Interview Protocol

Introduction

1. Can you tell me a little bit about the way you use your Google Assistant?
2. Can you walk me through your data donation experience?
3. Did you consider removing points from your donation? Why (not)?

    Ask a subset of the following questions:
    A. What feelings or emotions did you have through this experience?
    B. What was the biggest challenge for you?

Introduce Interview Structure

- Introduce the two main questions:
1. How acceptable is it for you to share [attribute] with [recipient]?
   Introduce subject (i.e., user), sender (i.e., Google Assistant prompted by user), and transmission principle (i.e., user is informed and notified).

2. How sensitive do you consider this [attribute] to be? We instructed donors to answer with a 5-point Likert scale from completely sensitive to completely not sensitive.
- Introduce the two stages of interview.

Single Interactions

Introduce the first half of the personal data canvas and the two main questions, here attribute corresponds to a single speech record.

Personal Data Canvas: Single Interactions

The first half of the personal data canvas (Fig. 1) corresponds to an interactive window with four attributes or examples of speech records (i.e., single interactions with a voice assistant) extracted from each interviewee's dataset. Each example included the transcript (as recorded by Google Assistant), the timestamp in human-readable format, and a clickable button to listen to the recorded audio. Each example belongs to a category: (1) training interaction, (2) playful interaction, (3) (common) request interaction, and (4) unintended interaction. Fig. 1 includes a description of the categories in red for clarity, but these were not displayed nor introduced during the interview.

Transcript: Trained voice model
Timestamp: Day, Month, Year @ Hour
(Intended: training)

Transcript: Hey Google, tell me a joke in Dutch
Timestamp: Day, Month, Year @ Hour
(Intended: playful)

Transcript: Hey Google, put on the sex lights
Timestamp: Day, Month, Year @ Hour
(Intended: common request)

Transcript: OK Google, what is the yard clendening
Timestamp: Day, Month, Year @ Hour
(unintended: background conversation)

Acceptability and Sensitivity Score

-2 -1   0   1   2
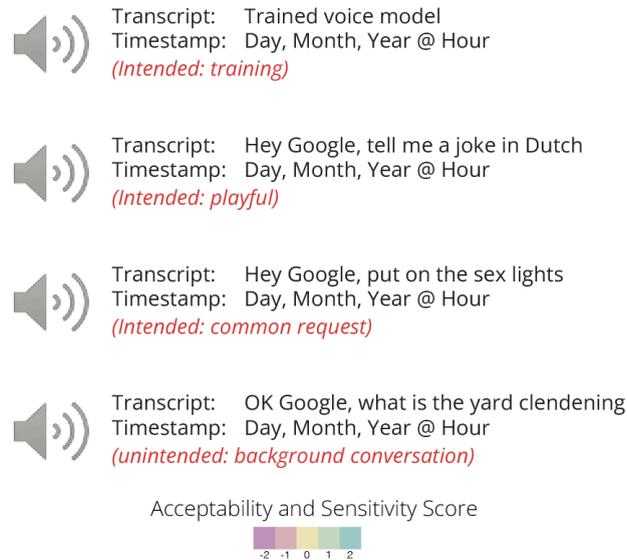
Fig 1. Personal Data Canvas: Single Interactions

For each attribute (i.e., speech record), reproduce the audio recording and read out loud the transcript and timestamp. Then invite donors to react to it and share their first impressions. Finally, ask the two main questions and follow-up questions.

Main Questions

1. How acceptable it is for you to share [attribute] with [recipient]? Instruct donor to answer with a 5-point Likert scale from completely acceptable to completely unacceptable.
2. How sensitive do you consider this [attribute] to be? Instruct donor to answer with a 5- point Likert scale from completely sensitive to completely not sensitive.

   Ask a subset of the following questions:

   A. Why do you consider this interaction (completely (not)) sensitive?
   B. What are some of the factors that led you to consider this interaction (completely (not)) sensitive?
   C. Can you describe what you think makes an interaction sensitive?
   D. Can you describe what you think makes an interaction not sensitive?
   E. How is this interaction different from the previous one for you?

   If intimacy was mentioned by the interviewee, ask a subset of the following questions:

   A. Can you elaborate on what you mean by intimate (or intimacy)?
   B. Can you describe what you think makes an interaction intimate?
   C. What are some of the factors that led you to consider this interaction intimate?
   D. How would you know an interaction is intimate?
   E. How is this interaction different from others (not intimate)?

Beyond Data Transactions: A Framework for Meaningfully Informed Data Donation

Recipients included: partner(s), family, friends, colleagues, and researchers. For each attribute, ask question one five times, one for each recipient, and question two once.

Aggregation of Multiple Interactions

Introduce the second half of the personal data canvas and the two main questions, here attribute corresponds to the information (i.e., patterns or inferences) derived from the aggregation of multiple speech records.

Personal Data Canvas: Multiple Interactions

The second half of the personal data canvas (Fig. 2) corresponds to an interactive window with multiple data visualizations. First, a bar graph with the distribution of requests across different categories (e.g., weather, music, time). Second, a heatmap with the number of interactions per category per hour of the day and day of the week. Third, a heatmap with the number of interactions per category per hour of the day during the weekdays and weekends. Fourth, a word cloud with the most frequent words.
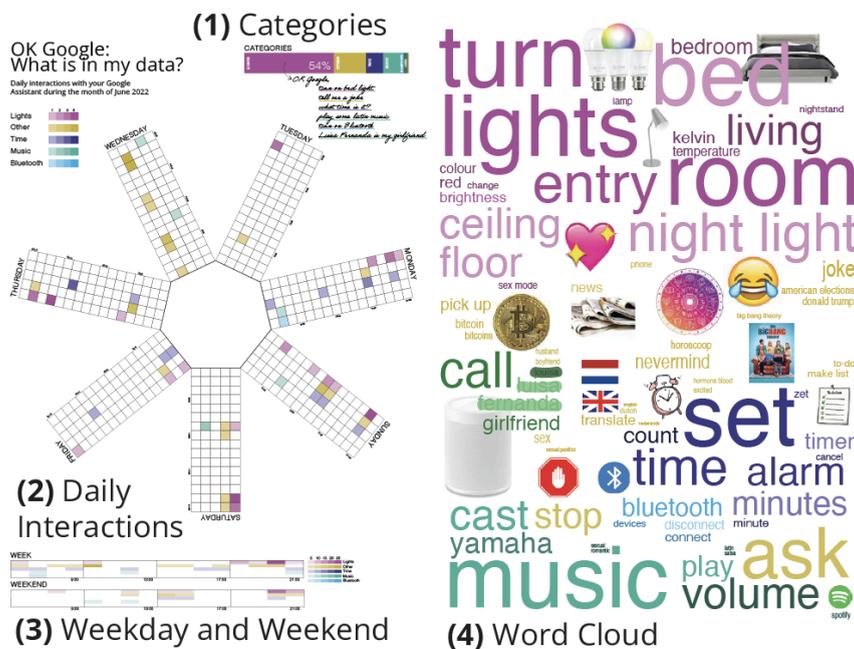
Fig 2. Personal Data Canvas: Multiple Interactions

Introduce the second half of the personal data canvas. Invite donors to explore and interpret it. Invite donors to identify their own patterns in time (e.g., on Wednesdays I'm not home) or sequence (e.g., I wake up and open an app), and inferences (e.g., you can tell I like [artist]). Used patterns and inferences identified by donors as attributes (i.e., information derived from the aggregation of multiple speech records). For each attribute, ask the two main questions and follow-up questions.

Main Questions

1. How acceptable it is for you to share [attribute] with [recipient]? Instruct donor to answer with a 5-point Likert scale from completely acceptable to completely unacceptable.
2. How sensitive do you consider this [attribute] to be? Instruct donor to answer with a 5-point Likert scale from completely sensitive to completely not sensitive.

   Ask a subset of the following questions:

   A. Why do you consider this information (completely (not)) sensitive?
   B. What are some of the factors that led you to consider this information (completely (not)) sensitive?
   C. Can you describe what you think makes information sensitive?
   D. Can you describe what you think makes information not sensitive?

   If intimacy was mentioned by the interviewee, *ask a subset of the following questions:*

   A. Can you elaborate on what you mean by intimate (or intimacy)?
   B. Can you describe what you think makes this information intimate?
   C. What are some of the factors that led you to consider this information intimate?
   D. How would you know information is intimate?
   E. How is this information different from others (not intimate)?

Recipients included: partner(s), family, friends, colleagues, and researchers. For each attribute, ask question one five times, one for each recipient, and question two once.

Data Donation Experience

1. What feelings or emotions do you have regarding your data donation experience, including this session?
2. Are there things you liked about your data donation experience? If so, what?
3. Are there things you learned from your data donation experience? If so, what?
4. Do you feel you gained something from participating in this process? Why (not)?
5. What would you like from a data donation experience in the future?
6. Do you have any questions for me?