

---

# **School of Computing Creative Inquiry Effective Keyword Search**

---

April 30, 2017

Timothy McCabe, Brendan Giberson, Parth Patel, & Duyen (Tracy) Nguyen  
Clemson University  
CPSC 4910-001  
Seminar in Professional Issues II  
Dr. Alexander Herzog

## Table of Contents

<b>1 EXECUTIVE SUMMARY .....</b>	<b>2</b>
<b>2 PROJECT DESCRIPTION</b>	
2.1 Client Background .....	3-4
2.2 Problem Statement .....	3
2.3 Project Goals and Benefits .....	3
2.4 Development Methodology .....	3
2.5 Challenges .....	3
2.6 Cost and Scheduling .....	4
<b>3 RESOURCES NEEDED .....</b>	<b>4</b>
<b>4 REQUIREMENTS</b>	
4.1 Programming Language .....	5
4.2 Functional Requirements .....	5
4.3 Nonfunctional Requirements .....	5
<b>5 DESIGN .....</b>	<b>5-6</b>
<b>6 IMPLEMENTATION .....</b>	<b>6-7</b>
<b>7 TEST PLAN AND RESULTS .....</b>	<b>7-8</b>

Link to GitHub Repo

<https://github.com/bgibers/SynSearch>

Link to Site Aesthetic

<https://people.cs.clemson.edu/~bgibers/CRTINQ/index.php?>

## **1 EXECUTIVE SUMMARY**

The goal of this project is a fully functional search bar implemented on Creative Inquiry's current site. While a search bar currently exists on the site, there is much room for improvement. Upon delivery the search bar that is implemented should display results, even if the user has an input error. The search function should display results directly related to the search query, while potentially displaying results that relate as well (i.e. synonyms of search query). It is an improvement of the currently existent search function within the infrastructure. We had intended to implement Google Search API to allow synonym search, while also checking for typos. We later decided to use an open source synonym library to search through the Creative Inquiry project listings. The plan was to implement the search algorithm on our dummy site and then meet with our clients to figure out how to integrate the algorithm into their database and get feedback from them. But we were not able to get in touch with our clients so this was not possible. Instead, we made our dummy site resemble the actual Creative Inquiry website as closely as possible and made sure the search functionality worked on it.

## **2 PROJECT DESCRIPTION**

### **2.1 Client Background**

Our client, Dr. Cora Allard-Keese, is the Associate Director of Clemson University Creative Inquiry. Her job here is to oversee Creative Inquiry, as well as ensure the program is likeable, comprehensive, and well managed.

### **2.2 Problem Statement**

After meeting with our client and discussing potential problems with the current (Creative Inquiry) web infrastructure, we found that the web search function could be improved. Currently the site uses a direct word search, which looks for a specific word in their database. This does not account for typos nor does it consider synonyms.

### **2.3 Project Goals and Benefits**

Our end goal is to improve the current search function within the infrastructure. We intend to implement a method to allow synonym search, while also checking for typos. This would allow a student to view far more results for a desired CI attribute. Rather than getting no results for a search of say “automobile”, a student may receive hundreds with the synonym “car”. This would greatly improve the functionality of the website while also potentially helping Creative Inquiry to reach its full potential.

### **2.4 Development Methodology**

Our first step in the development process was to research various types of algorithms and APIs. This is to determine which would best fit the system. Next we would chose one and implement it onto a dummy site for further development and testing. Once fully operational, we would like to place the search bar onto the site’s aesthetic and eventually the actual live website itself.

### **2.5 Challenges**

The biggest challenge we faced was figuring out how we can create a way for synonym searching to actually link to the database. The challenge is tying our implementation to their database and banner because of possible security issues.

Another challenge was to figure out how to update the sitemap. A sitemap is a document that you can give to Google to tell the search engine the structure of the website. It essentially helps webcrawler navigate and find relevant information.

Contacting the Webmaster that’s overseeing the Creative Inquiry website to understand how the existing search worked.

## 2.6 Cost and Scheduling

The cost of the project implementation and maintenance starts at \$0, and more search features are provided by Google Custom Search for extra fees. The time given to the project is roughly nine weeks.

## 3 RESOURCES NEEDED

- Access to open-source APIs
- Privilege of creating database servers via Clemson
- Knowledge of HTML, CSS, MySQL

Dr. Cora Allard-Keese (Associate Director of Creative Inquiry) and Tullen Burns (Events Coordinator) was interviewed three separate times to gather information about the project.

*First meeting:* the idea of improving the current search function within the infrastructure was proposed by Dr. Allard-Keese.

*Second meeting:* they gave feedback and approved the project process description that our group had structured:

- Research various search algorithms and APIs to determine a better technique to develop and improve the search function of the Creative Inquiry keyword search engine
- Implement chosen search method on a dummy site while developing and testing
- Integrate the search bar into the site's aesthetic
- Migrate new search feature onto a live website

*Third meeting:* we attempted to get in contact with the Website Developer, Jonathan Harp through Dr. Allard-Keese, but it's still was not successful. Our group needed to find out if we are able to set up a Google search functionality to the CI Current Project's page. Instead, Dr. Allard-Keese provided more information and resources for our group.

*Fourth meeting:* Clients failed to attend the final meeting. Our group waited at the scheduled meeting area for 15 minutes for them. Dr. Allard-Keese did give an in-person apology on the missed meeting.

## **4 REQUIREMENTS**

### **4.1 Programming Language**

As a product owner for Creative Inquiry, I want the Effective Keyword Search platform to be written in PHP and CSS, so that our webmaster can extend its capabilities in the future.

### **4.2 Functional Requirements**

As a user, I want a platform that uses a synonym search-based algorithm, so that I can be served a more effective result set.

As a user, I want the platform to be web based so that I can access it using a browser without the need of installing a desktop application.

### **4.3 Nonfunctional Requirements**

As a product owner, I want the platform to showcase the Creative Inquiry logo, so that it is brand-aligned.

As a user, I can use major browsers like Chrome, Firefox, Safari, and Internet Explorer, so that it can be versatile.

## **5 DESIGN**

The design of this project was aimed at being as close to Clemson Creative Inquiry's actual website as possible. Their site features their logo with a list of the available creative inquiries along with their associated number. At the top of the page there is a search bar which performs a simple direct word search. For our site we replicated all of these aspects with a small adjustment to the search. The following images are from Clemson Creative Inquiry's current site design. The first image is the page displayed upon entering the site(Creative Inquiry : Find a Project). The second is our site replicated very similarly.



Find Creative Inquiry projects that need students with your interests.

Project #	Project Details	Interested?
1231	<p><a href="#">Minorities-Majorities Gaps in Life, Death and Migration</a></p> <p>This experience welcomes those interested in the study of fertility, mortality, and migration to decompose the heterogeneity of these behaviors by sex, age, race, and ethnicity. Decomposition techniques have pointed to interesting paradoxes in population studies, such as the epidemiological "Hispanic paradox", the "Male-Female Health-Survival Paradox", or the "Immigrant fertility" paradox. This hands-on project uses public-use worldwide demographic data on health, migration and reproductive behavior by race, ethnicity, and sex, which trigger the opportunity of applying statistics and demographic techniques in finding new demographic paradoxes for a broad range of developing and developed countries.</p> <p><b>Majors Wanted</b>            Sociology            Health Science            Applied Economics and Statistics            Economics            Any/All Social Science            Psychology</p>	<p>Deadline Aug 13, 2017 <input type="checkbox"/></p>
1230	<p><a href="#">Sustainable Spools</a></p> <p>This project will focus on reducing the impact of 3D printing on the environment by developing sustainable methods of producing printing filament by utilizing waste plastic to generate 3D printing filament for use in prototyping.</p> <p><b>Majors Wanted</b></p>	<input type="checkbox"/>



Find Creative Inquiry projects that need students with your interests.

Project #	Project Details
1	<p><a href="#">Effects of environment stress on cell growth</a></p> <p>Cells respond to their environment by modulating their growth and/or changing their behavior. This CI will focus on understanding how environmental stressors such as low-dose radiation and heat affect human cells. While these stresses are not large enough to kill cells they can significantly affect the cells behavior. This project will involve experiments as well as computational modeling techniques. By using a combination of experimental, analytical and computational techniques, the team will work on various problems related to cell growth modeling in response to environmental stress.</p>
2	<p><a href="#">Mirrors and Windows: A Children's Literature NewsI</a></p> <p>The purpose of this CI project is to create a newsletter that focuses on children's literature intended for youth in grades K-6. The books will be diverse in terms of genres (e.g., poetry, fantasy, nonfiction, etc.) as well as the race of the authors and illustrators. Students will read 6-8 titles and select their favorite book to review.</p>
3	<p><a href="#">Conservation of Marine Resources</a></p> <p>The Conservation of Marine Resources creative inquiry team participates in marine ecology research to benefit the Florida Keys National Marine Sanctuary. Current projects include studying (1) the direct and indirect impact of parrotfish grazing on the survival of coral transplants, and (2) the impact of mass sponge mortality on the survival of juvenile Caribbean spiny lobsters.</p>
4	<p><a href="#">Courage Research</a></p>

## 6 IMPLEMENTATION

Implementing the synonym parser with a php script took some time. First, a database was created. Placed in this database were a dozen or so projects taken from the Creative Inquiry website. Next a generic website was developed that simply just displayed the information coming from the database. After ensuring the data was

displaying properly, it was placed into a neat table making the site easier on the eye. After learning a bit of php, I created a search bar that searched through the database and grabbed any string with the exact query. The query takes any set of letters so long as the length is above three. Upon ensuring the direct word search was properly working, the next step was to implement the synonym searcher (Thesaurus web service). After playing with the prewritten php that outputted to json I was able to see how the words were being displayed. I edited their script until I was only shown synonyms rather than antonyms and the definitions as well. Finally I linked the two scripts together by piecing lines together until the desired result was achieved. After testing several searches based on a synonym it was concluded that the script was successful. The final step was creating a CSS stylesheet as well as a better looking table. The site was tweaked until it closely resembled Clemson CI's.

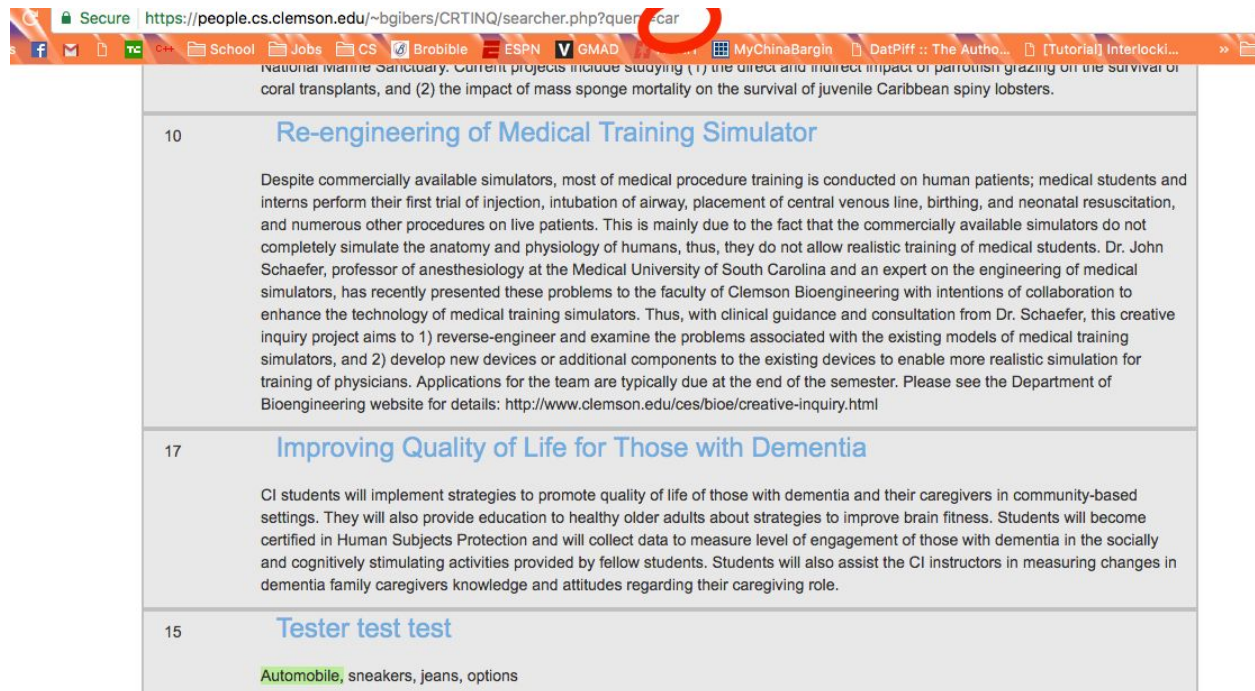
## 7 TEST PLAN AND RESULTS

Our test plan involved testing the site we created with numerous different queries. We then tested the same queries on Clemson CI's site and compared the results. Our site was a success as it does in fact handle numerous synonyms for just about any word. The following image shows searching the word car on Clemson CI's website. Using the find command I searched automobile. Only one result was displayed, and this was due to the fact the word car was also in the paragraph.

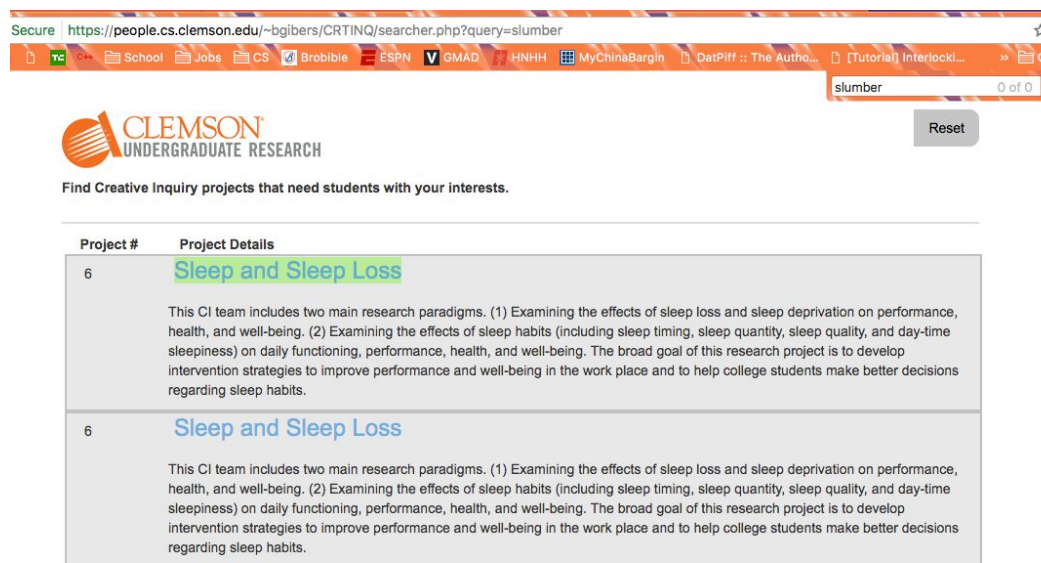
Majors Wanted	Any/All Majors
921	<p><b>CEDC 14: Health and Transportation Infrastructure Development</b></p> <p>This project is attempting to find solutions to problems faced by the Haitian population. Our medical equipment team will help other organizations redesign donated equipment by utilizing materials accessible to the Haitian people. Success for both of these teams means improving health care in Haiti and increasing the capacity of lake transportation in rural Haiti.</p>
1191	<p><b>Project Gateway</b></p> <p>Project Gateway will offer students an interdisciplinary opportunity that exists at the intersection of art and technology. The mission is to leverage the educational ecosystem of students and industry to re-purpose existing <b>automobiles</b> from Cuba for use in the new open region of tourism in Cuba. The project will be a 4-semester, pioneering effort that employs a systems integration approach to preserve the artistic design of a vintage vehicle while incorporating current vehicle technology advancements to combine the look of yesterday with the technology of today. Ideally, all student participants would be sophomores or juniors. The first semester, students will create a business plan for Project Gateway. The second semester, they will implement the branding/marketing and fundraising plan. The second year, students will build the <b>cars</b>.</p> <p><b>Majors Wanted</b>            Computer Engineering            Electrical Engineering            Ceramic and Materials Engineering            Management (Entrepreneurship Emph)            Management (International Mgt Emph)            Financial Management            Marketing            Accounting            Visual Arts            Any/All Business</p>
630	<p><b>Human Powered Vehicle Design</b></p> <p>Human Powered Vehicle Challenge (Additional information is available: <a href="https://www.asme.org/events/competitions/human-powered-vehicle-challenge-(hpvc)">https://www.asme.org/events/competitions/human-powered-vehicle-challenge-(hpvc)</a>)            Human-powered transport is often the only type available in underdeveloped or inaccessible parts of the world, and if well designed, can be an increasingly viable form of sustainable transportation. ASME's international Human Powered Vehicle Challenge (HPVC) provides an opportunity for students to demonstrate the application of sound engineering design principles in the development of sustainable and practical transportation alternatives. In the HPVC, students work in teams to design and build efficient, highly engineered vehicles for everyday use—from commuting to work, to carrying goods to market.</p> <p>Contact: Dr. Mocko (gmocko@clemson.edu) or Alix Griffin (team lead - alix@g.clemson.edu)</p>
878	<p><b>Seeking Insight on the BioMedical Applications of Magnetic Nanoparticles: Imaging and Therapy</b></p> <p>The ultimate goal of this project is to bridge the synthetic capabilities of laboratories at Clemson with analysis of the cellular and animal interactions at the Instituto de Ciencia de Materiales in Madrid, Spain (ICMM) and University of British Columbia (UBC). The end goal of the project is to evaluate the effect nanoparticle size has on the way particles</p>



The following image is our site where the same test was performed. Circled in red is the search query car. If you look at the bottom in green the result automobile was displayed without the word car being in the paragraph.



Several other tests were performed such as searching the term slumber on our site. Using the find command not a single result for the query slumber came up. Yet our site was still able to show results based off the query. The query pulled words such as sleep and displayed their results.



### **8 References**

1. "Creative Inquiry : Find a Project." *Clemson University*.  
[http://ci.clemson.edu/forms/find\\_a\\_project/](http://ci.clemson.edu/forms/find_a_project/).
2. *Thesaurus web service*. <http://thesaurus.altervista.org/>.