**ABSTRACT**

 This paper is concerned with information extraction from top-k web pages, which are web pages that describe top k instances of a topic which is of general interest. Examples include “the 10 tallest buildings in the world”, “the 50 hits of 2010 you don’t want to miss”, etc. Compared to other structured information on the web (including web tables), information in top-k lists is larger and richer, of higher quality, and generally more interesting. Therefore top-k lists are highly valuable. For example, it can help enrich open-domain knowledge bases (to support applications such as search or fact answering). In this paper, we present an efficient method that extracts top-k lists from web pages with high performance. Specifically, we extract more than 1.7 million top-k lists from a web corpus of 1.6 billion pages with 92.0% precision and 72.3% recall.

*Index Terms*—Web information extraction, top-k lists, list

Extraction, web mining

**The Architecture**

****

**EXISTING SYSTEM:**

 Nowadays We Have So many Search engines. But in that engine there is no Mostly member using pages search option. That’s why we will go for Mostly member using pages search option. In proposed system we will provide this option.

*Disadvantage:*

1. Searching is difficult to the users.
2. It’s time consuming process.

**PROPOSED SYSTEM:**

 In proposed system we will provide Mostly member using pages search option to the users. It is very use full to the user. It saves user’s time also. In this project we will go to develop so many things. Like web search and image search.

*Advantage:*

1. Searching is easy to the users.
2. It is save users time.
3. At a time we can search lot of information

**Hardware Requirements:**

* System : Pentium IV 2.4 GHz.
* Hard Disk : 40 GB.
* Floppy Drive : 1.44 Mb.
* Monitor : 14’ Colour Monitor.
* Mouse : Optical Mouse.
* Ram : 512 Mb.
* Keyboard : 101 Keyboards.

**Software Requirements:**

* Operating system : Windows XP.
* Coding Language : ASP.Net with C#
* Data Base : SQL Server 2005.