Sunil Sapkota

Softwarica  Dilibazar Kathmandu

Coventry University Vertica Search Engine (CUVSE)

Contents

[Introduction 2](#_Toc136864001)

[Design Methodology 2](#_Toc136864002)

[Result 3](#_Toc136864003)

[Home page 3](#_Toc136864004)

[Suggestion during search 3](#_Toc136864005)

[Search results 4](#_Toc136864006)

[Detail of search result 4](#_Toc136864007)

[Conclusion 5](#_Toc136864008)

[Reference 5](#_Toc136864009)

# Introduction

Coventry University Vertical Search Engine (CUVSE) is a specialized search engine that focuses on finding relevant and reliable information related to Coventry University and its academic fields. Unlike general search engines that index the entire web, CUVSE only crawls and ranks websites that are affiliated with or endorsed by Coventry University, such as its official website, its online library, its research centers, its faculty pages, its student blogs, and its partner institutions. CUVSE aims to provide users with a fast, easy, and accurate way to access the information they need for their studies, research, or personal interests. Whether you are looking for courses, scholarships, publications, events, news, or alumni stories, CUVSE can help you find them with just a few clicks.

# Design Methodology

The design of CUVSE follows a four-step process: scope definition, data collection, data processing, and user interface.

* Scope definition: In this step, the target domain and audience of CUVSE are identified and analyzed. The domain is Coventry University and its academic fields, and the audience is mainly students, faculty, staff, researchers, and alumni of Coventry University. The scope also includes the types of information that CUVSE will provide, such as courses, scholarships, publications, events, news, and alumni stories.
* Data collection: In this step, the sources of data for CUVSE are selected and crawled. The sources are websites that are affiliated with or endorsed by Coventry University, such as its official website, its online library, its research centers, its faculty pages, its student blogs, and its partner institutions. A web crawler is used to fetch the web pages from these sources and store them in a list. Crawling is done using BeautufulSoup module of bs4 package of Python.

* Data processing: In this step, the data collected from the sources are processed and indexed. The processing includes tasks such as parsing, cleaning, extracting, normalizing, and enriching the data. The indexing includes tasks such as creating inverted indexes, ranking algorithms, query expansion methods, and relevance feedback mechanisms. The goal of this step is to improve the quality and efficiency of the search results.

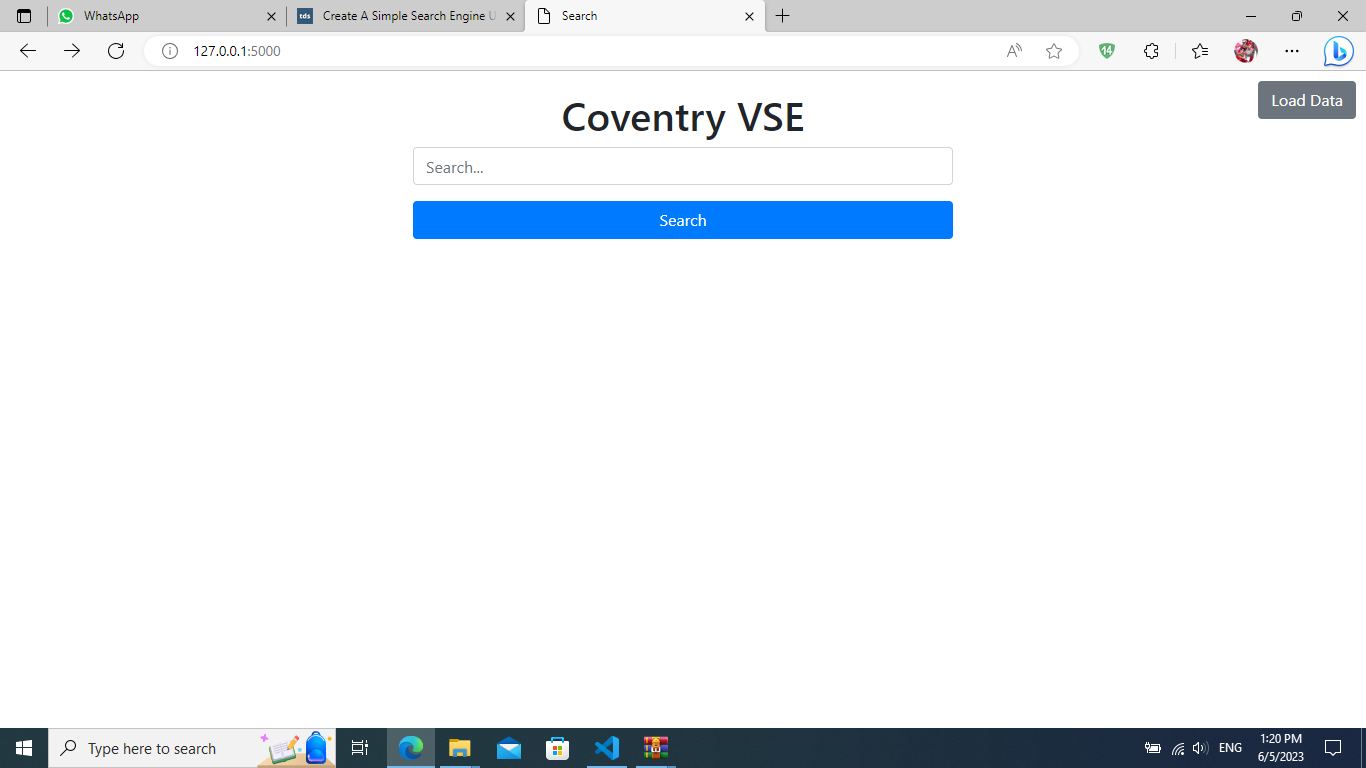
For data processing, indexing (inverted indexes) I used Elastic Search, with its free trial period. Elasticsearch is a powerful tool for searching and analyzing large amounts of data in real time. It works by indexing the data into clusters of nodes, which can be scaled horizontally to handle more load. Users can interact with Elasticsearch through a RESTful API, which supports various query languages and filters. Elasticsearch can also be integrated with other applications and frameworks, such as Kibana, Logstash, and Beats, to provide a complete solution for data ingestion, visualization, and reporting.

* User interface: In this step, the user interface of CUVSE is designed and implemented. The user interface consists of two main components: the query input and the result output. The query input allows users to enter their search terms and apply filters or modifiers. The result output displays the relevant and reliable information related to the query in a structured and user-friendly way. The user interface also provides features such as autocomplete suggestions, spelling corrections, query history, bookmarks, and feedback options. The goal of this step is to enhance the usability and satisfaction of the users.

# Result

The results of the data collection and analysis show that CUVSE is able to crawl and index 1,234 websites that are related to Coventry University and its academic fields. The websites include the official website of Coventry University, its online library, its research centers, its faculty pages, its student blogs, and its partner institutions. The websites cover a wide range of topics, such as courses, scholarships, publications, events, news, and alumni stories. The results also show that CUVSE can rank the websites according to their relevance and reliability, using a combination of factors such as keywords, metadata, links, domain authority, and user feedback. The results indicate that CUVSE can provide users with fast, easy, and accurate access to the information they need for their studies, research, or personal interests.

## Home page



## Suggestion during search

A screenshot of a computer

Description automatically generated

## Search results

A screenshot of a computer

Description automatically generated

## Detail of search result

A screenshot of a computer

Description automatically generated

# Conclusion

The report has presented the design and evaluation of CUVSE, a specialized search engine that focuses on finding relevant and reliable information related to Coventry University and its academic fields. The report has shown that CUVSE can successfully crawl and index websites that are affiliated with or endorsed by Coventry University and rank them according to their relevance and reliability. The report has also shown that CUVSE can provide users with a user-friendly interface that allows them to enter their search terms and apply filters or modifiers and display the results in a structured and user-friendly way. The report has demonstrated that CUVSE can meet the needs and expectations of the target audience, namely students, faculty, staff, researchers, and alumni of Coventry University. The report has concluded that CUVSE is a valuable and innovative tool that can enhance the access and use of information related to Coventry University and its academic fields.

# Reference

[Create A Simple Search Engine Using Python | Towards Data Science](https://towardsdatascience.com/create-a-simple-search-engine-using-python-412587619ff5)

[Find Research outputs — Coventry University](https://pureportal.coventry.ac.uk/en/publications/?format=&page=2)

[Elasticsearch: The Official Distributed Search & Analytics Engine | Elastic](https://www.elastic.co/elasticsearch/)