

Academic Search Engine (ACASEEK)

Dr. L. H. Patil, Prajakta Joshi, Rutika Sonak, Surabhi Parthabe

Department of Computer Science, Rashtrasant Tukadoji Maharaj University, Nagpur, Maharashtra, India-440027

Abstract— In this paper, we present information about search engine which will work on academic level. Accessing and collecting exact subject wise notes is a complex task, therefore it is necessary to have a correct access to notes when they are needed. It becomes difficult for the teachers to provide the soft copies of the various notes,* either they have to mail each and every student or the students have to come to the teachers asking for notes. While doing any of this, if the student or the teacher gets lenient it may become an issue of irresponsibility. So, to reduce the burden of providing and getting the notes by teachers and students can take help of this Academic Search Engine. This problem of import export and management of notes is solved by this project.

This project is about providing a platform for teachers and students, where teachers can upload the important notes and students can have an easy access to them. In this search engine teachers can have their own profile where they can upload the notes they want to provide to the students. These notes can be further accessed by the students by searching on the home page of the search engine. Not only the notes uploaded by the teachers can be accessed by the students but also students can have an access to any other information they need. The information uploaded by the teachers can be in any form like in form of pdf, videos or images. Using this search engine shall prove useful everyone being in any field, this will help all the students and teachers in managing their work.

Keywords— Academic, subject wise notes, profile, upload, pdf, videos, and images.

I. INTRODUCTION

The Academic Search Engine makes the notes and other important things related to the particular subject easily available to the students. To make them available to the students at any point of time they want and at any place they want.

To reduce the work load of teachers of providing the notes to students and worrying to ensure have all the students have got the notes. Even Google provides all information about the subjects but we aspire make a search engine which provides a concentrated set of information, clearly stressing on the appropriate information needed for a subject.

Academic Search engine will be a great benefit to all the students as all the essentials related to subject will be easily available to them whenever and wherever they want.

It will be a boon to the students who are unable to attend the college regularly and never have notes at the time of exams. They will not have to worry as the notes will be always available to them on the search engine. It will be provide a great advantage to teachers as they always stay worried about whether all the students have the notes or not till the college exams. Teachers will feel free of their responsibility once they upload the notes on their profile. This academic Search Engine can be used for all types of organizations be it a school, college or office or any other big organization.

Academic Search Engine is a smaller version of a search engine. This search engine will work on departmental or college level providing all the essentials for the particular subjects.

These essentials consist of notes, PPTs, images, PDFs and videos related to the specific subject.

Academic Search Engine will have a home page like any other search engine, it will consist of a search bar in which students can search about the notes and other things related to the subject.

Search results are displayed i.e. all the information related to the search are displayed which is stored in the database. Also the result page will consist of options on top left of the page (like *All, Images, Videos, News, More and Search tools* on top left of Google searched result page). Here, the options will be pdf, notes, images and videos.

On the other side the teachers will be provided with a user id and password, through which they can access their profile and upload the important notes and other things which will be accessible to the students.

Also if any other information is searched by students, a wiki link related to that topic will be provided.

II. LITERATURE REVIEW

[2] Now a days, search engine is the practical application of information retrieval techniques to large-scale text collections. A web search engine is the obvious example, it is also known as an evaluation corpus (plural corpora). The term “search engine” was initially used to refer to specialized hardware for searching text.

Search engines can be used with small collections, such as a few emails and documents on desktop or extremely large collections, such as the entire Web. There may be only a few or many thousands users of a given application.

Search engine support major function, as query process. The query process uses those structures and a person’s query to produce a ranked list of documents. The major components in query processing are user interaction, ranking, and evaluation. The user interaction component provides the interface between the people doing the searching as well as the search engine. There are two important tasks. First task for this component is accepting the user’s query and transforming it into index terms and second task is to take the ranked list of documents from the search engine and organize it into the results shown to the user. [6] The ranking component is the core of the search engine. It takes the transformed query from the user interaction component and generates a ranked list of documents using scores based on a retrieval model.

[4] Document Transformation is the strategy which is simple enough to become part of a large-scale search engine. It involves modification of document representation space so

that documents are brought closer to the queries to which they are relevant.

Direct Hit (www.directhit.com) is the only current search engine that claims to adapt to judgments. The owners of this system say the system learns by monitoring which sites searchers select from the results page, and how much time the searchers spend at these sites. However, the algorithms used by Direct Hit have never been published. It is believed that algorithms for exploiting user histories are too valuable and also the best strategies will only be found if the research community develops an interest in this area. Web search engines are straining against the limitations of speed imposed by current technology [1].

[5] This paper gives an introduction about the problem of storing a large amount of data on the web. Academic Search Engine containing as many fields related to Academics total. These fields allow rapid calculation of a web page's "Page Rank", an objective measure of its citation importance that corresponds well with people's subjective idea of importance. Because of this correspondence, Page Rank is an excellent way to prioritize the results of web keyword searches.

[3] Ranking functions are used so that no particular factor can have too much influence. In order to rank a document with a single word query, Google has to look at the document's hit list for that word and then it considers each hit to be one of several different types (title, anchor, URL, plain text large font, plain text small font ...), each of which has its own type-weight. The type-on their

Google counts the number of hits of each type in its hit list. Then every count is converted into a count-weight, this count-weights increase linearly with counts at first but quickly become less active so that more than a certain count will not help. Then the dot product of the vector of count-weights with the vector of type-weights is taken to compute an IR score for the document. Finally, the IR score is combined with Page Rank to give a final rank to the document.

III. PROPOSED WORK

Search engine like Google cannot be done overnight. Google has a unique NOSQL database called Big Table where they store the entire search data. Due to space and time constraints we will be making a smaller version of a search engine. This search engine will work on departmental or college level providing all the essentials for the particular subjects. These essentials consist of notes, PPTs, images, PDFs and videos related to the subject. Distribution of notes becomes very difficult for faculties, either they have to mail each and every student or the students have to come to the teachers asking for notes. While doing any of this, if the student or the teacher gets lenient it may become an issue of irresponsibility. Our main objective is to reduce the burden of providing and getting the notes by teachers and students can take help of this Academic Search Engine. This problem of import export and management of notes is solved by this project.

This project is about providing a platform for teachers and students, where teachers can upload the important notes and students can have an easy access to them.

The developed search engine will be used by teachers, students and administrator and the utilities provided by this search engine are as follows:

Registration Page: This page provides the facility to the faculties to register an account by making their profile. For the registration, faculties have to provide their first name, last name, e-mail id, user id, password and the subject which they are associated with.

The e-mail id which will be provided by the faculties will be displayed on their profile. This e-mail id will be useful to the students in order to contact with the faculty in case the student has any difficulties or issues regarding the notes provides by the teachers as well as related to the subject.

The user id plays an important role for the faculty registration process. As the user id provided to each faculty will be unique. Similarly, with the password i.e. the password for each faculty will be unique.

After the registration process for the faculty is completed, the faculty can upload the required data i.e. notes, pdf, images, etc. related to the subject. All of this uploaded data will be stored in the database and will be further useful for the searching process.

Home Page: This page will consist of a search bar in which the users will search anything related to the various subjects. This page will be connected to the database and this database will contain all the data which will be uploaded by the faculties in their profile. Also the search engine will be connected to Wikipedia so, if anything which is searched by the user has no results in the database then the results will be provided by Wikipedia. Also on the home page at the top right corner there will be options for faculty registration as well as for faculty login. Also on the page there will be options to filter the users' search like if the user wants only image or videos or theory as the results of the search then their will be options for that on the home page (like it is on the Google home page).

Ranking: Search engine ranking refers to the position at which a particular site appears in the results of a search engine query. Ranking is applied to the keyword entered in the search box.

Search engine will check for pages within their index that are a closest match. A score is assigned to these pages. These pages (or images & videos) will then be displayed to the user in order of score. A site is said to have a high ranking when it appears at or near the top of the list of result.

A search engine keeps an index of all web pages, and when a user types in a query search, the engine browses through its index and counts the occurrences of the key words in each web file. The winners are the pages with the highest number of occurrences of the key words. These get displayed back to the user.

Simple concepts of Ranking

- *Scoring:* The scoring makes up the initial rankings.
- *Boosting:* It is another element or signal that might raise a page's position in the rankings. Various personalization can also use a boosting element to re-rank results.
- *Dampening:* A dampening factor is an element that would lower the rankings of a web page after the initial scoring

process. This may *seem* like a penalty but it is a dampening element.

IV. CONCLUSION

In the existing scenario, the notes distribution process by teachers is done manually i.e. teachers have a great burden of making notes available to the students before the examination. To overcome this problem we provide a platform to teachers through this academic search engine (ACASEEK) where teachers can get relieved by uploading the required notes on their profile provided in academic search engine (ACASEEK). In future, this academic search engine can also have student study groups.

REFERENCES

- [1] Geetanjali Tyagi, Megha Sharma, Kumar Kaushik "Search Engine Optimization Technique Increasing Website Traffic and Online Visibility", January 2015.
- [2] W. Bruce Croft, Donald Metzler, Trevor Strohman "Search Engines Information Retrieval in Practice", published by: Pearson Education, Inc., April 2015
- [3] Dirk Lewandowski "New Perspective on Web Search Engine Research", March 2012.
- [4] Charles Kemp, Kotagiri Ramamohanrao "Long Term Learning for Web Search Engines", December 2002.
- [5] Sergey Brin and Lawrence Page "The Anatomy of a Large-Scale Hyper textual Web Search Engine", November 2000.
- [6] David Hawking "How Things Work- Web Search Engines", January 1999.