

# Architecture for Gujarati Search Engine

Hemang Desai<sup>1</sup>, Dr. Birajkumar V. Patel<sup>2</sup>

<sup>1</sup>Vimal Tormal Poddar BCA College, Surat, Gujarat, India

<sup>2</sup>P.G.Department of Computer Science, S.P. University, Vallabh Vidhyanagar, Gujarat, India

**Abstract:** This paper presents architecture for the Gujarati topic-oriented search engine technique. The proposed work is founded as admin-client architecture. This architecture could organize information from various different abstraction levels by tag of topic of queries and information. The existing search engine available for the English language, the proposed architecture is the block diagram of how topic-oriented search engine would look like and also introduced the functionalities and responsibilities of each modules of the architecture.

**Keywords:** Gujarati Search Engine, Topic-oriented Search Engine, Architecture, Informative Search Engine, Topic-oriented Framework.

## I. INTRODUCTION

Searching on the internet has become widely use for the people in society to get the information in the right manner [1]. The search engine has become as the platform for the searching. The optimization of these searching is become essential as the web is a large collection of information [2]. Here, introduced the new type of search engine that produces the all types of information to the user as in the form of images, blogs, documents, videos, and much more. The new basic idea behind proposed research work is to generate the architecture and optimization tool that can search in Gujarati language.

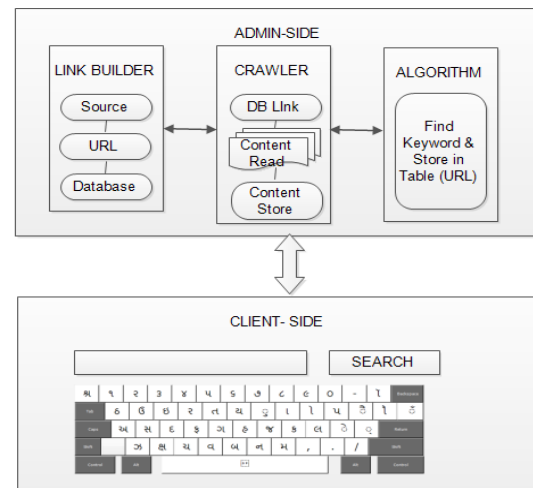
In this research paper, proposed Gujarati topic-oriented search engine. Instead of searching and display the information in English language. Our framework support the searching as well as results in Gujarati language.

## II. METHODOLOGY

### A. The Admin-side Architecture

The whole work can be demonstrated in block diagram as Admin- Client Architecture. Where the admin side is classified into three

major modules: a) Link Builder, b) Crawler and c) Algorithm as shown in Figure 1.



**Figure 1: Architecture of Admin-Client**

1) Link Builder module: The task of Link builder module is to fetch information from the web where lots of web content is available. In proposed architecture, one can accomplish the task by developing the Link Builder, which can work on reading the URL and fetching the information from different web pages. The fetched URL are automatically store in the database, where the

duplicate entries removed from the records, then the data needed for the model could be automatically fetch to present final results.

2) Crawler Module: A web search engine is a kind of search tool whose task is to search for hit the keywords and fetch the relative information from the World Wide Web [3]. The retrieved information is in the form of web pages, which they retrieve from the HTML mark-up of the pages. These pages are fetched by Web crawler (also known as a spider) — an automated Web crawler considers every link on the site. The crawler stores every linked URL in the database and also read the content and Meta tag. The main functionality of web crawler is to fetch and store the content links and read the contents from relative URLs.

3) Algorithm: Finally, at the scope of algorithm, the query from the user will be converted into keywords. The task of algorithm is to find that keyword and store in the table. The place of those keywords would be in the table as of where the Link builder and crawler have already stored the URLs.

#### **B. The Client Architecture**

The user can input search text to get the required content in the text box. Proposed architecture provide the Gujarati keyboard to the user for the easy typing and can search and hit the keywords in Gujarati language.

### **III. CONCLUSION AND FUTURE RESEARCH**

In this paper, we proposed Gujarati topic-oriented search engine which takes input as of Gujarati text from the user also explained how those text act from different abstraction level at the admin site. We also discovered the architecture for working of our proposed work and found the admin-side and client-side working of our proposed work. Our Future work is the practical implementation of proposed modules, also will work on how the searching and its result take place at all different place depending on the requirement.

#### **REFERENCES**

- [1] Berman, Ron, and Zsolt Katona. "The role of search engine optimization in search marketing." *Marketing Science* 32.4, 644-651, 2013.
- [2] Dover, Danny, and Erik Dafforn. *Search engine optimization (SEO) secrets*. Wiley publishing, 2011.
- [3] Beel, Jöran, Bela Gipp, and Erik Wilde. "Academic Search Engine Optimization (aseo) Optimizing Scholarly Literature for Google Scholar & Co." *Journal of scholarly publishing* 41.2, 2009.