

# Info Retrieval: Project Proposal

Raj Patel  
University of Utah  
raj.patel@utah.edu

Blaze Kotsenburg  
University of Utah  
bkotsenburg@gmail.com

Brandon Ward  
University of Utah  
bmanmo14@gmail.com

## ABSTRACT

Our project proposal is an exploratory project in hopes of developing a search engine based on what we have learned in class. We propose to develop a search engine for a dataset that we found for Wikipedia webpage documents. In this exploratory project, we plan to include methods for ranking algorithms, clustering of web docs after a user query, developing a user interface (UI), and using several evaluation metrics to compare relevance of the queries.

## KEYWORDS

search engine

### ACM Reference Format:

Raj Patel, Blaze Kotsenburg, and Brandon Ward. 2020. Info Retrieval: Project Proposal. In *Proceedings of ACM Conference (Conference'17)*. ACM, New York, NY, USA, 1 page. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

## 1 INTRODUCTION

During the course for Introduction to Information Retrieval (CS 6550), we have covered a multitude of topics and methods regarding current standards for information retrieval (IR). The topics covered provide a large amount of possibilities for a class project.

Our group's common interest is to develop a simple, yet efficient, search engine that contains a healthy amount of topics

and methods covered in the CS 6550 course. This project provides us an opportunity to apply what we have learned during class and assignments to develop a search engine tool. We plan to put our own twist on the search engine by including a post-query clustering for documents containing similar topics.

The proposed search engine project also gives us a great opportunity to combine our back-end learnings from the course with a front-end UI. We believe that this will be a great challenge as simplicity in UI design remains a crucial component for any search engine. We believe that displaying the clustered documents in a simple UI will be the most challenging part for the front-end.

Since this project is exploratory and its purpose is to help us further our knowledge and understanding of IR, our goal is to make sure that our search engine is functional, user friendly, and as efficient as possible based on our current knowledge of IR.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*Conference'17, July 2017, Washington, DC, USA*

© 2020 Association for Computing Machinery.  
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00  
<https://doi.org/10.1145/nnnnnnn.nnnnnnn>