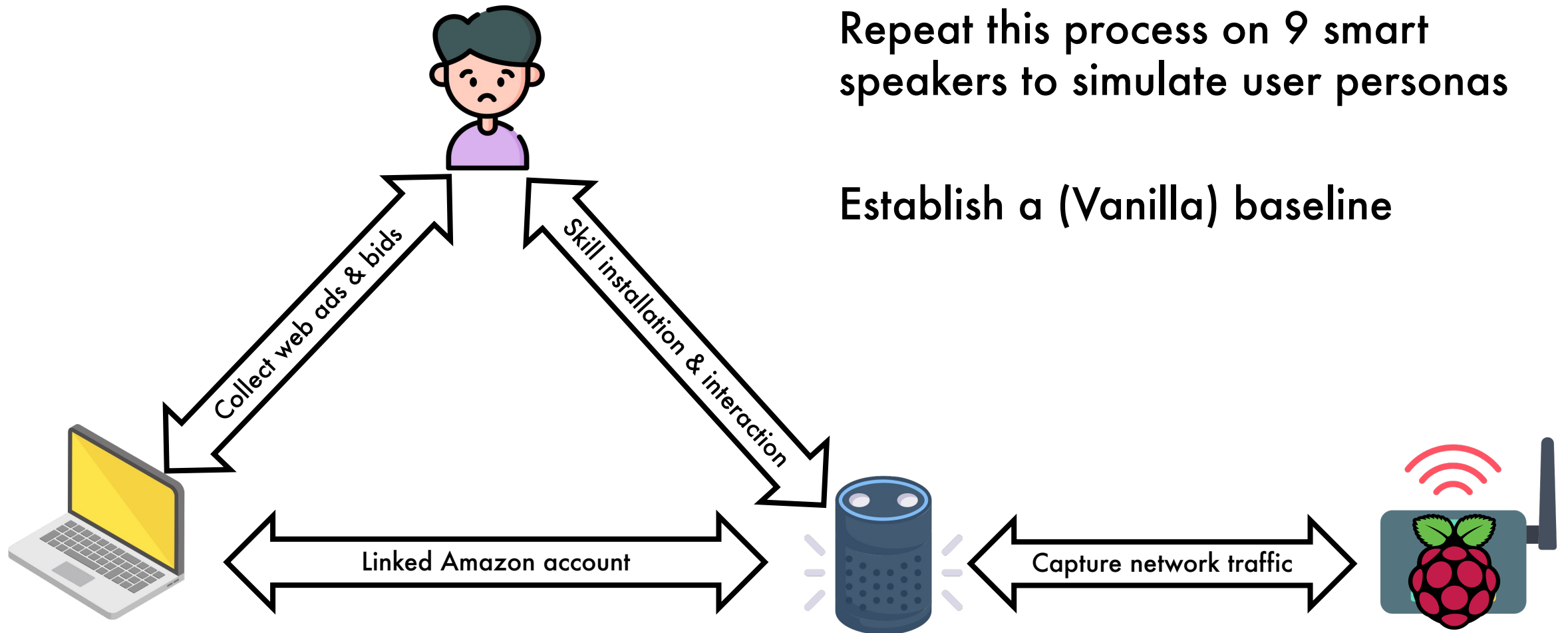
Higher bids when data is exposed

Advertisers pay more when they have information about the user

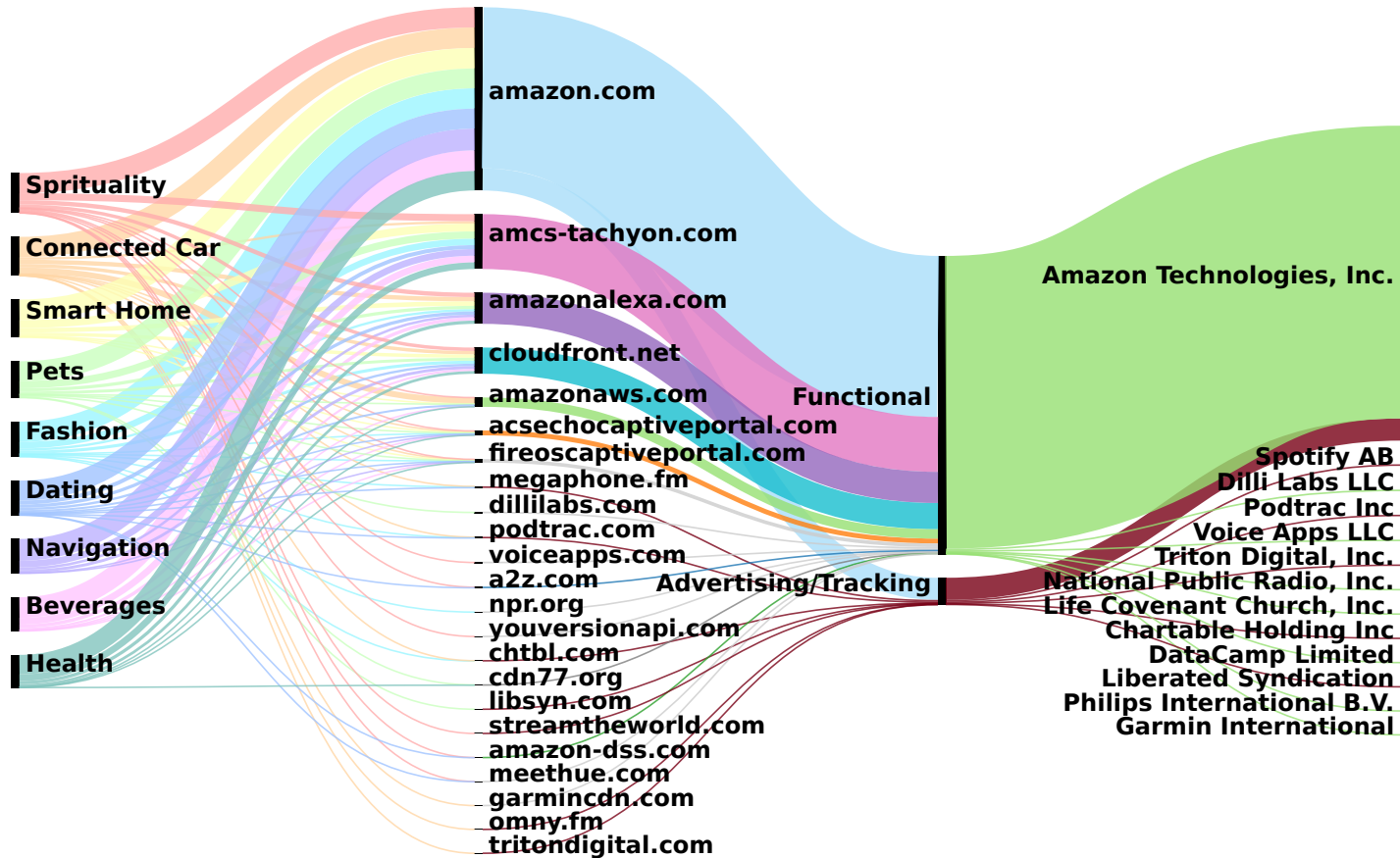


amazon alexa

Exposing data & Inferring usage



Data collection by platform & third-parties



All skills share data with
(mediator) Amazon

8.3% of traffic goes to
advertisers/trackers



Data is used to profile users



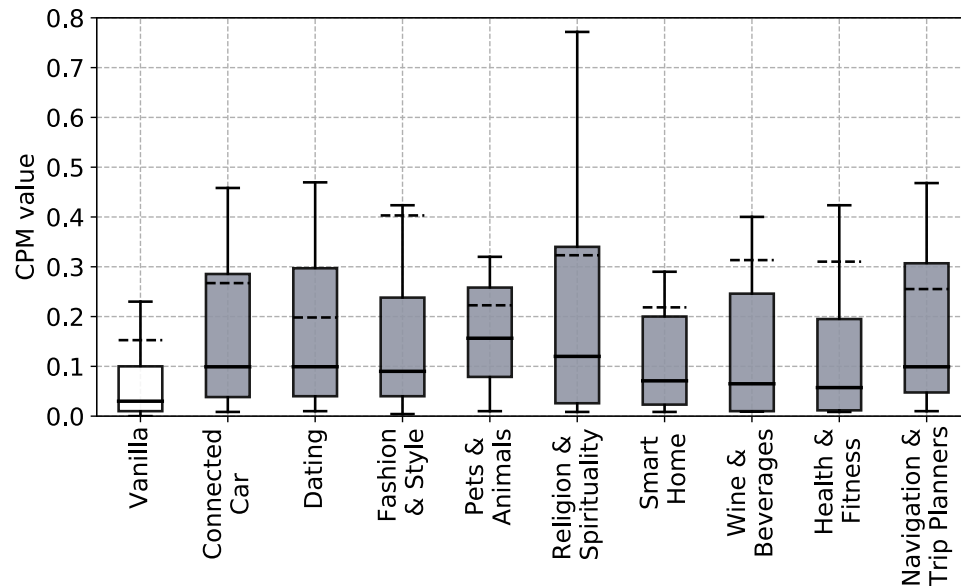
Amazon processes user interactions (e.g., meta data) to infer user interests

Smart Home: Electronics, Home, Kitchen

Fashion & Style: Beauty, Personal Care, Fashion

Data is used in ad targeting

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



Ads relevant to the installed skills
and unique across personas

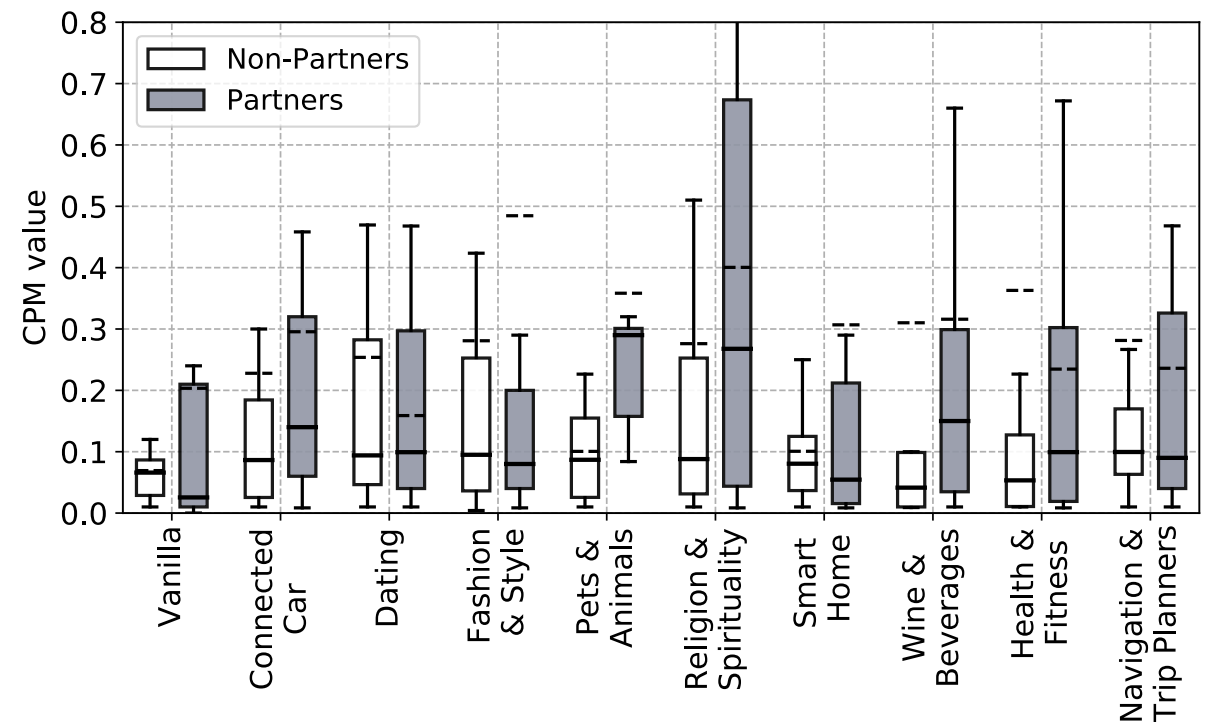


Data is potentially shared

41 third parties synced their cookies with Amazon

For most personas bids were higher from partners

Exceeding 3X for Pets & Animals and Religion & Spirituality



Disclosures by skill vendors



Third parties do not clearly disclose their practices

~58% of skills do not provide a privacy policy

~68% of privacy policies do not mention "Amazon" or "Alexa"



Policheck to evaluate privacy policies

~2% skills clearly disclose their data collection

Disclosures by Amazon



No clear disclosure about the usage of data

Privacy policy
Alexa Privacy Hub
Alexa Device FAQs

The New York Times

“(We) do not use voice recordings to target ads.”

Certain experiences on Alexa, such as streaming music or news videos on your Echo Show device, may contain ads. Some of those ads may be interest-based. Interest-based ads, sometimes referred to as personalized or targeted ads, present you with features, products, and services that might be of interest to you based on your interactions with certain content or services. For instance, if you ask Alexa to order paper towels or for recipe ideas, you might see or hear ads related to cleaning products or cooking utensils. We offer you settings to manage your interest-based advertising preferences.

Managing advertising preferences
on Alexa

Disclosure to regulators



Key takeaways



Framework to audit smart speakers
Requires no cooperation from manufacturer
Generalizable across IoT devices



Alexa Echo interaction data is collected by Amazon and third parties
Amazon uses Echo interactions to infer user interests and target ads



Data collection and usage practices are not clearly disclosed by Amazon and third parties