TRIBHUWAN UNIVERSTIY

INSTITUTE OF ENGINEERING

Pulchowk, Lalitpur.

Project Proposal on

Knowledge Graph, A search engine

Submitted by:
Kishan K.C.
Roll No : 067BCT519

Submitted To:
Department of Electronics and Computer Engineering

Abstract

The main target of this project is to develop search engine to search the details of the keywords given by the users. Anyone can use this system to search about the data. The system is provided with the number of data as database in form of files. The system will then ask for the two inputs. The system analyses the relation between the two words matching their keywords with that of database. The system takes two input, then searches the keyword in input into the database, then determines the relation between the inputs on the basis of matching keywords.

Table of Contents

Introduction

We need to search for the various kinds of data over the internet. The searching facility is provided by the search engine. This search engine understand the words that we want to enter. These words describe the real world objects. This system helps in collecting information about objects in real world. Objects could be a person, a book, movie and many more other types. For Example. For a famous person, we collect relevant data like date of birth, Appearances, height, etc. The knowledge graph connects that person to closely related other objects. Knowledge graph relate the data with closely related other data with reference to various things. Search engine looks for the word you typed in, but sometimes also looks for synonyms or related terms.

Literature Review

Searching s been the great subject of interest. Nowadays, Search engine are that much highly functioning that it searches our query as the objects not just as the word. Google has been using the concept of knowledge graph. The relative way of representing the relation has been the major advantages to perform these kinds of searching.

Google Search

Figure 1: Google Knowledge Graph

Google Search starts with the web. It's made up of over individual pages and it's constantly growing. Google navigates the web by crawling. It means google follows links from page to

page. Google sorts the pages by their contents and their factors and keep the track of the data in the index. When we search for anything in the google, it starts to look for the clues to better understand what we mean. Based on these clues, google pulls up the relevant documents from the index. Then google ranks those document on the basis of over 200 factors. All this happens in $1/8^{th}$ of a second. [1].

Facebook Entity Graph

Facebook's Entity Graph, a complicated data set which maps the social network's 100+ billion connections between people, places and interests. The Entity Graph became the backbone for the entire Social Graph, and the technology behind Graph Search.

Facebook's Entities team was first founded with the task of transforming the plain text descriptions in user's profiles into structured data. Put simply, the team worked to add meaning to this text, the importance of which cannot be understated: without being able to process its user's entries, Facebook would have no way to do anything with your data; no ad targeting, no business model and no helping you stalk your old friends from high school.

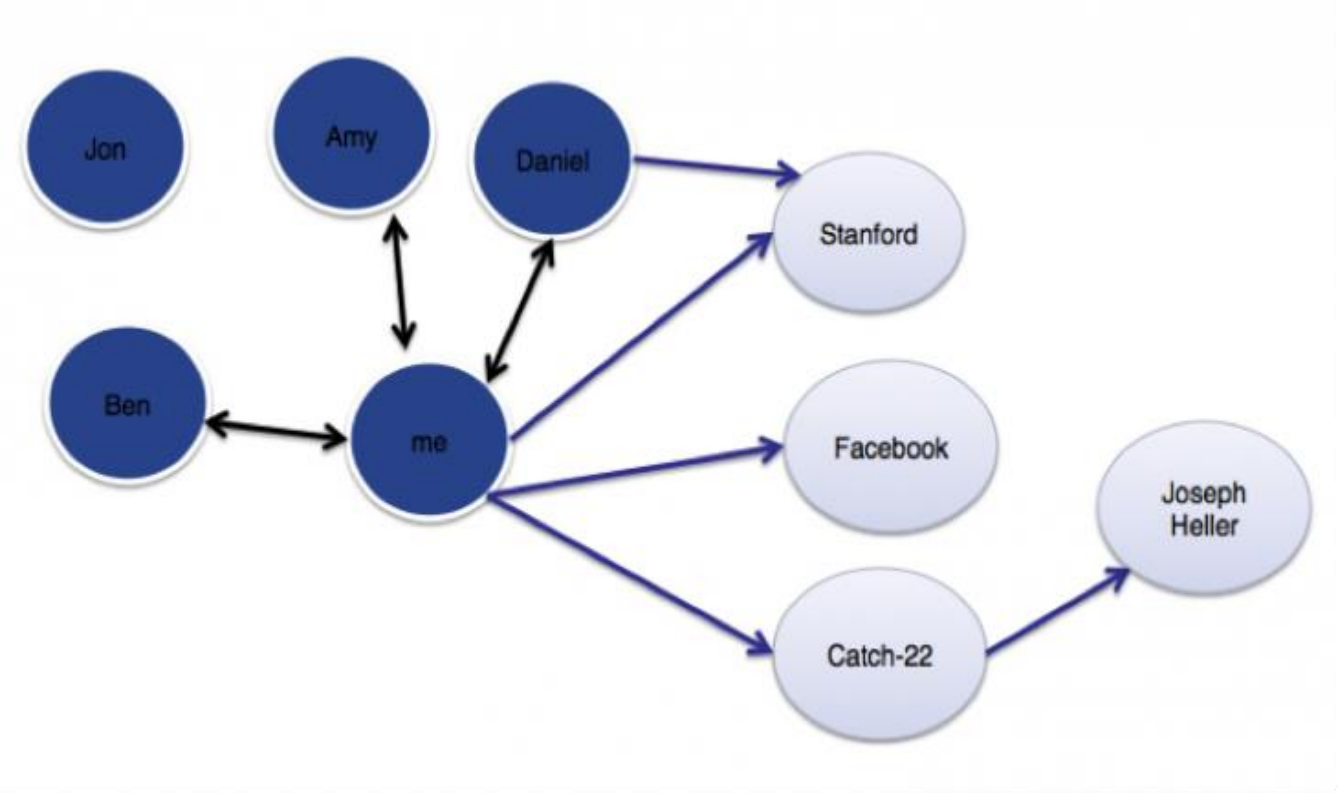

Figure 2: Facebook Entity Graph

In order to take advantage of all of those juicy details in your profile, Sun said his team had to find a data set to represent a seemingly unlimited number of interests. Their solution was to tap into Wikipedia, which powered Facebook's creation of "millions of 'fallback' pages." Facebook heavily relies on Wikipedia to this day.

These fallback pages were matched to interests that couldn't be connected to pre-existing pages. Afterwards, they were manually vetted for duplicates; ones which didn't receive any connections were deleted. According to Sun, many other quality control challenges came into play, including the handling of film remakes, "shortened forms of book names," as well as "ASCII art and obscenities."[2].

Methodology

I have studied about the graph theory for searching to carry the project work smoothly. I have collected various information about the searching technique. According to project information that I have collected, I have tried to build my project into block diagram for better understanding of my project and work smoothly to complete the project. I have studied about the related system and come up with the system architecture. The System architecture of my system is

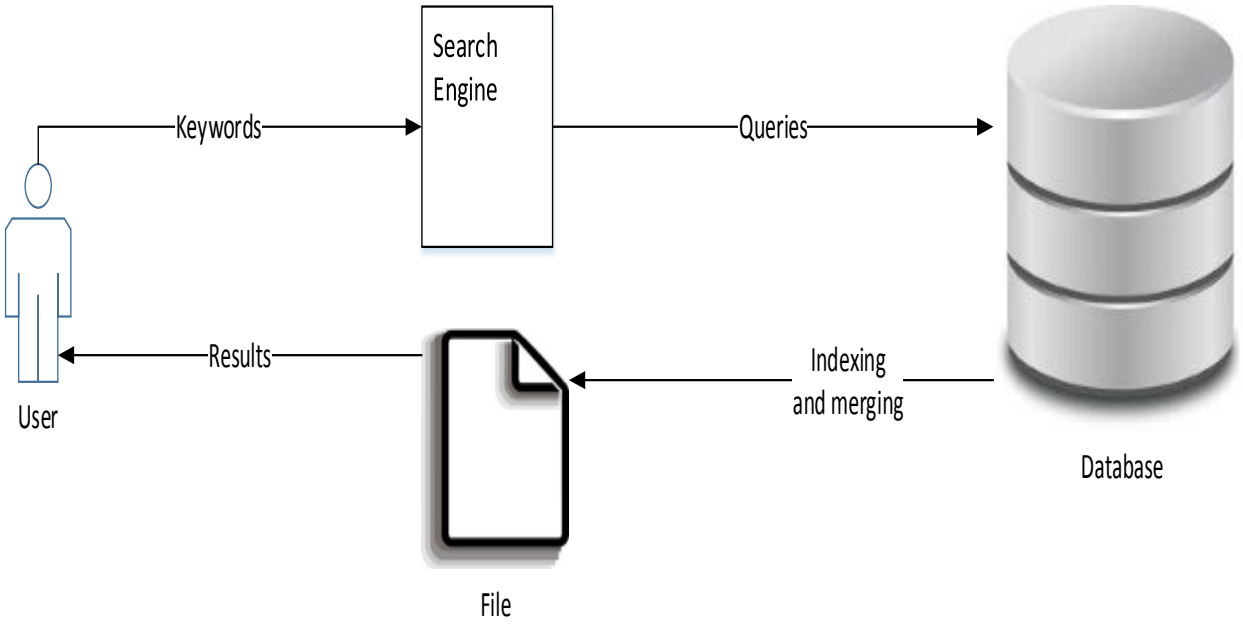

Figure 3: Block Diagram Of search Engine

Work Plan

I have already finished studying and collecting information required to carry out project work smoothly. I will be carrying out the software development with the logical design phase followed by physical design. Then I will be developing and testing the system. So, to carry out the project development smoothly in accordance to the above approach, I have made a schedule to complete the development of the system.

ID	Task Name	Start	Finish	Duration	Jul 2013
					6/23 6/30 7/7 7/14 7/21 7/28
1	Literature Study	6/25/2013	6/27/2013	3d	
2	Program Design	6/28/2013	7/4/2013	7d	
3	Functioning prototype	7/4/2013	7/4/2013	0d	
4	Coding	7/4/2013	7/23/2013	20d	
5	Complete Code of The System	7/24/2013	7/24/2013	0d	
6	Testing	7/24/2013	7/25/2013	2d	
7	Final Test cases	7/26/2013	7/26/2013	0d	
8	Documentation	7/27/2013	7/30/2013	4d	
9	Final Document	7/31/2013	7/31/2013	0d	

Table 1: Project Development Schedule

Development Tools

To complete my project in time, I will be using various development tools. The tools that will be used can be listed as:

1. CodeBlocks
2. C++
3. Files as Database

Bibliography

1. Google. (2013). How Search works. In Google Inside Search. Retrieved June 28, 2013, from http://www.google.com/intl/en/insidesearch/howsearchworks/thestory/.
2. Harrison Weber. (June 6, 2013). How Facebook's Entity Graph evolved from plain text to the structured data that powers Graph Search. In The next web. Retrieved June 29, 2013, from http://thenextweb.com/facebook/2013/06/06/the-evolution-of-facebooks-entity-graph-the-structured-connections-behind-graph-search/.
3. SEW Staff. (March 13, 2007). How Search Engines Work. In Search engine watch. Retrieved June 25, 2013, from http://searchenginewatch.com/article/2065173/How-Search-Engines-Work.
4. Search engine details. Retrieved June 25, 2013, from http://ads.harvard.edu/pubs/A+AS/2000A+AS..143...61E/node4.html.
5. Linda Barlow. (November 05, 2004). How To Plan The Best Search Strategy. In How To Use Web Search Engines. Retrieved June 26, 2013, from http://www.monash.com/spidap1.html.