

Voices and Voids: Subverting Voice Assistant Systems through Performative Experiments

Desjardins, A., Psarra, A., Whiting, B. A.
(2021)

Desjardins, A., Psarra, A., Whiting, B. A.
(2021, June 22). *Voices and Voids: Subverting Voice Assistant Systems through Performative Experiments*. *ACM International Conference Proceeding Series*.
<https://doi.org/10.1145/3450741.3466807>

Scope

Reclaim, examine and transcode VAs data through interdisciplinary performative experiments of Post Internet Art.

- ↳ AI voice assistants data is the **artistic material**
- ↳ Challenge AI and ML technologies

Introduction

Compared to the research done in human interaction with VAs, the data that these devices collect is unexplored.

VA data as the center for understanding the complex interactions between people, large tech corporations, **data economy, hidden labour and surveillance capitalism.**

This interaction are explored through the sonic qualities and semantic meaning of the VA data.

? *When we use voice assistant data as creative material what is revealed about voice assistant systems?*

Online exhibition: <https://voicesandvoids.net/>

why am I infatuated with my rabbit

does the rabbit feel the emotion of love

how do I know when my rabbit likes me

what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate
what's my heart rate

"time": "2019-01-19T22:39:59.288Z",
how wide is the average fingernail

"time": "2019-01-19T22:40:07.364Z",
how many centimeters in an inch

"time": "2019-01-19T22:40:46.519Z",
how many centimeters in a yard

"title": "Said set the timer for 5 minutes",

"time": "2019-01-19T22:42:47.718Z",

kie dokie that is very interesting from the snails perspective

Context – Related Work

New Media Art works such as “[Someone](#)” by Lauren McCarthy and “[Chthonic Rites](#)” by Wesley Goatley use VA technologies to engage with the topics of technological solutionism, power politics, surveillance capitalism and the ghost workers behind these systems.

Musical References: Inspired by the chance and simultaneity of John Cage and Cristian Wolff, the soundscape practice of Hildegard Westerkamp and Maryanne Amacher.

VAs in HCI and design:

- ↳ Main use of VAs are everyday life tasks, and intimate and health related questions.
 - ↳ Trust, respect, conversational expectations set important guidelines for future conversational agents design.
 - ↳ Use of speculative design to imagine the future of VA interaction or to explore them from a non-human perspective.
 - ↳ Concerns regarding AI & ML systems human labour, surveillance capitalism, privacy and security.
 - ↳ Explainable and transparent AI and how to make data trails more visible and actionable.
- ◆ This paper contributes to this context providing a first-hand experience of these topics through artistic exploration.

Methodology

A **multidisciplinary** (new media art + music + interaction design) approach to:

- ↳ research various lines of inquire
- ↳ combined different method to cover more ground exploring VAs and their data

Outcome: 12 different experiments exploring VA's data.

Topics of examination:

1. Intimate and mundane interactions → drive the close relationship between the user and assistant.
2. Hidden labour behind ML systems → understanding the amount of labor performed by ghost workers.
3. Creative potential of VAs → transforming the data into performative experiences to reveal what is physically hidden.
4. Surveillance capitalism → the local experiments were connected to the massive corporate infrastructure

Methodologies used:

- ↳ Research-through-design
- ↳ Practice based research in the arts
- ↳ Performance as research

A research team comprised of the 3 authors plus 5 students, and 2 artists/researchers in the field of HCI.

Methodology - Tools

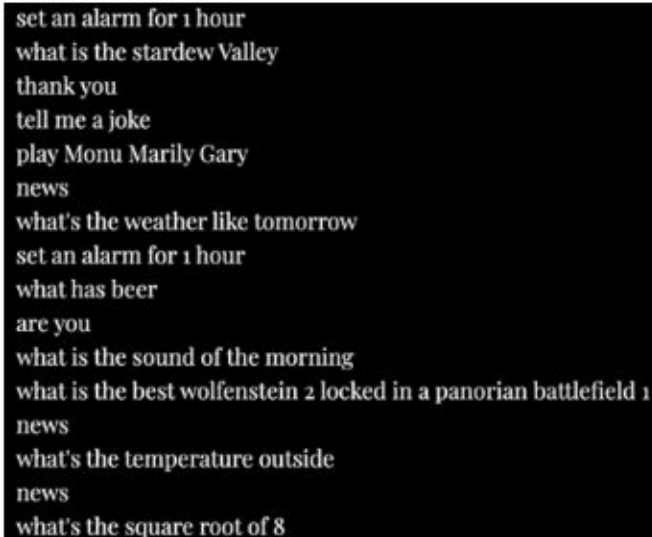
Amazon Mechanical Turk data collection: 26 MTurk workers were paid 10\$ to share their Google Assistant data logs (audio + JSON). Their identity remained anonymous.

Custom GUI: To search and play audio clips from the most evocative entries as well as TTS and to apply audio effects and parameter control via OSC.

RNN: textgenrnn poetry generation model trained on the MTurk logs.

Hacked percussion instruments and wearable voice assistant: Exploring the idea of embodiment/disembodiment, the instruments had actuators triggered by the VAs wake up words.

GIF and Hypertext: Due to Covid-19, the project was presented online. GIFs become functional buttons to navigate the experiments. Hypertext is used to generate narrative.



set an alarm for 1 hour
what is the stardew Valley
thank you
tell me a joke
play Monu Marilyn Gary
news
what's the weather like tomorrow
set an alarm for 1 hour
what has beer
are you
what is the sound of the morning
what is the best wolfenstein 2 locked in a panorian battlefield 1
news
what's the temperature outside
news
what's the square root of 8

Figure 3: Excerpt from the neural net poem Play The Earth

Voices and Voids Vignettes

12 different vignettes (experiments) **reveal what is hidden** in VA systems.

Grouped in 4 tactics:

1. Intentional glitches > Push AI models to their limits to cause VAs to respond via errors.
 - a. The function of these devices is to advertise products. When the assistant would play a random song or book eventually, it would ask to purchase it.
 - b. If not a word is not understood, the system would tend to use the closet one in its advertising catalogue.
2. Emphasizing intimacy and personal narratives > Explore the how VA data can hold and reveal intimate parts of one's life. Data expanding four years of life, voice timbre, accent, background noises, everyday task questions to questions about religion, love and sex.
3. Translations through embodiment > Explores the idea of agency and musical recontextualization of the VA data utterances.
4. Improvisation and juxtaposition: embodiment through instrument playing and choreography.

ORIGINAL	ALEXA VOICEPAD/TEXT TO VOICE
rakete bee bee? rakete bee zee bee zee. Fümms bö wö tää zää Uu, Uu zee tee wee bee fümms.	look at the baby? rocket bezos saying saying. whole foods bowl full the zoo ETV be funs.
rakete rinnzekete rakete rinnzekete rakete rinnzekete rakete rinnzekete rakete rinnzekete rakete rinnzekete rakete rinnzekete rakete rinnzekete	racketeering ditty racketeering details racketeering dirty rocket hitting details racketeering ditty racketeering details racketeering dirty rocket hitting details

Figure 2: Comparison between excerpts from the original UrSonate text, vs the Alexa Voicepad translation



Figure 1: Wearable voice assistant system

Discussion

This paper highlights:

- the intimate and mundane relationships between VAs and their users.
- The hidden labor in AI systems
- The creative potential of VA data
- Issues of surveillance capitalism.

These topics cannot be explored individually, they overlap as part of a complex system.

These matters create an unsettling rhythm that makes it impossible to just “use” a voice assistant or to just “have one at home”.

Multidisciplinary practices expose limitations of the system that reveal the above mentioned topics.

Working with the MTurk system was ethically challenging. How to acquire data and compensate the workers fairly. The identity of the workers was protected.

Conclusion

The experiments reveal the materiality of ML systemes:

- Hidden human labour.
- Vast processing power.
- Large data centers.
- Repetitive and mundane nature of our interaction with VA.

Through these practices of using data as a creative material:

- The VA becomes an active agent
- New human and non-human relationships are created
- The VA meaning of an assistant devices is change to cultural artifact
- Technological solutionism is challenged