

Research Questions

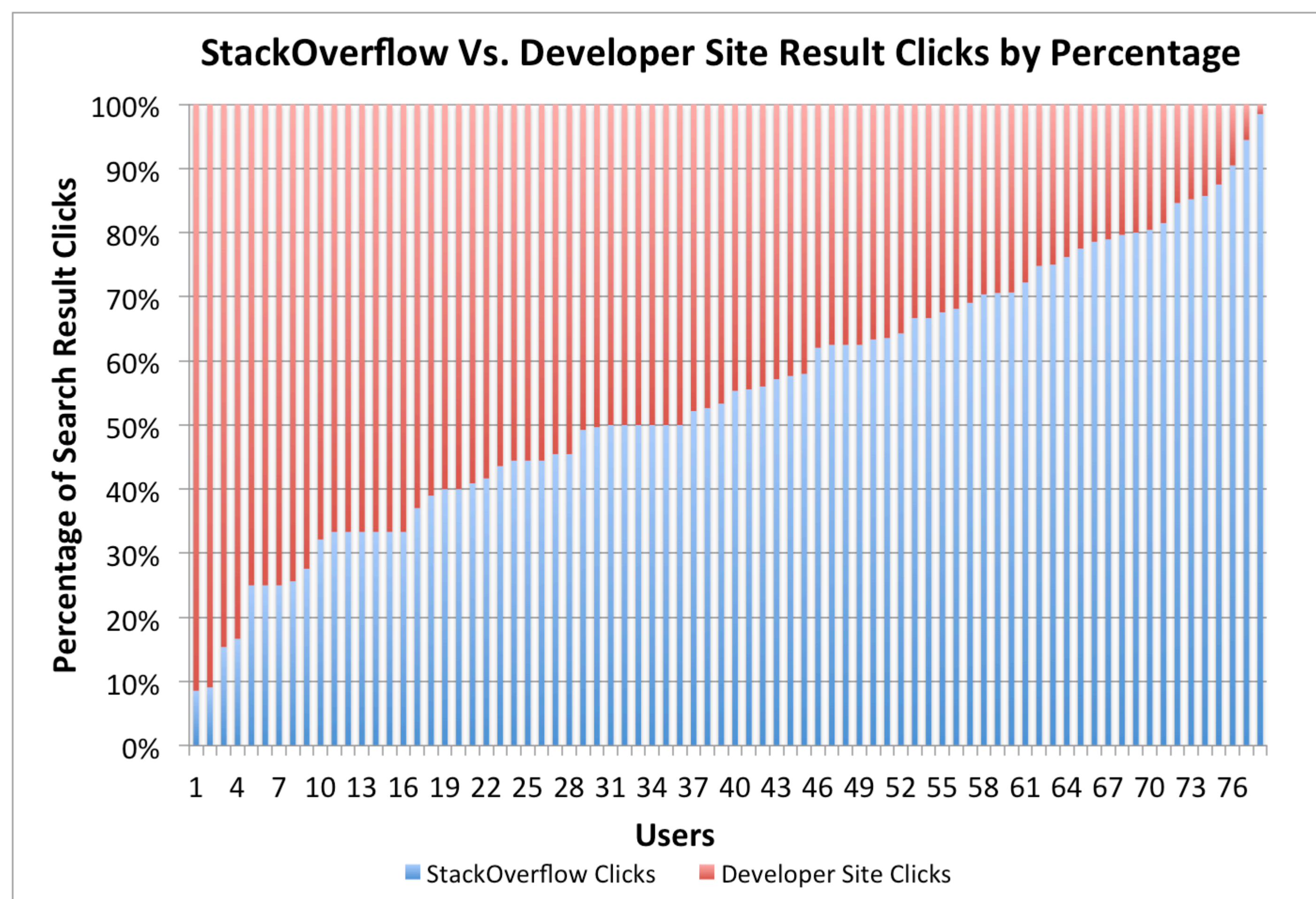
- **RQ1:** How prominent are crowd-based documentation websites (e.g., StackOverflow) in developer searches?
- **RQ2:** What sources of information are viewed in lieu of crowd-based documentation?
- **RQ3:** What are some difficulties associated with technical search queries?

Study #1: Observing Developers in the Wild



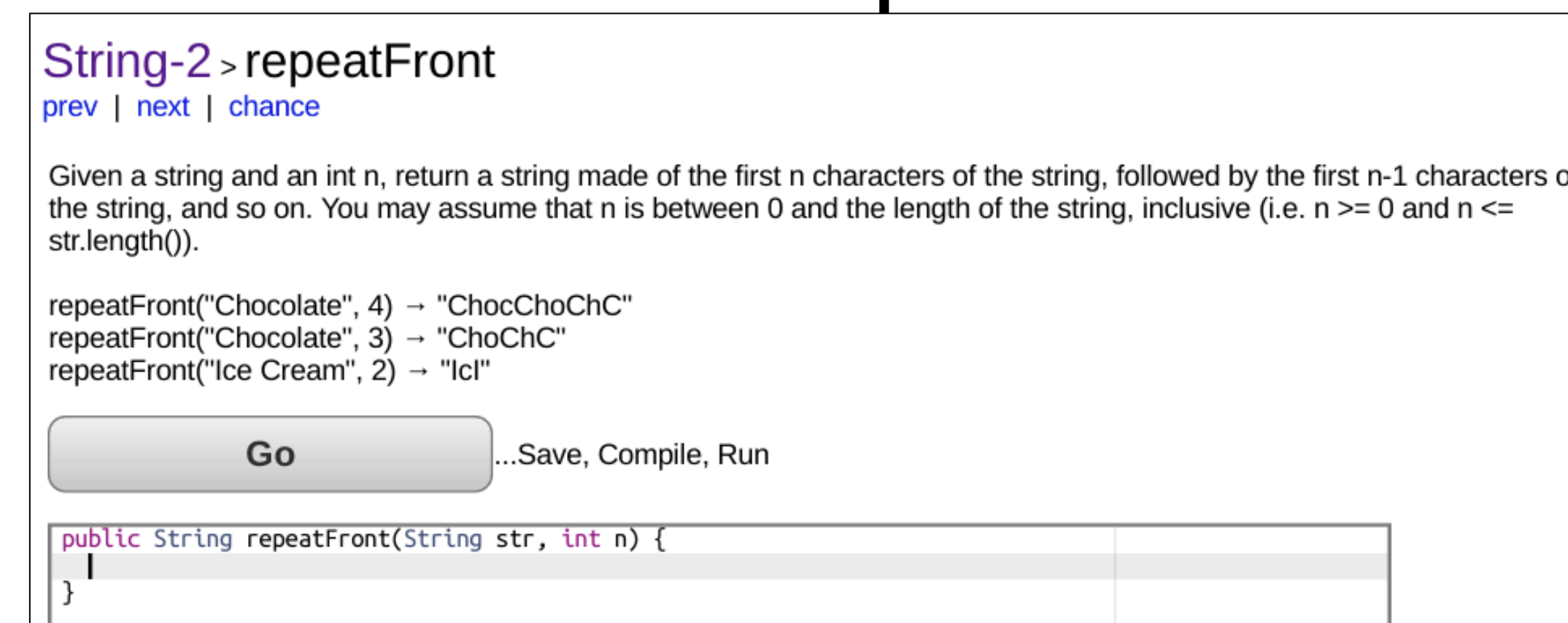
Studying **79** developers that visited **4,180** developer related websites over **5** months, we found that on a per user-basis::

- Median number of StackOverflow result clicks: **10**
- Average number of StackOverflow result clicks: **75**
- Median number of other developer result clicks: **8**
- Average number of other developer result clicks: **76**



Data obtained from Codealike: <https://codealike.com/>

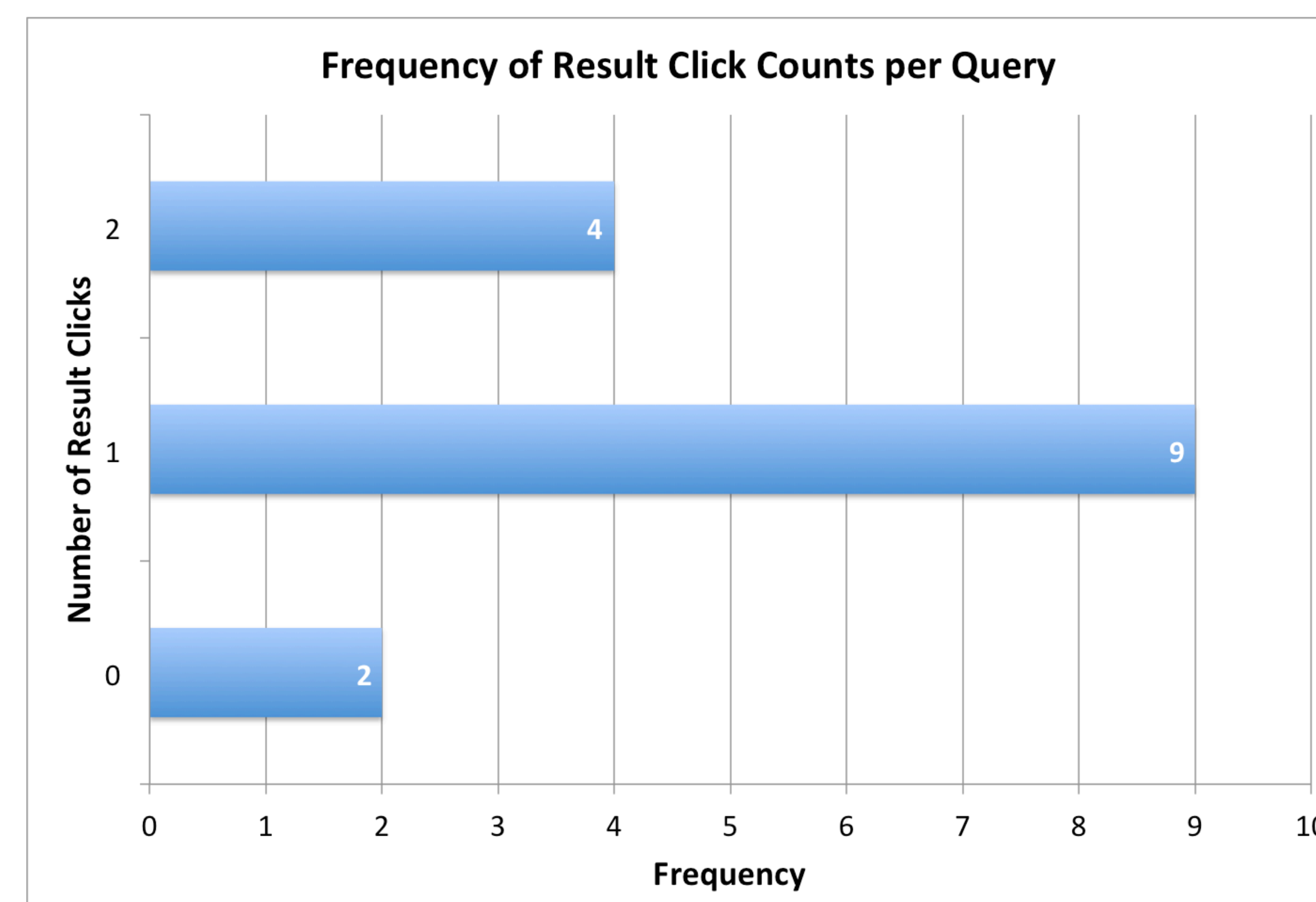
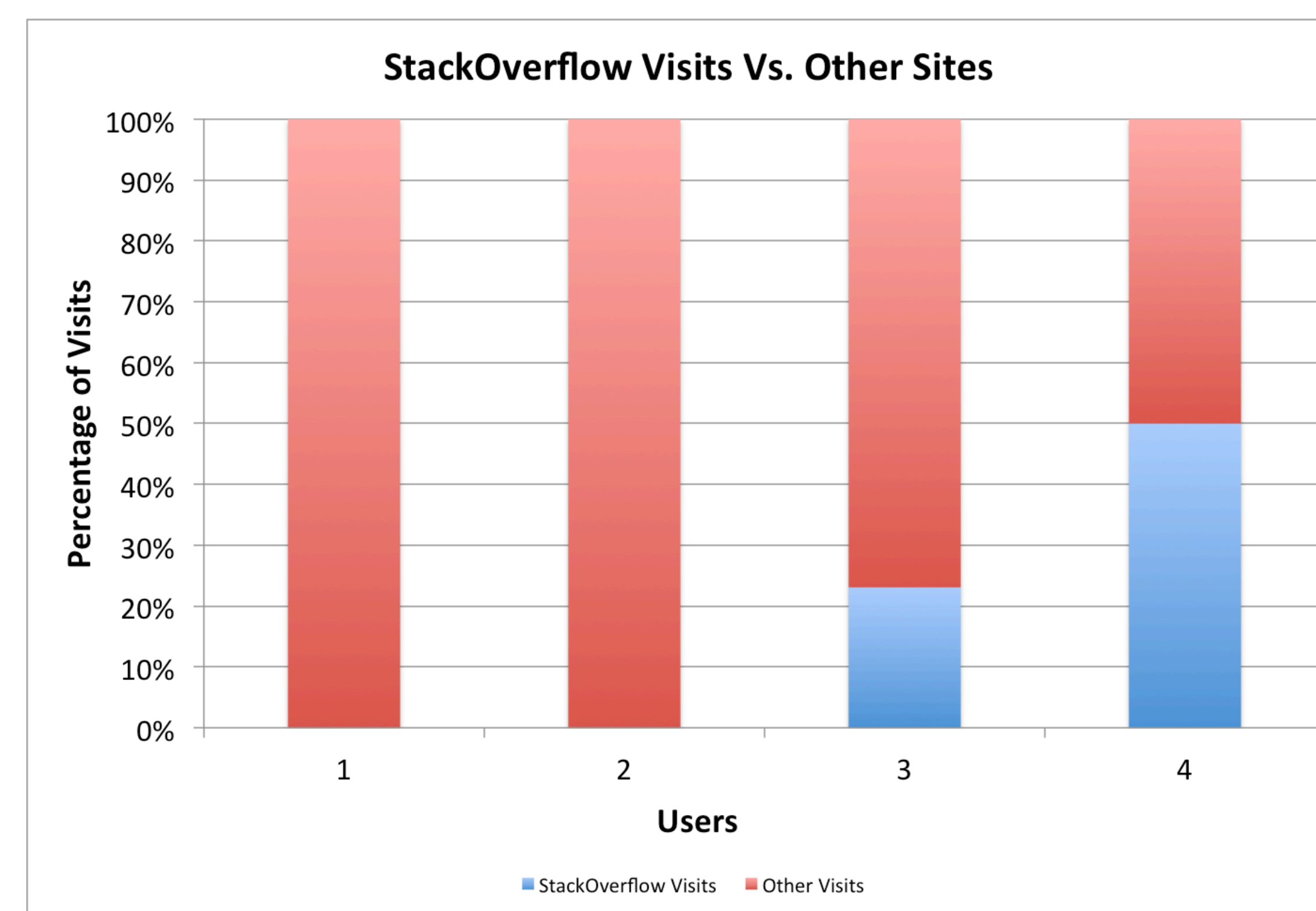
Study #2: Investigating Search Behavior, In-Depth



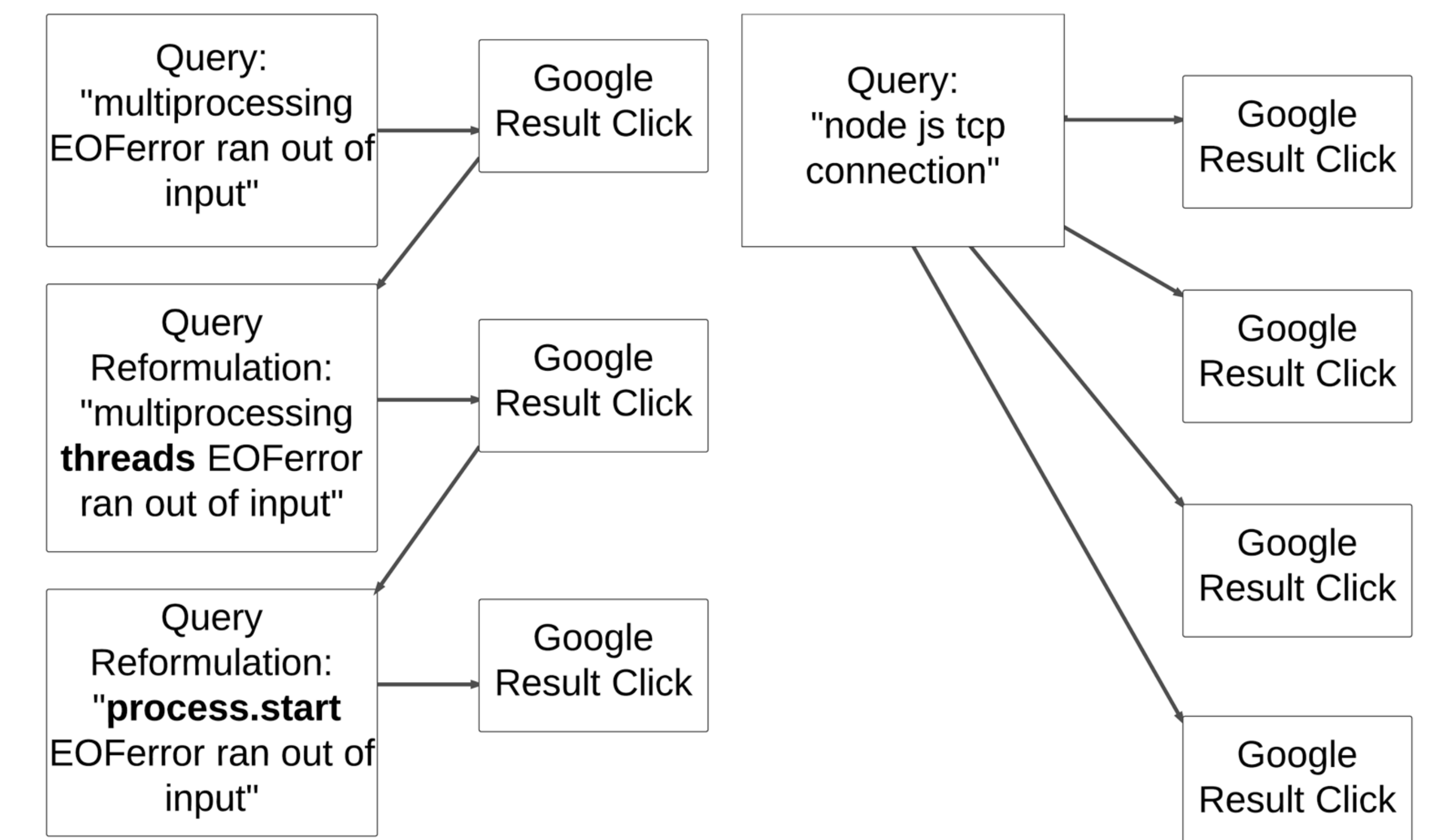
Study Setup:

Users completed three Java programming tasks, similar in difficulty and scope. These include manipulation of strings, arrays, and ArrayLists. Test cases, which were exposed to participants, were used to determine when each task was complete. Participants were permitted to perform as many web searches as needed.

The four participants were undergraduate students; each participant completed all the tasks, and the average completion time for all tasks was **19** minutes.



Which Search Pattern fits you?



“Zig-Zag” searching

The user makes a series of queries, clicking on one search result each time. Frequent backtracking, possibly with a revised query, is observed.

“Fan-Out” searching

The user makes one query, but opens multiple search results simultaneously. Minimal backtracking to search page is observed.

In Study #2, all participants followed the Zig-Zag pattern!

Future Work

Study #1 Future Work:

- Investigate the data collection mechanism used by Codealike to improve tracing user workflows.
- Create a more robust method of categorizing developer related websites and search queries.
- Look for common search reformulation patterns among developer queries.

Study #2: Future Work:

- Study larger population with greater variation in experience level.
- Explore the types of questions for which crowd documentation is best suited.
- Compare these results with more open-ended tasks.
- Find out if the “other sites” category is dominated by a different class of developer-related websites.