

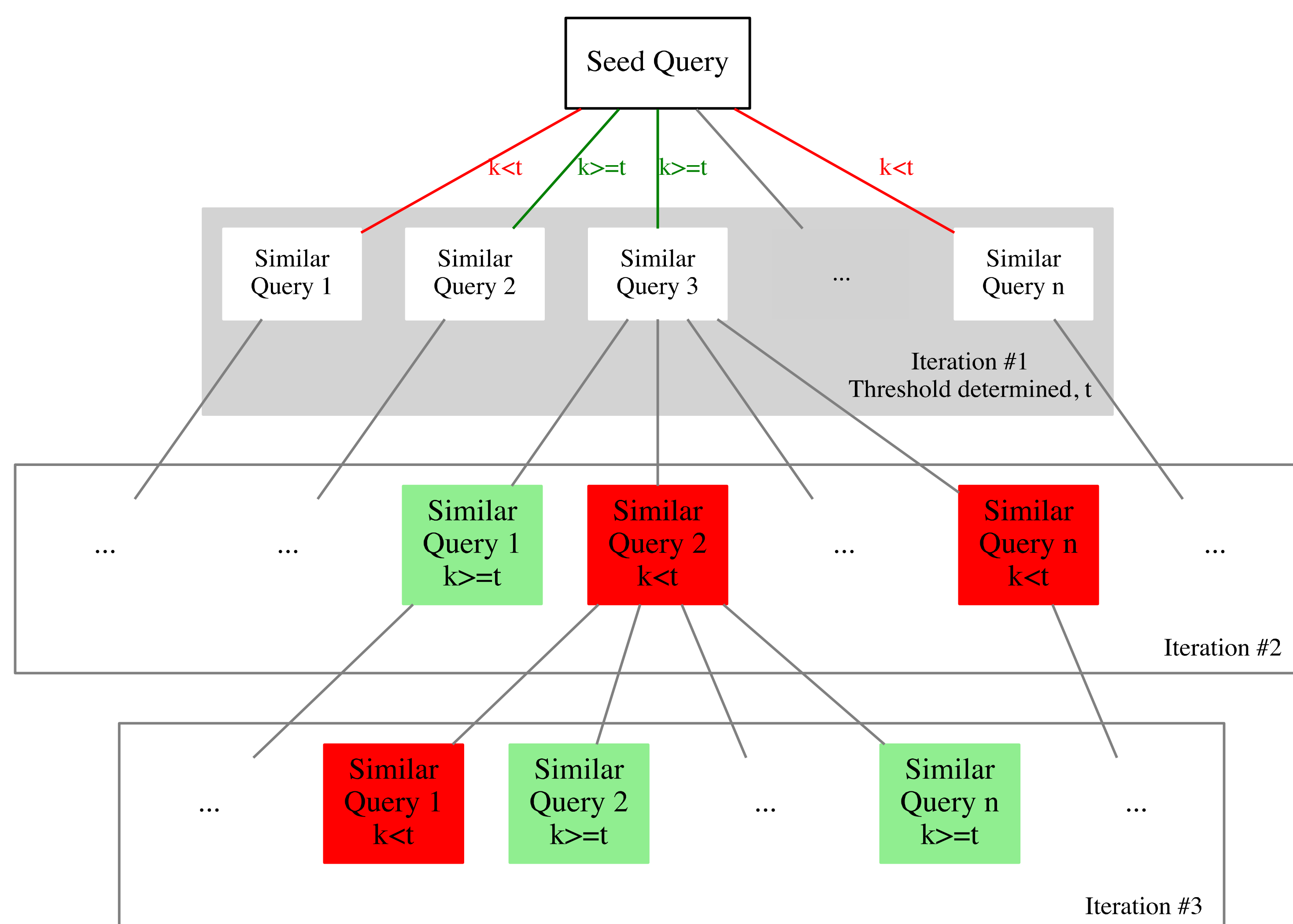
Leveraging Search Query Trends as a Privacy Preference Proxy

Motivation

- In 2015, Pew Research Center published a report discussing the change in online activities and habits of US adults since the Edward Snowden revelations.
- We want to verify this change in privacy-related behavior due to major privacy incidents like the Snowden revelations through online search queries.

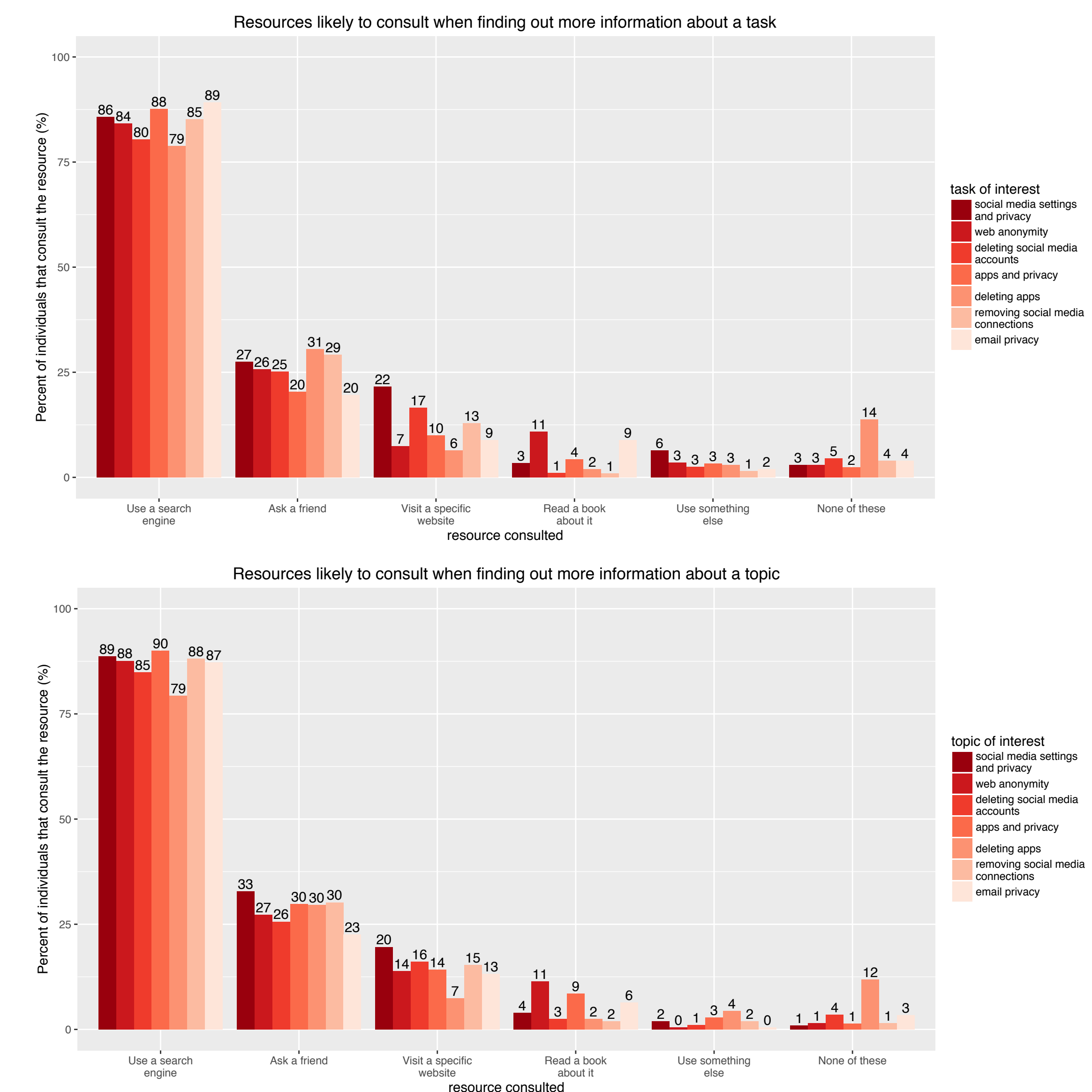
Methodology

- We constructed a set of representative search queries from Google covering the major categories of privacy-related behavior in the Pew report.
- Fetch Google Trends time-series data for the search query set and compare the distribution of the trends data using Earth Mover's Distance and/or Kullback-Leibler divergence.



Methodology (cont.)

- Use Amazon Mechanical Turk (AMT) to determine the likelihood of a user consulting a specific resource for a privacy-related task or topic.
- Collect sample queries from AMT to test the recall of the representative query set.



Limitations

- Google's suggested query similarity can be ambiguous.
- Google Trends data does not quantify query volume.
- The representative set contains queries that often do not have any trend data due to them being long-tailed.