

1 Question

How do use patterns vary over time in human interactions with conversational AIs? Does this pattern vary between transparent interfaces and black box interfaces for these systems?

2 Dataset

This project will utilize two interaction datasets from third party creators of conversational AI products- one product with a transparent interface (e.g. Amazon's Alexa), and one with a black box interface (e.g. Apple's Siri). If only one dataset is obtainable, the second part of the question may be omitted.

3 Variables

The analysis for this project includes three variables:

1. Type: is the interface transparent or black box? This will be consistent within each dataset.
2. Frequency: how many interactions occur per user per time interval?
3. Duration: how long is each interaction? This would ideally be measured as a number of discrete tasks per interaction; however, depending on what the dataset includes, this may or may not be feasible.

4 Why Observational Data?

Effective human-computer interaction and user experience research rely heavily on observational data; one of their core purposes is to identify *how* people actually use or interact with systems without influencing their behavior. This often utilizes a modified ethnographic framework (cognitive ethnography) which comes with a strict protocol on observation without influence, which is in line with how Salganik defines observational data.

In *Bit by Bit*, Salganik further categorizes big data as observational, and identifies two sub-types: government administrative records, and business administrative records. This project utilizes the digital traces that result from business administrative records. This is appropriate, as I wish to objectively identify patterns in how people use a certain type of technology without influencing their use and because this technology is owned and productized by private firms.

5 Why Big Data?: Positive and Negative Characteristics

■ Big

Salganik identifies three ways in which big data can be big: "many people, lots of information per person, or many observations over time". This dataset qualifies as big based on all three of these parameters: it contains information from many people, lots of information per person, and many observations over time.

■ Always On

This attribute of big data is particularly useful because it enables the study of unexpected events, as well as real time measurement, according to Salganik. For this experiment, the unexpected events part of this attribute is salient: it is possible that there are breaks in the data, or that its patterns are not consistent over time, particularly given that this data is based on human behavior. This data is particularly always on- products that utilize this technology are often always "listening" for a wake word to activate upon.

■ Non-Reactive

Effective interaction research is founded on not influencing the user, which can be incredibly difficult to achieve. Thus, this, given that “measurement is much less likely to change behavior in big data sources” attribute makes big data especially suitable for this kind of research [Salganik]. This dataset is non-reactive because it is comprised of data on things that have already happened “in the wild”.

■ Incomplete

This dataset is incomplete; it includes only what a third party or third parties are willing to share.. Because this particular project does not attempt to infer causality in differences of patterns over time, or segment users into specialized populations, operationalizing this data is straightforward: the only aspect contingent upon construct validity is the interface type, which is easy to categorize according to existing heuristics.

■ Inaccessible

Any research using this data is predicated on a relationship with a third party, and their willingness to share data.

■ Not Representative

The users of these technologies are not representative of society as a whole, nor are they a random sample; rather they likely represent a relatively affluent demographic that skews toward technology early adopters. As Salganik mentions, it is important to think about the data that we don't have- different populations could use these technologies very differently.

■ Drift

All three of Salganik's characterizations of drift: population, behavioral, and system could become issues with this dataset. As new technologies catch on, more people adopt them, and the population drifts. As people use once novel technologies for longer, they learn behaviors, and their behavior drifts. As companies iterate on their designs and algorithms, the system itself drifts. On a positive note, these technologies still require an update with release notes. Thus, it should be relatively easy to cross reference this information (i.e., when the system changes) to mitigate issues caused by system drift.

■ Sensitive

Some instances of these technologies are always “listening”. This is useful, but also makes it likely that they will capture very sensitive information. If this dataset contains interaction content, some of it would likely be potentially embarrassing to the person who generated it. However, it seems highly unlikely that content data would be made public, so it possible to circumvent issues caused by sensitive information.

6 Feasibility Assessment

This project is tenable as a hypothetical thesis project. Data collection seems the biggest potential concern. However, I am working on a (still general) project on augmented intelligence and conversational AI with James Evans, which will likely result in obtaining a related, usable dataset. I think it would also be possible to obtain a similar dataset by leveraging connections with the professors I know from UCSD.

Based on the University of Chicago IRB guidelines, this project shouldn't require IRB approval; the only way that information would be personally identifiable would be in the event of an anonymization failure. Interaction data would presumably be primarily structural and coded with anonymized user IDs, but it could potentially contain locations or content of individual interactions, which would make it sensitive. The easiest way to address concerns around this is to not publish the dataset, which would likely already be a stipulation of the third party who provided the data.

7 Sources

By Bit: Social Research in the Digital Age. "Bit By Bit: Social Research in the Digital Age." Bit By Bit: Social Research in the Digital Age. Accessed October 24, 2016. <http://www.bitbybitbook.com/>.

"Social & Behavioral Sciences Institutional Review Board." Does My Research Need IRB Review? Accessed October 24, 2016. <https://sbsirb.uchicago.edu/page/does-my-research-need-irb-review-0>.