

1 Project Proposal

(a) Project title?

“Search Engine for Children”

(b) What is it going to do?

Based on previous research, it has been found that children prefer to use a popular search engine, such as Google Search, over one specifically built for them and do not gain much in using a search engine built specifically for children. However, children have difficulties in using popular search engines including: formulating queries, interpreting query suggestions, and analyzing returned results. For these reasons this website will have three main parts. The first is to interpret a query written by a child, that may be a longer natural language query instead of a keyword query, and produce query suggestions. The second part is to offer query suggestions in a way that is easy for a child to interpret, or to simply bypass a user’s choice and select a query suggestion for them. The third and final part is to analyze the returned results, remove links that are inappropriate (including sexually explicit, content that children would not be interested in, content that is not at a child’s reading level), and re-rank results to bring the most child-relevant links to the top. The UI should still look like Google Search and may even use Google search API on the back end.

(c) Who is the target audience?

Children, Teachers/Educators, Parents of young children

(d) What sort of data will it manage?

The main data would be text to process, such as queries, as well as links and logging data. Other data may include an index, tf-idf or cosine similarity scores, or links/websites crawled. May also create surveys or a database of benchmark data for research purposes that could be stored/processed through the website.

(e) What web technologies will it use?

PHP Laravel for the framework. Apache Lucene, Solr, and/or Hunch for indexing and crawling websites. Google Search API. MySQL database and/or Elasticsearch for a few things?

(f) Stretch Goals?

Develop own indexing and web crawling instead of using Google Search. Both would be good for comparison for research purposes. I also want to include logging, built in surveys, and to create a database for storing benchmark data.

2 Category Requirements

(a) Is it a website?

Yes. I would like to gain experience with AWS so I might try and host there.

(b) Is it a developer tool, library, or framework?

No

(c) Will it be proprietary to your current employer?

No

(d) Is it a browser plugin?

No

(e) Is it a game?

No

(f) Is it an API?

No