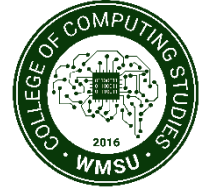




Republic of the Philippines
Western Mindanao State University
College of Computing Studies
DEPARTMENT OF COMPUTER SCIENCE
Zamboanga City



Knowledge Sharing and Intellectual Property Preservation through an Online Repository of Electronic Theses and Dissertations (ETD): The Crimson's Legacy

A Thesis presented to the faculty of
Department of Computer Science
College of Computing Studies

In partial fulfillment of the requirements for the degree of
Bachelor of Science in Computer Science

Ronald M. Arcilla
Emmanuel L. Toledo
Researchers

Mr. Mr. Gadmar M. Belamide
Adviser

March 25, 2022

Western Mindanao State University
College of Computing Studies
DEPARTMENT OF COMPUTER SCIENCE
Zamboanga City

Approval Sheet

The Thesis attached hereto, entitled “**Knowledge Sharing and Intellectual Property Preservation through an Online Repository of Electronic Theses and Dissertations (ETD): The Crimson’s Legacy**”, prepared and submitted by <Researchers Ronald M. Arcilla and Emmanuel L. Toledo, in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science, is hereby **recommended for Oral Examination**.

Mr. Gadmar M. Belamide
Adviser

APPROVED by the Oral Examination Committee on _____ with a rating of **PASSED**.

<Panel 1>
Chairperson

<Panel 2>
Member

<Panel 3>
Member

ACCEPTED in partial fulfillment of the requirements for the degree of **Bachelor of Science in Computer Science**

ODON A. MARAVILLAS JR, MSCS
Head, Department of Computer Science

RODERICK P. GO, Ph.D.
Dean, College of Computing Studies

Acknowledgment

First of all, congratulations and thank you to the researchers' effort to finish this study. Hope they are proud of themselves.

To their adviser, Mr. Gadmar M. Belamide, for his great guidance throughout the research and his trust towards the researchers which gave them the fuel they needed during the final defense for this study.

To the Arcilla family, Ma. Luisa B. Arcilla and Ronalyn M. Arcilla provided the researcher with his needs which opened a path for him to accomplish the study. And to Roven Niño M. Arcilla for being an inspiration to the researcher.

To Toledo Family

To the College of Computing Studies Faculty, which has been a big part of the researchers' progress for 3 and a half years which led to the success of this study.

To Axl Cuyagan, who taught Linux programming to the researcher, helping him deploy the website which is a goal of the study.

And to all the friends and special people of the researchers. They have helped the researchers overcome this challenge in a more tolerable manner and environment.

The researcher owes this study to these people. Thank you so much.

Abstract

Electronic Theses and Dissertations (ETD) is simply a digital representation of a thesis or dissertations. Through preliminary research it has been ruled out that the use of ETD is not common here in Zamboanga City. Every generation of graduating students has been contributing to intellectual resource of the world in the form of a thesis consuming papers and increasing the number of physical documents stored inside a school. With the help of ETD, physical storage become irrelevant since ETD are stored in an online repository which secures intellectual property preservation and the publicity of this intellectual resources will multiply because of the internet, the information super highway.

This study proposes a website that is suitable to be an online repository for ETDs. With the use of HTML and CSS, Flask and Python, MySQL, Natural Language Processing for a search result and search recommendations, and a web-hosting service from Hostinger to deploy a functional website along with an added functionality of a plagiarism checker with the use of Copyleaks API.

The finished system is then subjected to testing from Western Mindanao State University Faculty and students to test the satisfaction the website can give to its users. And the results gathered from these tests shows that the users are satisfied with the end product.

Keywords: Electronic Thesis and Dissertations, Website, Plagiarism Checker

Table of Contents

Approval Sheet.....	i
Acknowledgment.....	ii
Abstract.....	iii
CHAPTER I INTRODUCTION	1
Background of the Study.....	1
Statement of the Problem	2
Objectives.....	2
Scope and Limitations.....	3
Significance of the Study.....	4
Theoretical Framework (if only applicable)	Error! Bookmark not defined.
Definition of Terms.....	5
CHAPTER II REVIEW OF RELATED LITERATURE	6
Related Studies	Error! Bookmark not defined.
Synthesis	Error! Bookmark not defined.
Conceptual Framework.....	15
CHAPTER III METHODOLOGY	16
Research Design	Error! Bookmark not defined.
Respondents.....	Error! Bookmark not defined.
Data Gathering Instruments, Techniques, and Procedures.....	Error! Bookmark not defined.
Statistical Tools (if only applicable)	Error! Bookmark not defined.
Analytical Tools.....	Error! Bookmark not defined.
Technical Tools.....	Error! Bookmark not defined.
Software Process Model	Error! Bookmark not defined.
System Architecture.....	Error! Bookmark not defined.

Implementation Plan	Error! Bookmark not defined.
CHAPTER IV RESULTS AND DISCUSSION	29
CHAPTER V CONCLUSION AND RECOMMENDATIONS	35
Conclusion	35
Recommendations	36
References	37
Appendix A Gantt Chart.....	39
Appendix B Evaluation Tool.....	Error! Bookmark not defined.
Appendix C Photo Documentation of Testing	Error! Bookmark not defined.
Appendix D Flowchart/Diagrams	40
Appendix E Algorithms	44
Appendix F Relevant Source Code.....	Error! Bookmark not defined.
Appendix G Screenshot/Picture of the System	46
Appendix H Test Cases and Results	50
Appendix I Bug Report	Error! Bookmark not defined.
Appendix J User Manual.....	Error! Bookmark not defined.
Appendix K Curriculum Vitae.....	51

List of Figures

List of Tables

List of Appendices

CHAPTER I

INTRODUCTION

Background of the Study

Since the establishment of the concept of Electronic Theses and Dissertations (ETD), the development of software for ETDs has been introduced to the world to create, publish, manage, and archive ETDs. Nations overseas have been implementing the e-only theses and dissertations where only ETDs are submitted in the university's chosen repository and they do not require students to submit physical documentation of their research. This movement has been operating to share researched knowledge and increase a scholar's research publication, which increases their reputations for their respective professional fields.

In the Philippines, not many universities have been participating in the concepts of ETDs. So far in research, the only entities that have been using ETDs are the Philippine Electronic Theses and Dissertations (PETD) that hosts databases for ETD for masters and doctorate degrees of colleges or universities, as well as the Digital Archives@UP Diliman (DA@UPD) which caters to the overall digital records of the institution where ETDs are included therein.

In Zamboanga City, the concept of ETDs is not yet practiced although there are some efforts in implementing the said concept such as Western Mindanao State University and their e-library in Facebook and another e-Library is implemented in Zamboanga City State Polytechnic College which may be a start for ETDs. May this research be of aid to the city so that ETDs may be implemented in the researchers' schools because their team believes in the goals and benefits of archived and shared knowledge to assist the future generation for successful advancement in each generation of researchers.

Statement of the Problem

At the last year of college students, a terminal requirement, which is mostly thesis, is their final challenge before graduating in their respective degrees. Anyone who has gone through knows how hard it is to find references and similarities, especially if the study is new in the world. Also, thesis hardcopies, besides that they add costs for bookbinding and printing, have a chance of being neglected or even discarded as the year and generations of graduating student passes by.

So, what is the best way to promote a shared library of knowledge and at the same time, cut costs for hardcopies and solve the problem of preserving physical documents of the thesis each year? The researchers are proposing a WMSU online compiler for Electronic Theses and Dissertations, which, if possible, be a success and be a start of a new era of creating and preserving thesis in Zamboanga, per school year.

Objectives

1.3.1 General Objectives

The General objective of this thesis proposal is to develop a website that will serve as an online repository of thesis and dissertations of WMSU students, for WMSU students and the public and to create a project that encourages the academe and other concerned agencies in Zamboanga City the utilization of ETDs.

1.3.2 Specific Objectives

1. To use HTML, JavaScript, and CSS for web design for the online repository website.
2. To deploy the proposed website online.
3. To manipulate and sort data with Python, Flask and MySQL in the system for results based on user's search query and produce abstract similarities.
4. To increase the website's functionality with a Plagiarism Checker API Copyleaks.

Scope and Limitations

Scope

The scope that is included in the researchers' study is the impact of a compilation of past thesis for the future generations of students to use and even, maybe, anyone who is interested in studying almost anything. The research shall survey a range of students who are currently enrolled in WMSU and is taking a thesis subject. The primary target users are thesis students at Western Mindanao State University, Zamboanga City. This research will also study the algorithms that are crucial for the development aspect which is sorting, searching, and comparing algorithms.

Limitations

The limitation of the researchers' study is that the research will not include students outside of WMSU and students who are not taking up a thesis.

Significance of the Study

The results of this research will be significant to the following:

Thesis Students. The main target beneficiary of this project is the students or the future researchers, specifically those who are graduating and under a thesis subject. The students shall benefit from the shared knowledge stacked by each generation of students through the researchers' online compilation of thesis in WMSU, cut printing and bookbinding costs, and preserve the output of their hard work, their research.

Professors and Instructors. The pursuit of knowledge is endless and even those who have a degree will benefit from shared knowledge.

Western Mindanao State University. The said university will have a solution in terms of research or thesis compiling as well as being the first university to promote ETDs in Zamboanga.

Other Colleges or Universities. The success of this project may be a start of a whole new era of thesis education.

Commission on Higher Education. Improved quality both for education and research.

Public. Even a regular person who is just interested and curious. The compilation of ETDs on the researchers' website will be open to the public for a better network of shared knowledge.

Future Researchers. For the next generations of students to come.

Definition of Terms

Only terms, words, or phrases which have special or unique meanings in the study are defined. Provide at least 10. Definitions may be taken from encyclopedias, books, magazines and newspaper articles, dictionaries, and other publications but the researcher must acknowledge his sources. Definitions taken from published materials are called conceptual or technical definitions.

Term	Definition
1. Chunking	the division of large information into smaller chunks which are easier to hold in memory.
2. ETDs	an acronym for Electronic Theses and Dissertation.
3. Inflected Words	word formation where the word's meaning is changed through adding another item on it.
4. Lemmatization	reduction of a word's inflected form into a single form through scanning WordNet corpus as reference.
5. Named Entity Recognition	identify named entities used within a sentence.
6. Parts-of-Speech tagging	identification of parts of speech for every word.
7. Stemming	cutting the inflected word with algorithm alone into their root form.
8. Tokenization	sensitive data turned into a non-sensitive data called token.

Table 1: Definition of Terms

CHAPTER II

REVIEW OF RELATED LITERATURE

2.1. Related Literature

Google Scholar and Microsoft Academics on Online Repository User Interface

Google Scholar and Microsoft Academics are famous for being an online search engine for studies, articles, thesis, dissertations, and court opinions from scholars and reporters. These products are obviously from Google and Microsoft, two of the successful companies that invent technologies both for hardware and software, and will be a very safe and secure reference to use for the researchers' proposed system's user interface which gives us all the insights and principles for the user experience that will surely be credible.

From the main search bar at the landing page to the sorting section on the side panel, these settings and formats shall be observed in the proposed website's UI. The results from the search query and all its linking will be integrated into the systems recommendation results which are also based on the user's search query.

These products' search methods will be closely studied and observed and implemented on the researchers' proposed website.

Networked Digital Library of Theses and Dissertations on Electronic Theses and Dissertations

The concept of Electronic Theses and Dissertations (ETD) has been around even since 1990, and organizations like the Networked Digital Library of Theses and Dissertations (NDLTD, 1996) and Philippine Electronic Theses and Dissertations (PETD, 2020) have been encouraging and advocating the use of ETD. But what is an ETD?

ETD is basically a digital representation of your thesis or dissertations which can be stored in a repository may it be online or in local computer memory. Here is a short history in ETD as discussed by NDLTD:

The concept of electronic theses and dissertations (ETDs) was first discussed at a 1987 meeting in Ann Arbor, Michigan, organized by UMI and attended by representatives from Virginia Tech, the University of Michigan, and two small software companies – Toronto-based SoftQuad and Michigan-based ArborText.

The project lay dormant for a few years, until 1991 when Virginia Tech's Dean Gary Hooper financed further critical development. Virginia Tech Computer Science professor Ed Fox and Graduate School dean John Eaton collaborated on the ETD project, investigating problems associated with production, archiving, and access. In the early 1990s, Fox and Hooper held a series of design and discussion meetings, working closely with the Coalition for Networked Information (CNI), the Council of Graduate Schools (CGS), UMI, and other interested groups. At the same time, the Virginia Tech University Library's Scholarly Communications Project developed procedures and systems for processing, archiving, and providing public access to Virginia Tech's graduate research works.

The said organization also has discussed publishing ETDs to repositories where it is stated that some major research universities, nowadays, provide a specialized thesis repository run by the university itself and that it is a graduation requirement at many institutions. Such statements strengthen the validity of this research's purpose.

Federis on Copyright Laws in the Philippines

On a website called Federis, one of the top filers of trademark and patent applications in the Philippines and is a law firms of highly qualified attorneys, lawyers, licensed patent and trademark agents, has deep insight and well-discussed topic on intellectual property protection which is one of the concerns of this research and development.

It states on their website that under Philippine law, a copyright infringement happens when any of the exclusive economic or moral rights granted to the copyright owner is violated. A person is considered liable for such infringement of a copyrighted article when they:

- a. Directly selling such article.
- b. Distribute such articles for trading.
- c. Publicly exhibit such article.

They have also discussed that copyrightable works (artistic or literary works) are protected from the moment they are created and enumerated works that are not protected by the said law which are:

1. Ideas such as a procedure or system method.
2. News of the day or press information.
3. Official text of legal nature.
4. Work of the Philippine Government.
5. Lectures, research, speeches, and regulation rendered in government agencies and meetings of public characters.

And lastly, the most important note, is that they have discussed who are the owners of copyrightable works which is no one but the author of the work themselves.

Ciencia: Online Publication of Research by WMSU

Ciencia is the official Science and Technology research journal of Western Mindanao State University and has an open-access website for public sharing of the researches of WMSU. Ciencia was formerly called WMSU Research Journal when it began its organization with the purpose to give faculty researcher an avenue to showcase their research, but now till the present day, it has evolved to be an organization with an International Standard Serial Numbers (ISSN) for both of their printed and online publication which is a great advantage since an ISSN gives a publication its identity among hundreds, if not thousands, of other publications. Ciencia is also indexed in the Philippine E-Journals. The organizations also envision being accredited by CHED and be indexed internationally.

the researchers' study and development also aim to produce the same results and outcome as Ciencia and this literature can give Crimson's Legacy a good path on where to start and where to head-on.

Forbes on 2020 Website Design Trends

User Interface design is very crucial when you want to develop a website that will last long or forever and anything constant should give comfort to the ever-changing and moving, users.

Based on Forbes' website blog is that at the present day, users tend to demand a fast and smooth interface versus the users 10 years ago who use websites that aren't very user-friendly. And here are some tips from Forbes we can use to upgrade their project's UI design.

1. Dark Mode – is a very mainstream trend and gives the user option for a less bright and cleaner set-up for those who use their devices for a longer time span.

2. Hand-drawn elements (digital art or scanned graphics) – everyone can always appreciate hand-drawn art.

3. Visible Grids – even in some artworks, not removing the guideline adds an element to finished art. Visible grids can give a website its clean and professional look.

4. White Space Frame – spans a user's attention towards the product, a good technique paired with minimalist design.

5. Larger-Than-Life Typography – good for any website themes. Command the reader's eyes even from afar!

6. Geometric Designs – a sure way to create a modern and minimalist design for the researchers' website. Shape speaks.

7. Minimalist Navigation – Less is more.

8. 1970's Color Scheme – to create a classic and timeless design.

Web Hosting by Hostinger

Their proposal is a website so they should study web hosting and this subtitle helps us understand what web hosting is. Hostinger is a web hosting provider. And as Hostinger explains:

“Web hosting is an online service that enables you to publish your website or web application on the internet. It is basically renting some space on a physical server where you can store all the files and data necessary for your website to work properly”

After purchasing their service, they now buy a domain name which is simply the name that the researchers' website will use. But some Hostinger package already has a free domain name which may be considered when purchasing the said service. This study will help this research and product to reach millions of people through the internet.

Edureka (YT) on Stemming and Lemmatization Tutorial: Natural Language Processing

To attain the complexity of taking a thesis' abstraction into keywords, They would need to study Natural Language Processing (NLP) and the stemming algorithm and lemmatization algorithm which will help us capture keywords in researches.

Based on the video NLP is the part of Computer Science and Artificial intelligence that deals with human languages which are perfect in dealing with researches. Also, the video stated that NLP has different steps which are:

- Tokenization

- Stemming

- Lemmatizations

- Parts of Speech tags

- Named Entity Recognition

- Chunking

Stemming and lemmatizations algorithms have been around since the 1960s and it amplifies a system's tagging system, indexing, SEO, web search result, and information retrieval which is crucial for the researchers' proposed system as it will be their system's core process.

2.2. Related Studies

Yale Fineman on ETDs

According to Yale Fineman in the research included in Portal: Libraries and Academy: Theses and dissertations are scholarly works that take years to research and to write and as secondary sources of information, theses and dissertations are particularly useful to researchers in the humanities because they are proven facts that are stored and archived for the future generations of the researcher. But, the vastness of the population conducting research each year creates a problem in terms of storing physical copies of the research in archives and university libraries. The best way to bring back published researches to light is to upload them electronically and to give students and researchers free and open access to these documents via the internet.

These findings can further increase the validity of the advantages of ETDs.

The Research Wizard by Kare Hein and Marc Davis

Their proposed system or website's contents are purely dependent on the inputs and creation of users or thesis students of Western Mindanao State University so it is only natural to study the nature of a Database-Driven Website with HTML and PHP/MySQL. They were also advised with a complexity where they need to find a way or an algorithm that'll match research topics based on its contents. And in this subtitle, they are going to dive deeper into the said topic.

This study was created when librarians had a problem with providing patrons access to electronic format resources in their libraries. The solution was a web application, The Research Wizard. Its main function was to amplify results in a search that will cater to patrons' or clients' demands. It uses a topical keyword research access to do its job and this study is perfect to adapt in the researchers' proposed system where they need to give users compared works through the abstraction of research.

Mubina Malik and Trisha Patel Database Security

One of the biggest factors in deploying a website is security since placing it on the internet means placing it in reach for the public. In their conclusion, they have stated that website access protection begins in studying who has access to the database and what type of data the attackers might desire. It is also stated that 48% of attackers are the database's authorized users which means 48% of users who misused their access and privileges and 48% is a big chunk in 100 even without the fact that authorized user should be trustworthy. This study focuses on control and attack methods in a database.

Comparison Table of Relates Systems

Attributes	This Study	Google Scholar	Microsoft Academics	Fedora	Ciencia
Deployed Website	✓	✓	✓	x	✓
Upload Studies	✓	✓	✓	✓	✓
Plagiarism Checker	✓	✓	x	x	x
Open for Public	✓	✓	✓	x	✓
For WMSU students	✓	x	x	x	✓
Focused on Theses and Dissertations	✓	x	x	x	x
Able to download ETDs	✓	✓	✓	✓	✓
Aims to Eliminate Paper Documentation	✓	x	x	x	x

Table 2: System Similarity Table

Synthesis

With all the related studies and literature that have been discussed in this section, they have summarized points that will be implemented in the researchers' system.

- Google Scholar and Microsoft Academic's user interface is created by great developers from these two great companies. It shall be observed when designing their proposed website
- Technologies that will be used in this study are:
 - HTML, JavaScript, and CSS for front end development
 - Python, FLASK, and MySQL for Backend Development
 - Hostinger for Web Hosting

The algorithms used will be of the field of machine learning and Natural Language processing which will collaborate to handle the data management of the ETDs.

Conceptual Framework

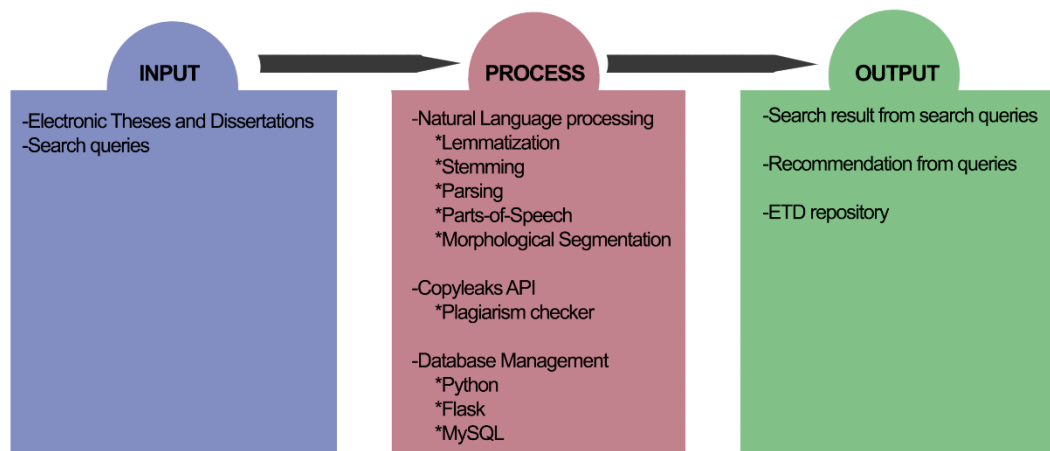


Figure 1: Crimson Legacy's Conceptual Framework

This conceptual framework gives us a proper understanding of how all the fields and subtitles of this study correlates with each other in the goal of creating an online repository for Electronic Theses and Dissertations. they can see in the diagram that the two main data collected on the researchers' website are search queries and ETD uploads.

The process handles the data with their corresponding fields and algorithm and is explained as follows:

Natural Language Processing – breaks the ETDs and search queries into small chunks of data which gives as the ability to clearly cluster and associate a whole abstract with another giving the researchers' website a powerful and accurate recommendation system for a user's search.

Copyleaks API - a ready to use developer tool that gives the website the ability to scan the uploaded ETD for a plagiarism score both locally (server storage) and globally (internet).

Database Management – the repositories for all data that will be used in all the process of the researchers' website.

CHAPTER III

METHODOLOGY

Research Design

This study will purely be applied research design as to prove a hypothesis and not formulate one. This study aims to develop a website which is an online repository for electronic theses and dissertations for WMSU that caters thesis submissions from thesis students with the use of Natural Language Processing and Copyleaks API

Research Locale

The research will be conducted online directed towards people who are students of an institution that is one of the pioneer state universities here in Zamboanga City

Validity of the Instrument

The research Instrument shall be validated by the thesis adviser and presented to the thesis panel.

Data Gathering Procedure

In gathering data for this study, the researchers use information from the internet such as ND LTD website which is an organization that has been campaigning and advocating the utilization of ETDs for universities. The website houses a lot of facts regarding Electronic Theses and Dissertations (ETD) and even online repositories.

The development aspect of the system was enforced by the studies made in the RRL (see chapter 2) for the technologies that will be used in the development such as HTML, JavaScript, CSS, Python, and MySQL, and as well as the computer science field that is present in this study which is Machine Learning and Natural Language Processing. All of these will also be practiced by the researchers before the development begins.

The dataset needed for Machine Learning was also all gathered online and will be integrated into the system after a data cleaning.

And lastly, the proposed website will heavily reference PETD's website as a guide towards a successful user interface for online repositories and therefore are closely observed for this development.

Statistic Tool

Technical Aspect of the Project

Natural Language Processing

- Tokenization
- Stemming
- Lemmatizations
- Parts of Speech tags
- Named Entity Recognition
- Chunking

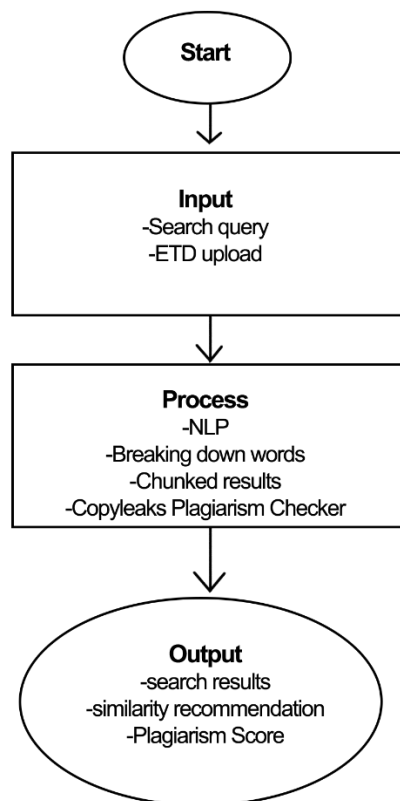


Figure 2. Natural Language Processing Flowchart

This algorithm will help us break down paragraphs, sentences, phrases, and inflected words so that algorithm can deal with the overall data of a search query

Search Queries Code Snippet:

```

1498 #!!!!!!!!!!!!!!topicsort
1499 @app.route("/public/srt/<s>", methods=['GET','POST'])
1500 def thssrtd(s):
1501
1502     mycursor = mysql.connection.cursor()
1503     mycursor.autocommit = True
1504
1505     tkn = word_tokenize(s)
1506
1507     stop_words = set(stopwords.words("english"))
1508
1509     stpd_list=[]
1510
1511     #stop words
1512     for word in tkn:
1513         if word.casefold() not in stop_words:
1514             if '%' != word:
1515                 stpd_list.append(word)
1516
1517     lemm = WordNetLemmatizer()
1518     lmd_words = [lemm.lemmatize(w, get_wordnet_pos(w)) for w in stpd_list]
1519
1520     syn = list()
1521     for synset in wordnet.synsets(s):
1522         for lemma in synset.lemmas():
1523             syn.append(lemma.name())    #add the synonyms to syn
1524
1525     print(str(syn))
1526     res = []
1527
1528     for w in lmd_words:
1529         srch = '%'+w+'%'
1530         mycursor.execute(("SELECT * FROM thesis Where thesis_title\
1531                             LIKE %s \
1532                             ORDER BY id_thesis DESC",(srch,))
1533         x = mycursor.fetchall()
1534         for y in x:
1535             res.append(y)
1536

```

Figure 3: Search Query Code Snippet 1

```

1528 for w in imd.words:
1529     srch = '%'+w+'%'
1530     mycursor.execute("SELECT * FROM thesis Where thesis_title\
1531                       LIKE %s \
1532                       ORDER BY id_thesis DESC",(srch,))
1533     x = mycursor.fetchall()
1534     for y in x:
1535         res.append(y)
1536
1537 if request.method == "POST":
1538     t = request.form.get('tpcsrft')
1539
1540     mycursor.execute("SELECT * FROM topic WHERE topic_name = %s",(t,))
1541     id = mycursor.fetchone()
1542     tpid = id['id_topic']
1543
1544     mycursor.execute("SELECT (id_thesis) FROM thesis_totopics WHERE id_topic = %s",(tpid,))
1545     thid = mycursor.fetchall()
1546
1547     srtd = []
1548
1549     for x in res:
1550         for y in thid:
1551             if y['id_thesis'] == x['id_thesis']:
1552                 srtd.append(x)
1553
1554     fnsrt = []
1555
1556     for x in srtd:
1557         mycursor.execute("SELECT * FROM thesis WHERE id_thesis = %s",(x['id_thesis'],))
1558         srthts = mycursor.fetchall()
1559         for y in srthts:
1560             fnsrt.append(y)
1561
1562     return render_template("public_home_searchdsrtd.html", fnsrt = fnsrt, s=s, t=t, stpd_list=stpd_list)
1563
1564 else:
1565     return redirect(url_for('home'))

```

Figure 4: Search Query Code Snippet 2

Search Recommendations Code Snippet:

```
1579 mycursor.execute("SELECT *, topic.id_topic, topic.topic_name\  
1580                     FROM thesis_totopics\  
1581                     JOIN topic\  
1582                     ON thesis_totopics.id_topic = topic.id_topic\  
1583                     WHERE id_thesis = %s",(id,))  
1584 tpc = mycursor.fetchall()  
1585  
1586  
1587 mycursor.execute("SELECT *, authors.id_author, authors.author_fname, authors.author_lname, authors.course\  
1588                     FROM thesis_toauthors\  
1589                     JOIN authors\  
1590                     ON thesis_toauthors.id_author = authors.id_author\  
1591                     WHERE id_thesis = %s",(id,))  
1592 aut = mycursor.fetchall()  
1593  
1594 thsabs = natlangproc(ths['abstract'])  
1595  
1596 thsabs_syn = syno(thsabs)  
1597  
1598 fn1 = ' '.join(thsabs)  
1599 fn2 = ' '.join(thsabs_syn)  
1600  
1601 fn = fn1+fn2  
1602  
1603 # print(fn1)  
1604  
1605 mycursor.execute('SELECT * FROM thesis')  
1606 thesis = mycursor.fetchall()  
1607  
1608 mycursor.execute("DELETE FROM absrel")  
1609 mysql.connection.commit()  
1610  
1611 for x in thesis:  
1612     tmpabs = natlangproc(x['abstract'])
```

Figure 5: Search Recommendation Code Snippet

```
1611 for x in thesis:  
1612     tmpabs = natlangproc(x['abstract'])  
1613  
1614     tmpabs_syn = syno(tmpabs)  
1615  
1616     fn12 = ' '.join(tmpabs)  
1617     fn2 = ' '.join(tmpabs_syn)  
1618  
1619     tmp = fn1+fn2  
1620  
1621     seq = difflib.SequenceMatcher(None,fn1,fn12)  
1622  
1623     rate = seq.ratio()*100  
1624  
1625     res = str("%.4f").format(rate)+"%"  
1626  
1627     if rate >= 1:  
1628         mycursor.execute("INSERT INTO absrel (id_thesis,rate)\  
1629                             VALUES (%s,%s)",(x['id_thesis'],rate))  
1630         mysql.connection.commit()  
1631  
1632  
1633     # print(res)  
1634     # print(fn12)  
1635  
1636 mycursor.execute("SELECT * FROM absrel ORDER BY rate DESC LIMIT 3")  
1637 rel = mycursor.fetchall()  
1638  
1639 # print(rel)  
1640  
1641 relabs = []  
1642  
1643 for x in rel:  
1644     mycursor.execute("SELECT * FROM thesis WHERE id_thesis = %s",(x['id_thesis'],))  
1645     r = mycursor.fetchone()  
1646     relabs.append(r)  
1647  
1648 # print(relabs)  
1649 return render_template("public_home_openthesis.html", ths = ths, tpc = tpc, aut = aut, relabs = relabs, rel = rel)
```

Figure 6: Search Recommendation Code Snippet

Phases of Development

Planning

The development of this system shall follow the agile method for a rapid development where sprints are tested in iterations until a successful website development is accomplished.

In machine learning machine training, a plan outline of training set, validation set, and test state. This will be closely observed in terms of training the machine learning model.

Analysis

System requirements:

Functional requirement

- ETDs uploading
- Search bar feature for research
- Recommendations from selected thesis
- Admin account
- Faculty Operator's Account
- Plagiarism/Similarity Checker

Non-Functional Requirements

- WMSU color theme (red)
- User interaction feedbacks
- Google Scholar and Microsoft Academics referencing

Design Process

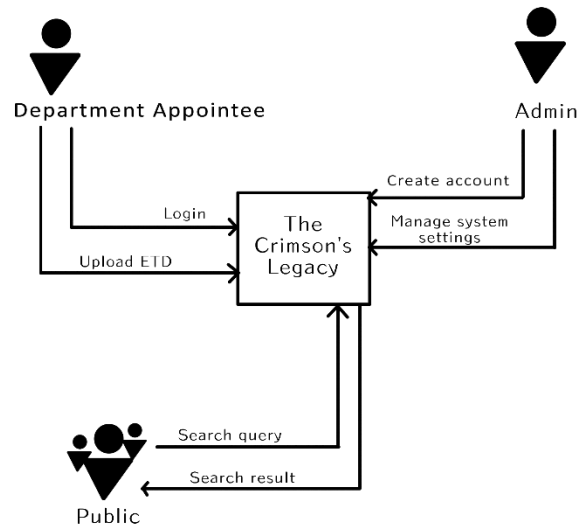


Figure 7. User Case DFD

the researchers' proposed website has 3 users. The thesis students who can register and login for ETD uploading, the admin for system settings and ETD approval, and the public, since the website is open for public who seeks information, everyone with internet access may use the website and search for certain research submitted by WMSU students.

Data

System's ERD

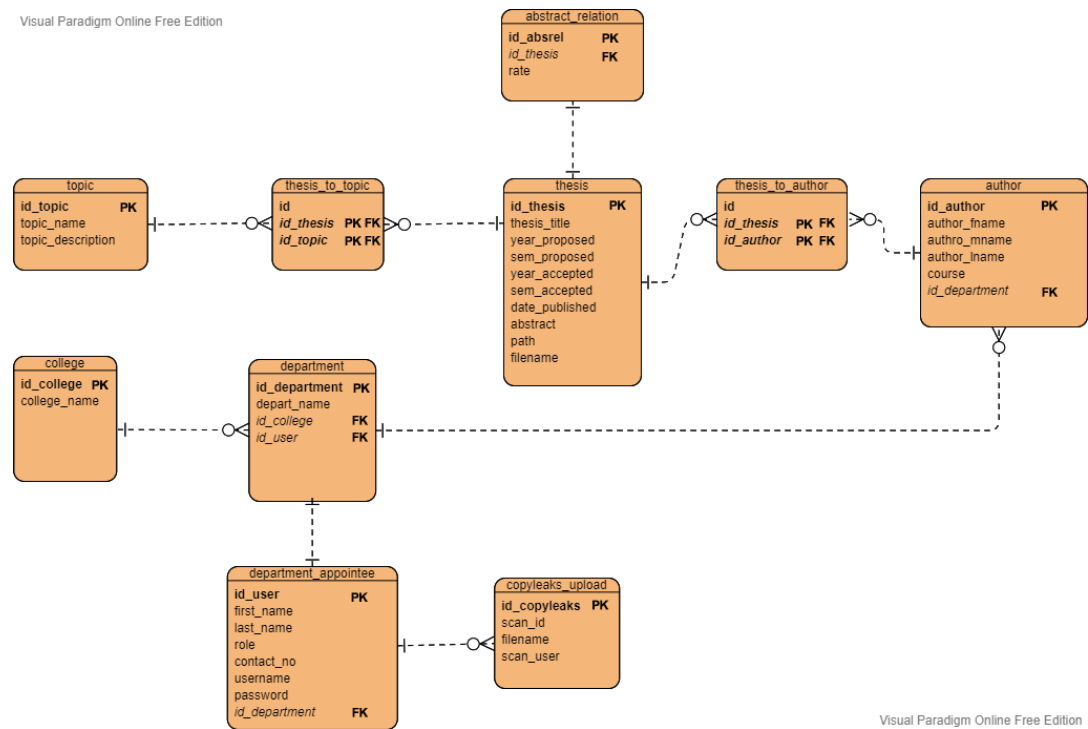


Figure 8. Crimson Legacy's ERD

In this ERD, they can see the relationship of a thesis, uploaded in table thesis, towards multiple tables making it the center of information for the researchers' system.

Logic

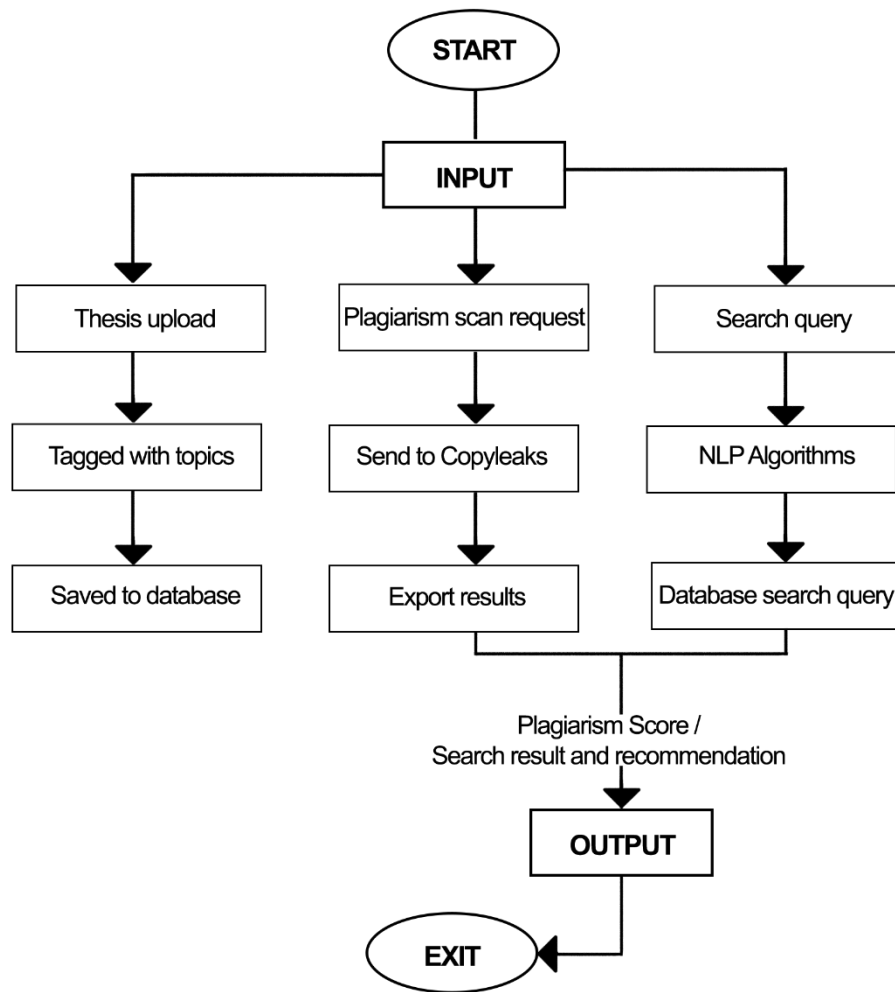


Figure 9. Crimson Legacy's Flowchart

In this flowchart, they can see the overall flow of the researchers' proposed system. It all starts with the input which is either ETD upload, plagiarism scan request, or search query. ETD uploads are tagged with corresponding topics by the uploader then stored into the database. Plagiarism scan request are sent to Copyleaks API and then returns with a plagiarism score. And, search query is broken down by NLP algorithms and then processed by the system's search algorithms which outputs searched thesis and recommendations for clicked thesis.

Implementation

Hardware requirement

-the system will be deployed online and can be accessed with mobile phone and a personal computer which will also be the requirement for the users to manage the website.

Software requirement

-the system is a website and can be accessed from any internet browser. HTML, JS, and CSS will be used to develop the front end and Flask, Python, and MySQL for the back end.

Network requirement

-the system requires internet for use and will be deployed online through the web hosting of Hostinger.

System Architecture

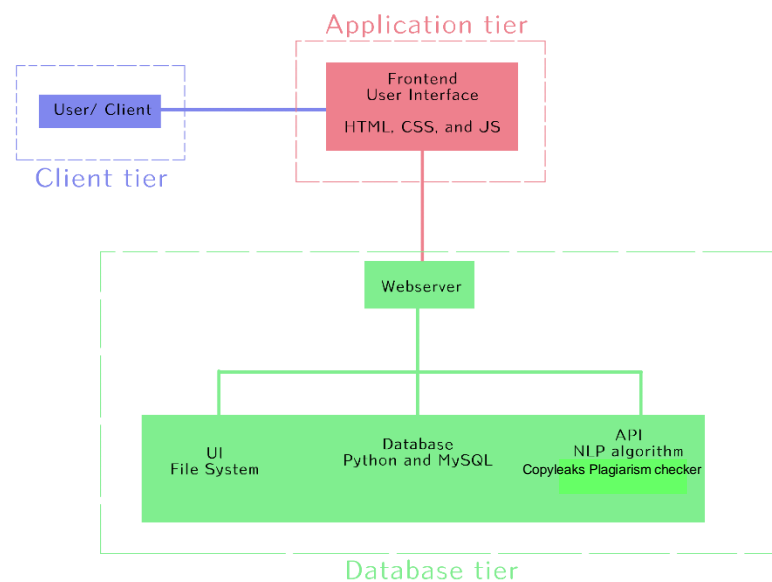


Figure 10. Website Architecture N-tier

The system architecture is an N-tier architecture where the sectors of the proposed system is divided by client tier, application tier, and database tier.

Prototype I

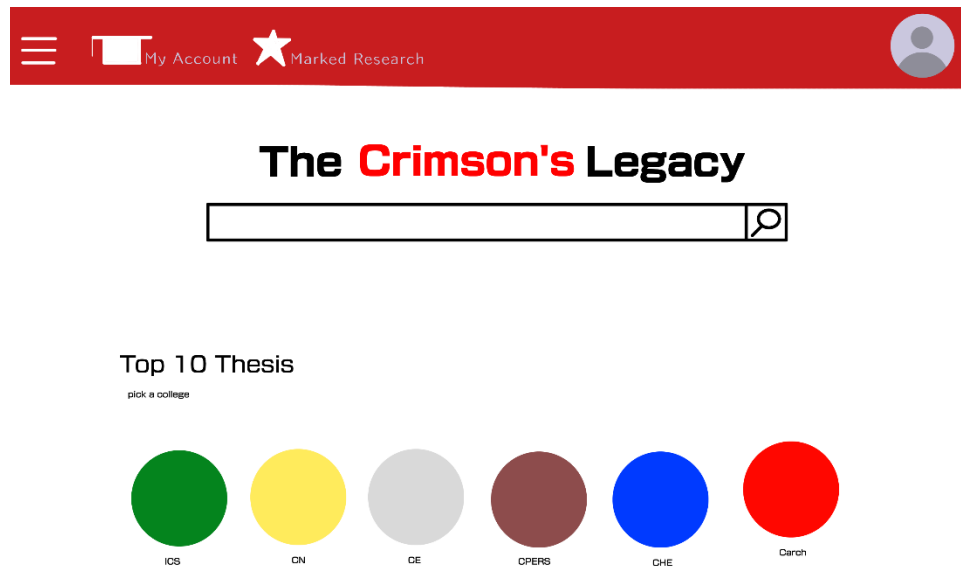


Figure 11. Landing page prototype.

As referenced with Google Scholar UI, this is the prototype of the researchers' system's landing page.

Prototype II

Home page

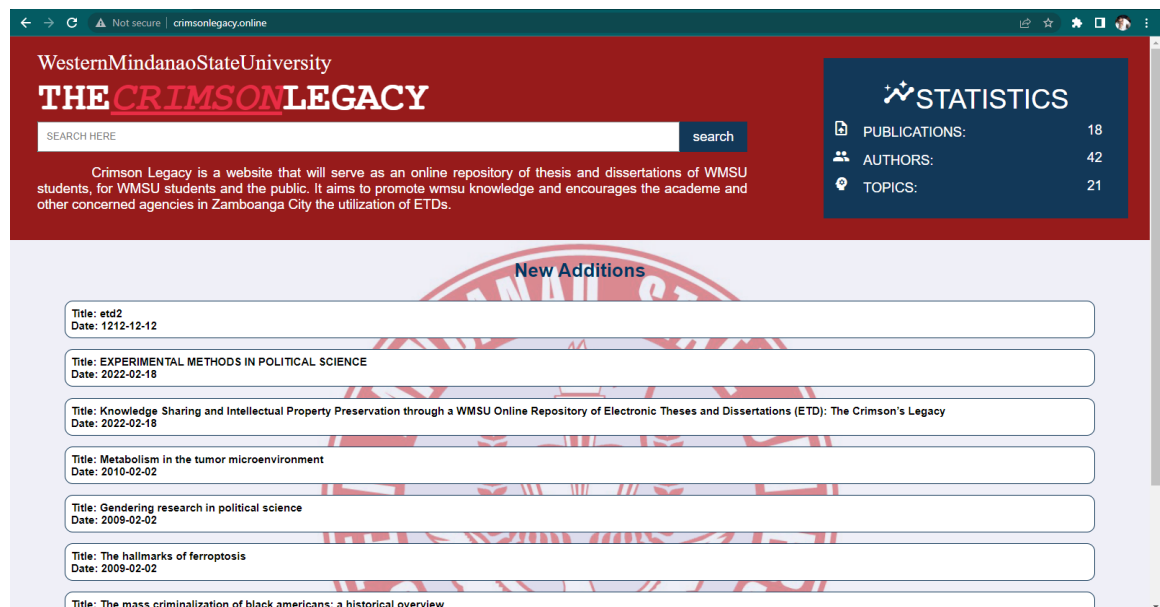


Figure 12: Prototype II Landing Page

Search Result

THE CRIMSON LEGACY

Select a topic to narrow down your search result

Showing results for "micro"

Thesis
Title: Metabolism in the tumor microenvironment
Date: 2010-02-02

Abstract

Experiments in culture systems where one cell type is provided with abundant nutrients and oxygen have been used to inform much of our understanding of cancer metabolism. However, many differences have been observed between the metabolism of tumors and the metabolism of cancer cells grown in monoculture. These differences reflect, at least in part, the presence of nonmalignant cells in the tumor microenvironment and the interactions between those cells and cancer cells. However, less is known about how the metabolism of various tumor stromal cell types differs from that of cancer cells, and how this difference might inform therapeutic targeting of metabolic pathways. Emerging data have identified both cooperative and competitive relationships between different cell types in a tumor, and this review examines how four abundant stromal cell types in the tumor microenvironment, fibroblasts, T cells, macrophages, and endothelial cells, contribute to the metabolism of tumors.

Thesis
Title: lot-based smart plant irrigation system for mustard greens using blynk application and node microcontroller unit
Date: 1997-02-02

Abstract

Irrigation is a major issue in the Philippine agriculture and as we face the pandemic which claimed to be the major adversity that the mass is still dealing right now impacting every aspect of living such as health and economy. People are forced to live a life that are far different from normal with constant change and adaptation. It is a novel setup for all of us that we never realized we could do until we confront a pandemic. This new social custom is now deemed as "new normal" and everyone is mandatory to abide to these alterations. Hence, people are bound to rigorously take care of themselves, towards others and the environment. Apart from that, since the beginning of pandemic, knowing that everyone is lockdown to their respective homes, gardening flourish and became a popular pastime for Filipinos. This paper presents an IoT-based irrigation system that will be utilizing a greenhouse (128cm in length, 60cm in width and 58cm in height) that serves as the controlled environment to our selected plant sample mustard greens which seeks to solve, counteract irrigation problems in residential houses, regulates water distribution and set the quantity of water to be provided to the mustard greens and also to control the temperature, humidity to its optimum condition required for growing a healthy mustard greens. The system was composed of a node microcontroller unit that will serve as the main controller of the irrigation of environmental contact sensor node and noncontact sensor node; contact sensors was made up of soil moisture sensor, temperature and humidity sensors together with the non-contact sensor which made up of ultrasonic sensor. Each node functions as an IoT device. The system will also make use of wireless transmission technologies to serve as an information gateway for the users through push notifications that is linked to the blynk server for transmitting data between blynk mobile application that will carry out as the receiver of the collected data from the sensors; for displaying the monitored environmental parameters, to control each node; and to configuration settings for irrigation. It displays the real-time and percentage data from the soil moisture, temperature and humidity sensor as well as the tank's water depth. The user will be able to manage this device in two ways: remotely or automatically.

Figure 13: Prototype II search result

Clicked Thesis

THE CRIMSON LEGACY

THESIS TITLE:
lot-based smart plant irrigation system for mustard greens using blynk application and node microcontroller unit

AUTHORS:
Ardlee Natividad
BS Electronic Engineering

Clyde Salaritan
BS Electronic Engineering

Raenic Karl Jalao
BS Electronic Engineering

DATES:
Date Proposed/Semester: 1997-02-02/1st
Date Accepted/Semester: 1997-02-02/2nd

ABSTRACT:
Irrigation is a major issue in the Philippine agriculture and as we face the pandemic which claimed to be the major adversity that the mass is still dealing right now impacting every aspect of living such as health and economy. People are forced to live a life that are far different from normal with constant change and adaptation. It is a novel setup for all of us that we never realized we could do until we confront a pandemic. This new social custom is now deemed as "new normal" and everyone is mandatory to abide to these alterations. Hence, people are bound to rigorously take care of themselves, towards others and the environment. Apart from that, since the beginning of pandemic, knowing that everyone is lockdown to their respective homes, gardening flourish and became a popular pastime for Filipinos. This paper presents an IoT-based irrigation system that will be utilizing a greenhouse (128cm in length, 60cm in width and 58cm in height) that serves as the controlled environment to our selected plant sample mustard greens which seeks to solve, counteract irrigation problems in residential houses, regulates water distribution and set the quantity of water to be provided to the mustard greens and also to control the temperature, humidity to its optimum condition required for growing a healthy mustard greens. The system was composed of a node microcontroller unit that will serve as the main controller of the irrigation of environmental contact sensor node and noncontact sensor node; contact sensors was made up of soil moisture sensor, temperature and humidity sensors together with the non-contact sensor which made up of ultrasonic sensor. Each node functions as an IoT device. The system will also make use of wireless transmission technologies to serve as an information gateway for the users through push notifications that is linked to the blynk server for transmitting data between blynk mobile application that will carry out as the receiver of the collected data from the sensors; for displaying the monitored environmental parameters, to control each node; and to configuration settings for irrigation. It displays the real-time and percentage data from the soil moisture, temperature and humidity sensor as well as the tank's water depth. The user will be able to manage this device in two ways: remotely or automatically.

DOWNLOAD PDF:
[Revised_Research_Proposal_Paper.pdf](#)

TOPICS:

Related Abstracts

#1 Title: lot-based smart plant irrigation system for mustard greens using blynk application and node microcontroller unit
Date: 1997-02-02

#2 Title: The symbolism and its meanings in the story of little prince
Date: 2002-02-02

#3 Title: Ideology of capitalist dystopia of Suzanne Collins' 'the hunger games'
Date: 0996-02-02

#1 Rate: 100.0
#2 Rate: 2.97596
#3 Rate: 2.86267

Figure 14: Prototype II clicked thesis

Faculty Login Page

THE CRIMSON LEGACY
User Login

Username

Password

Login

Figure 15: Prototype II login page

Admin Page

THE CRIMSON LEGACY
welcome Admin: Ronald Arcilla

Statistic

Manage Users

Manage College & Department

Manage Thesis

Manage Authors

Manage Topics

Account Settings

Logout

STATISTICS

Number of Appointee	10
Number of Theses	18
Number of Topics	21
Number of Authors	42
Date	2022-04-17
Department without Appointee	1

Figure 16: Prototype II admin page

Department Appointee Page

THE CRIMSON LEGACY
welcome Department Appointee: User1 Bernardino

Manage Thesis

Plagiarism Checker

Logout

THESES

Search here

Title	Proposed Date	Accepted Date	Total Page	Actions
eld2	1212-12-12 2ndSemester	1212-12-12 1stSemester	123	Open
EXPERIMENTAL METHODS IN POLITICAL SCIENCE	2021-02-18 1stSemester	2022-02-18 2ndSemester	33	Open
Knowledge Sharing and Intellectual Property Preservation through a WMSU Online Repository of Electronic Theses and Dissertations (ETD): The Crimson's Legacy	1996-02-18 1stSemester	2022-02-18 2ndSemester	36	Open
Metabolism in the tumor microenvironment	2010-02-02 1stSemester	2010-02-02 2ndSemester	26	Open
Gendering research in political science	2009-02-02 1stSemester	2009-02-02 2ndSemester	25	Open
The hallmarks of ferroptosis	2009-02-02 1stSemester	2009-02-02 2ndSemester	22	Open
The mass criminalization of black americans: a historical overview	2008-02-02 1stSemester	2008-02-02 2ndSemester	29	Open
Integrating viral metagenomics into an ecological framework	2007-02-02 1stSemester	2007-02-02 2ndSemester	29	Open
Nyan nyan go: a basic mathematics android-based mobile game for grade 2 students of zamboanga city	2005-02-02 1stSemester	2005-02-02 2ndSemester	89	Open
Language analysis on bong joon-ho's snowpiercer and parasite in relation to conflict theory and marxism	2004-02-02 1stSemester	2004-02-02 2ndSemester	143	Open

[1](#) (current)
[2](#)
[Next](#)

View Thesis

Add Thesis

Figure 17: Prototype II dept. faculty page

Plagiarism Checker Page

THE CRIMSON LEGACY
welcome Department Appointee: User1 Bernardino

Manage Thesis

Plagiarism Checker

Logout

PLAGIARISM CHECKER
powered by
Copleaks

Search username here

Scan ID	Filename	User	Status	Actions
76345	final_hope.txt	user1	scanning	wow
29318	zampenproj.docx	user1	scanning	wow
67863	final_hope.txt	user1	scanning	wow
29471	final_hope.txt	user1	scanning	wow
97212	Arcilla_BSCS4A_RIZALAssign1.docx	user4	scanning	wow
52765	Arcilla_BSCS4A_RIZALrefpaper.docx	user4	scanning	wow
72348	Arcilla_BSCS4A_RIZALrefpaper.docx	user1	scanning	wow
97747	Arcilla_BSCS4A_RIZALAssign1.docx	user1	scanning	wow

View Scan

Submit Scans

Figure 18: Prototype II plagiarism checker

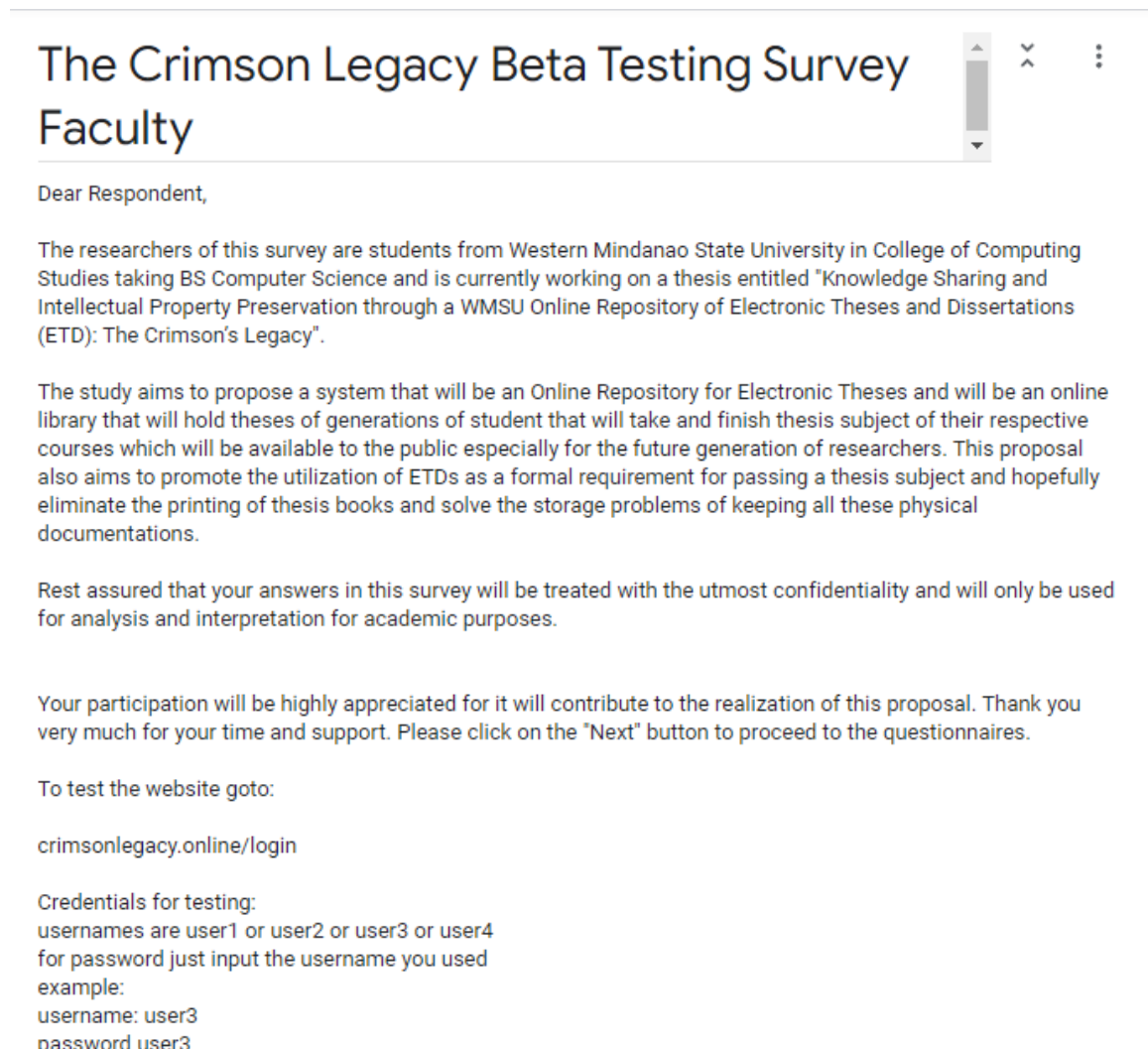
This is the prototype of this website that has been deployed for this project.

CHAPTER IV

RESULTS AND DISCUSSION

After the development process, the researchers' way to test the system is to let the public use the website as it is meant to be. The researchers contacted 5 faculty and 5 students from Western Mindanao State University to test the website. And here are the results:

Faculty Survey



The Crimson Legacy Beta Testing Survey
Faculty

Dear Respondent,

The researchers of this survey are students from Western Mindanao State University in College of Computing Studies taking BS Computer Science and is currently working on a thesis entitled "Knowledge Sharing and Intellectual Property Preservation through a WMSU Online Repository of Electronic Theses and Dissertations (ETD): The Crimson's Legacy".

The study aims to propose a system that will be an Online Repository for Electronic Theses and will be an online library that will hold theses of generations of student that will take and finish thesis subject of their respective courses which will be available to the public especially for the future generation of researchers. This proposal also aims to promote the utilization of ETDs as a formal requirement for passing a thesis subject and hopefully eliminate the printing of thesis books and solve the storage problems of keeping all these physical documentations.

Rest assured that your answers in this survey will be treated with the utmost confidentiality and will only be used for analysis and interpretation for academic purposes.

Your participation will be highly appreciated for it will contribute to the realization of this proposal. Thank you very much for your time and support. Please click on the "Next" button to proceed to the questionnaires.

To test the website goto:

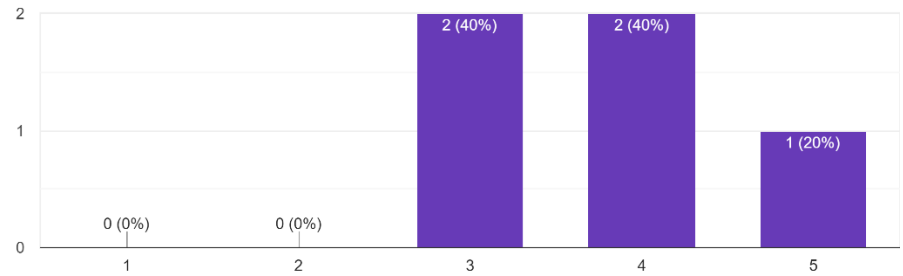
crimsonlegacy.online/login

Credentials for testing:
usernames are user1 or user2 or user3 or user4
for password just input the username you used
example:
username: user3
password user3

Figure 19: Faculty Testing Survey Introduction

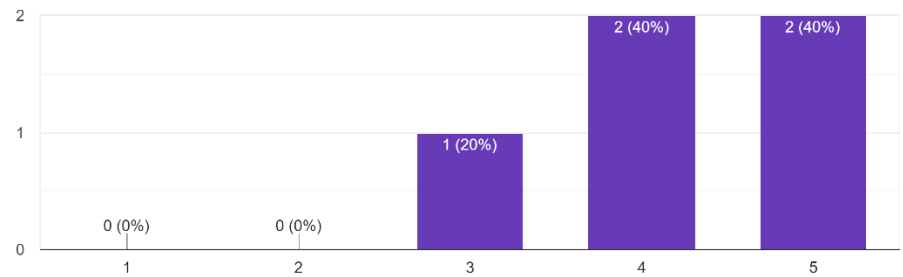
The website looks appealing

5 responses



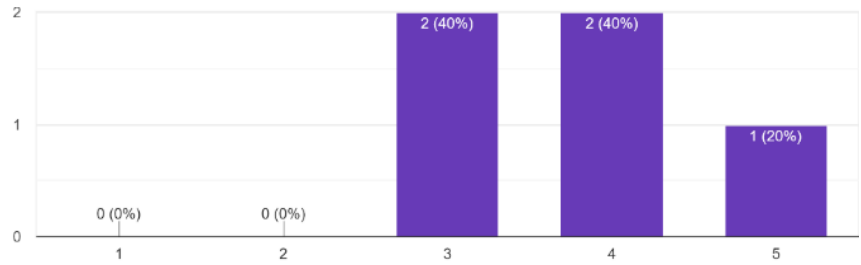
The website is easy to navigate

5 responses

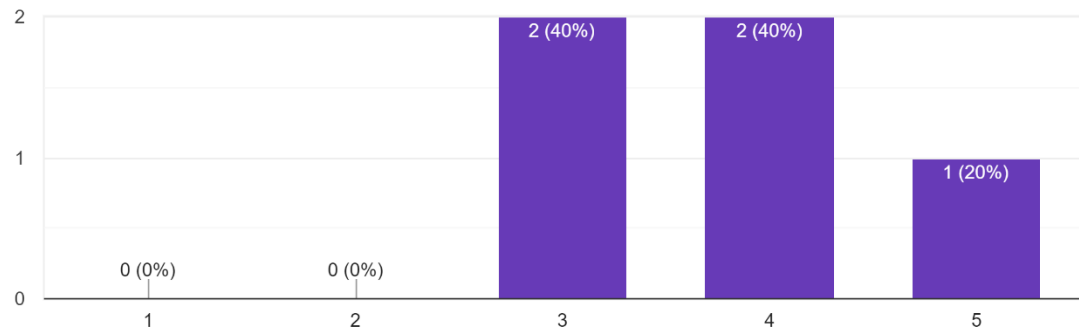


The website is "newbie" friendly

5 responses



The website successfully accomplished the goals of managing and promoting WMSU thesis online
5 responses



How likely would this website help your students in their thesis year?
5 responses

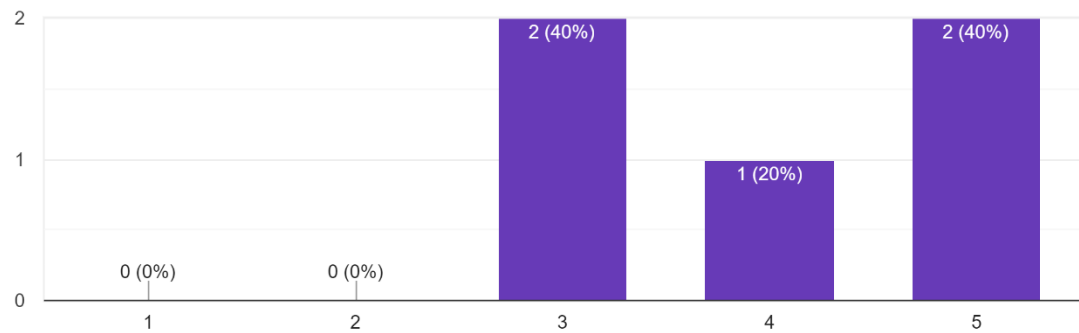


Figure 20: Faculty Testing Survey

Student Survey

The Crimson Legacy Beta Testing Survey

Student

Dear Respondent,

The researchers of this survey are students from Western Mindanao State University in College of Computing Studies taking BS Computer Science and is currently working on a thesis entitled "Knowledge Sharing and Intellectual Property Preservation through a WMSU Online Repository of Electronic Theses and Dissertations (ETD): The Crimson's Legacy".

The study aims to propose a system that will be an Online Repository for Electronic Theses and will be an online library that will hold theses of generations of student that will take and finish thesis subject of their respective courses which will be available to the public especially for the future generation of researchers. This proposal also aims to promote the utilization of ETDs as a formal requirement for passing a thesis subject and hopefully eliminate the printing of thesis books and solve the storage problems of keeping all these physical documentations.

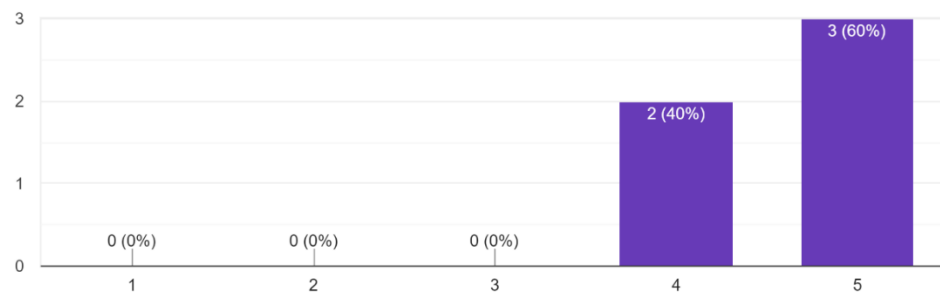
Rest assured that your answers in this survey will be treated with the utmost confidentiality and will only be used for analysis and interpretation for academic purposes.

Your participation will be highly appreciated for it will contribute to the realization of this proposal. Thank you very much for your time and support. Please click on the "Next" button to proceed to the questionnaires.

Figure 21: Student Testing Survey Introduction

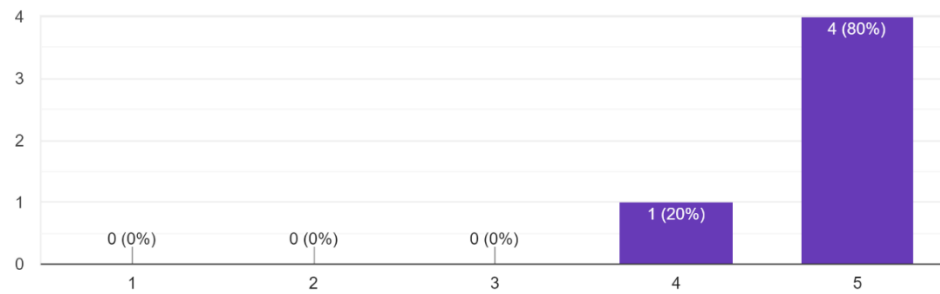
The website looks appealing

5 responses



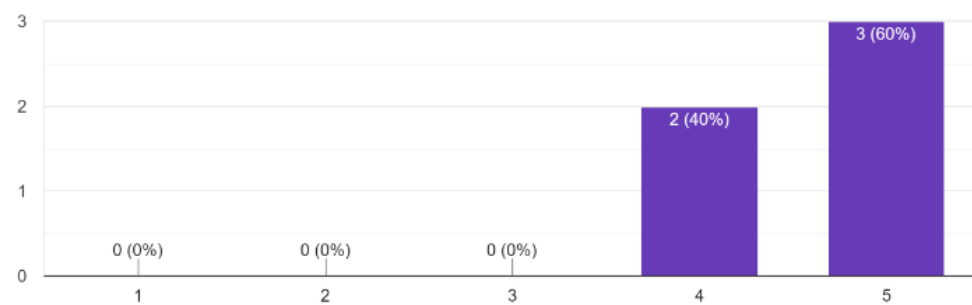
The website is "newbie" friendly

5 responses

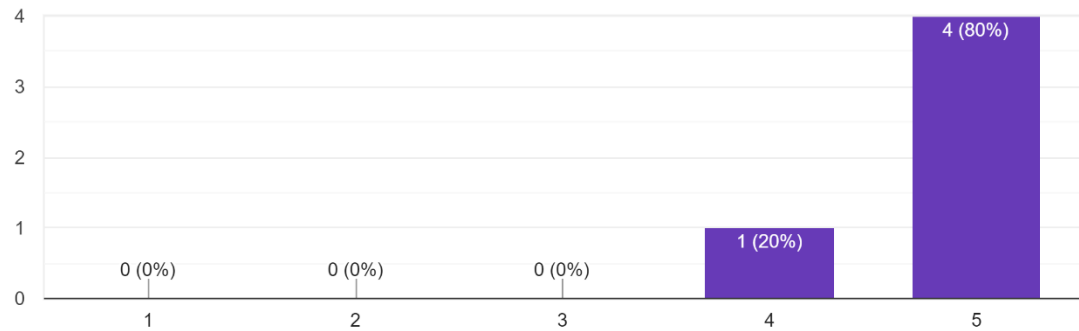


The website is easy to navigate

5 responses



The website successfully accomplished the goals of managing and promoting WMSU thesis online
5 responses



How likely would this website help you in your thesis year?
5 responses

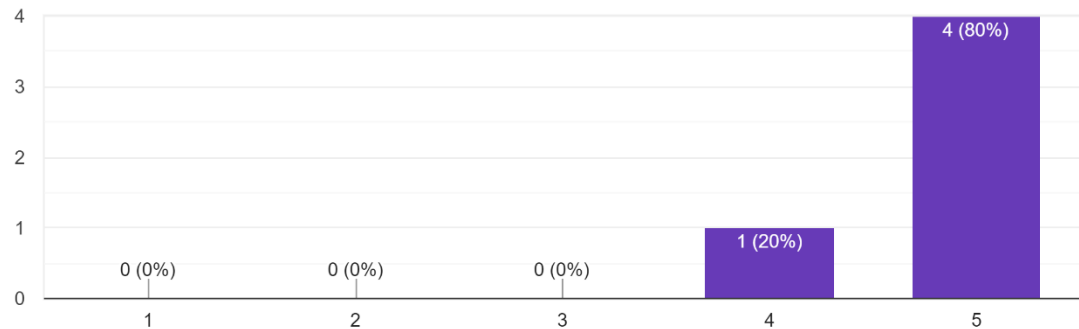


Figure 22: Student Testing Survey

As shown in the survey the website received a mid-passing score from the faculty while they get a high passing score from the students. This shows that the system is working with its functions but can use a little upgrade on functionality and design on the faculty user's pages.

CHAPTER V

CONCLUSION AND RECOMMENDATIONS

Conclusion

The study aims to promote Electronic Theses and Dissertations through creating a website that holds this electronic file for public use and knowledge preservation.

The work has four (4) specific goals and it has all been achieved with a 4-point explanation:

1. The website is fully functional and deployed on the internet with the help of web-hosting company's service, Hostinger.
2. The website's frontend is built with HTML and CSS and is successfully outputting required content and result based from user's query
3. The website's backend is built and manipulated with python and flask and MySQL is used for database functionality where we can store and manage data
4. The website has an increased functionality with the help of Copyleaks. The website is able to do some plagiarism check and give plagiarism score to our faculty users called department appointee.

With these four (4) points. The researchers have accomplished their general objective which is to create a website that will hold as a thesis repository of Western Mindanao State University students which helps in the knowledge preservation and the publication of the Crimson's research.

With Internet being the "information super highway" no information travels as fast as the information uploaded in the internet. May this study start the utilization of Electronic Thesis and Dissertations in our city for as it increases the publicity of our pride studies to reduce the use of paper and to reach other researchers all over the globe. It is easier to develop and study new topics when existing and supporting studies are easy to find.

Recommendations

The proposed system is fully functional and completed based on the objectives of this research. After all the work done in this study, some recommendations are observed for the future researchers. The following recommendations are considered:

- a. The researchers suggest that when no Machine Learning is included in the study, use different backend languages or tools like PHP which is much faster and more suitable in terms of web development.
- b. The researchers suggest to add animation to frontend pages which makes it look less static
- c. The researchers suggest to use CSS frameworks to reduce development time with a better design.
- d. The researchers suggest that if Machine Learning is one of the future researcher's strengths, improvise your own plagiarism checker which may help you innovate ideas and make customize it according to your proposed system
- e. The researchers suggest that if money is not a problem, purchase a higher web-hosting service and APIs to aid on your study since what is used here in this study are as cheap as it can be.

References

NDLTD. (n.d.). Mission, Goals, and History: Our History. Retrieved May 5, 2021, from <http://www.ndltd.org/about>

Federis: Intellectual Property Law. (n.d.). Copyright. Retrieved May 5, 2021, from <http://www.federislaw.com.ph/faqs-resources/copyright/>.

Ciencia. (n.d.). About Ciencia. Retrieved May 7, 2021, from https://wmsu.edu.ph/research_journal/?page=home.

ISSN International Centre. (n.d.). What is an ISSN?. Retrieved May 7, 2021, from <https://www.issn.org/understanding-the-issn/what-is-an-issn/>.

Andrew Oziemblo. (2020, March 4). 10 Website Design Trends for 2020. Forbes. <https://www.forbes.com/sites/forbesagencycouncil/2020/03/04/10-website-design-trends-for-2020/?sh=1e5f0c8c7645>.

Domantas G. (2021, March 09). What is Web Hosting? Web Hosting Explained for Beginners. Hostinger. <https://www.hostinger.ph/tutorials/what-is-web-hosting/>.

Edureka!. (2020, November 29). Stemming and Lemmatization Tutorial | Natural Language Processing (NLP) with Python | Edureka [Video]. https://www.youtube.com/watch?v=p1ccbR2P_xA.

Fineman, Y. (2003). Electronic Theses and Dissertations. Portal: Libraries and the Academy 3(2), 219-227. doi:10.1353/pla.2003.0032. Retrieved May 5, 2021, from <https://muse.jhu.edu/article/42858/summary>.

Karen Hein, & Marc Davis. (2008). The Research Wizard: An innovative Web Application for Patron Service. Taylor & Francis Online. https://doi.org/10.1300/J136v07n01_01.

Mubina Malik, & Trisha Patel. (2016.) Database Security – Attacks and Control Methods. International Journal of Information Sciences and Techniques. https://d1wqtxts1xzle7.cloudfront.net/57896420/dbms_security_2.pdf?1543647875=&res

[ponse-content-](#)

[disposition=inline%3B+filename%3DDATABASE_SECURITY_ATTACKS_AND_CONTR
OL_ME.pdf&Expires=1620673656&Signature=Cg~reEXaqO3rdTnzOKB2CX7iSAa0VRc
wNqU8YZKbe0h83rNfBmmXDzjg9eyvdZmznADlaVlg-
Pkw4McimG3vaoDMe7VEwt8s3VzJ1oiHJ4w98dBQ-
ii4e7CHNxFSYALx4RfJE3SsNztX1Qgy-
8YdLs9NZeyXmyufieMq2~PV2AoJmNDIV1qpl08WW1JCCszvsL14222880LH6QIGDzB6
qca473AJ~NrHFnWzQLnKjOSWkTzqjHRthtfn3JztgLm3no1lxdASNocc91hEyZ5vi65kLd
gn8vTrsMuPT9tNVGERu46ZPHW1ILGZXX3vUbvbrBGUjcl0oxTSU3fkpZ~A_&Key-
Pair-Id=APKAJLOHF5GGSLRBV4ZA](#)

Google. (N.D.). *Google Scholar*. Google Scholar. <https://scholar.google.com/schhp?hl=en>
Microsoft. (N.D.). Microsoft Academics. Microsoft Academics.
<https://academic.microsoft.com/home>

Mohri M., Rostamizadeh A., and Talwalkar A.. (2018). *Foundation of Machine Learning*.
The MIT Press.
[https://books.google.com.ph/books?hl=en&lr=&id=dWB9DwAAQBAJ&oi=fnd&pg=PR5&
dq=machine+learning&ots=AypSTPv1l0&sig=hsmQnabizip5pG9FDdLvseelMNo&redir_e
sc=y#v=onepage&q=machine%20learning&f=false](https://books.google.com.ph/books?hl=en&lr=&id=dWB9DwAAQBAJ&oi=fnd&pg=PR5&dq=machine+learning&ots=AypSTPv1l0&sig=hsmQnabizip5pG9FDdLvseelMNo&redir_e
sc=y#v=onepage&q=machine%20learning&f=false)

Siraj Raval. (2016, December 17). *The Best Way to Prepare Dataset Easily*. [Video].
https://www.youtube.com/watch?v=0xVqLJe9_CY.

Appendix A

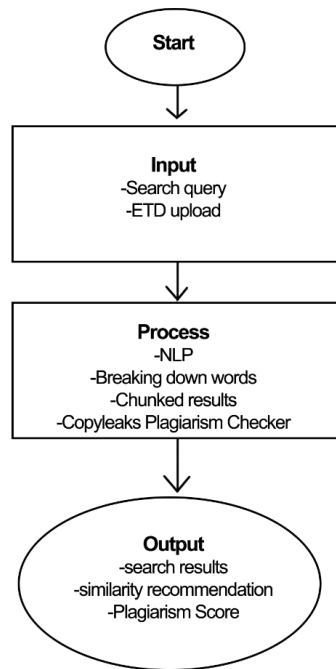
Gantt Chart

GANTT CHART										
Western Mindanao State University - Crimson Legacy Development										
Task	AUG		SEP		OCT		NOV		DEC	
Local Frontend										
Local Backend										
Online Deployment										
Add API to system										
Theis paper documentation										

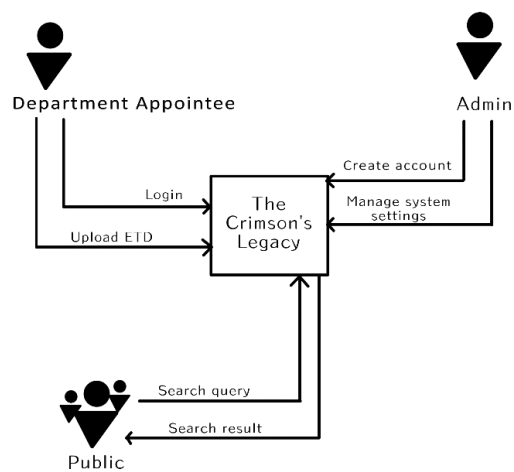
Appendix B

Flowchart/Diagrams

Search Query Flowchart

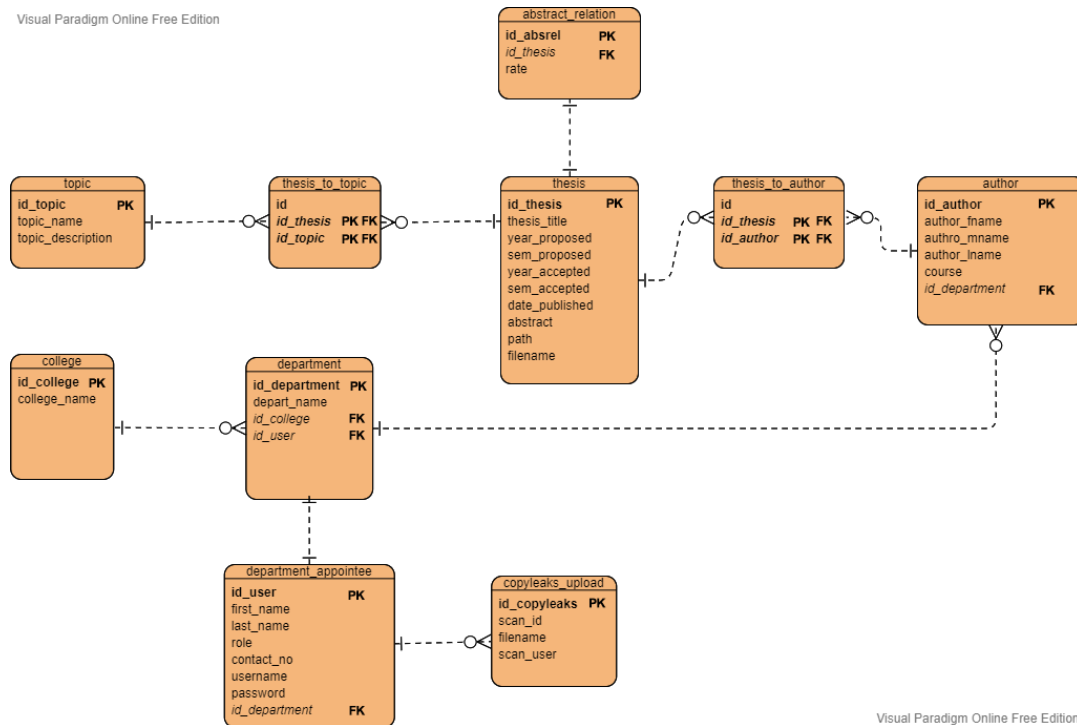


Crimson Legacy Process



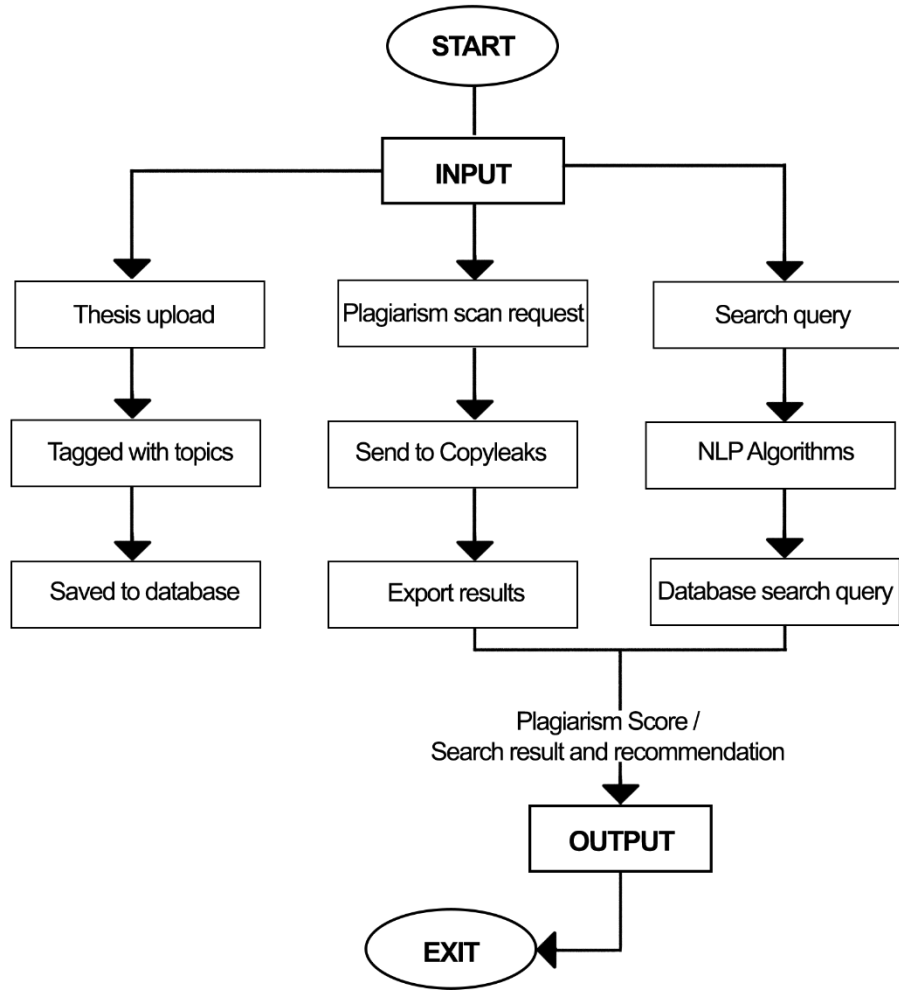
Crimson Legacy ERD

Visual Paradigm Online Free Edition

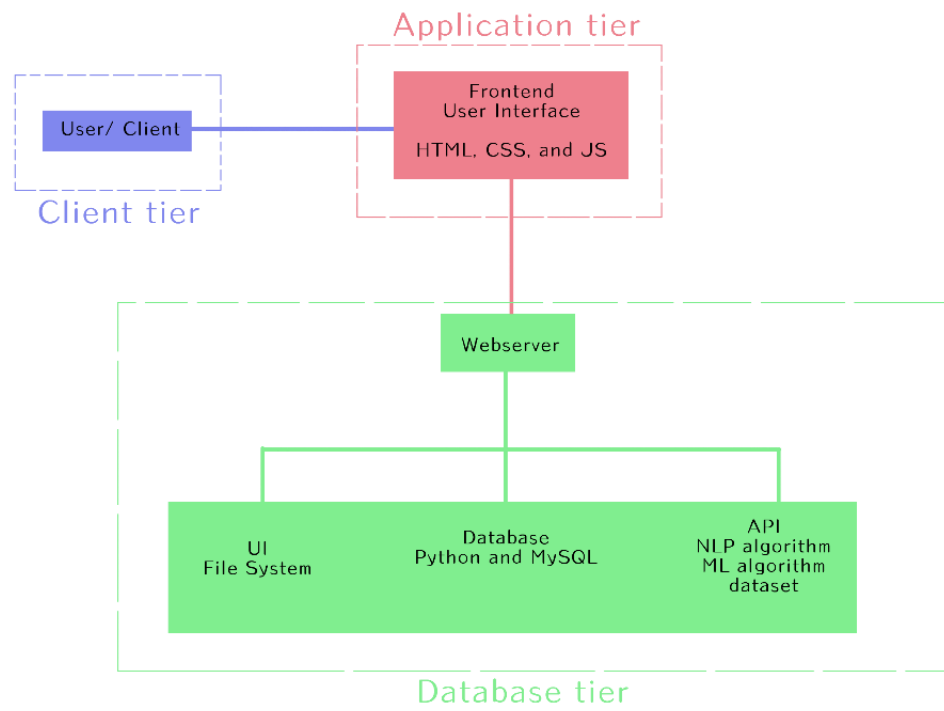


Visual Paradigm Online Free Edition

Crimson Legacy Flowchart



Website Architecture



Appendix C

Algorithms

Search Queries Code Snippet:

```

1498 #!!!!!!!!!!!!!!!!!!!!!!topicsor:]
1499 @app.route("/public/srt/<s>", methods=['GET','POST'])
1500 def thssrtd(s):
1501
1502     mycursor = mysql.connection.cursor()
1503     mycursor.autocommit = True
1504
1505     tkn = word_tokenize(s)
1506
1507     stop_words = set(stopwords.words("english"))
1508
1509     stpd_list=[]
1510
1511     #stop words
1512     for word in tkn:
1513         if word.casefold() not in stop_words:
1514             if '%' != word:
1515                 stpd_list.append(word)
1516
1517     lemm = WordNetLemmatizer()
1518     lmd_words = [lemm.lemmatize(w, get_wordnet_pos(w)) for w in stpd_list]
1519
1520     syn = list()
1521     for synset in wordnet.synsets(s):
1522         for lemma in synset.lemmas():
1523             syn.append(lemma.name())    #add the synonyms to syn
1524
1525     print(str(syn))
1526     res = []
1527
1528     for w in lmd_words:
1529         srch = '%'+w+'%'
1530         mycursor.execute("SELECT * FROM thesis Where thesis_title\
1531                             LIKE %s \
1532                             ORDER BY id_thesis DESC",(srch,))
1533         x = mycursor.fetchall()
1534         for y in x:
1535             res.append(y)
1536
1537
1538     for w in lmd_words:
1539         srch = '%'+w+'%'
1540         mycursor.execute("SELECT * FROM thesis Where thesis_title\
1541                             LIKE %s \
1542                             ORDER BY id_thesis DESC",(srch,))
1543         x = mycursor.fetchall()
1544         for y in x:
1545             res.append(y)
1546
1547     if request.method == "POST":
1548         t = request.form.get('tpcsrt')
1549
1550         mycursor.execute("SELECT * FROM topic WHERE topic_name = %s",(t,))
1551         id = mycursor.fetchone()
1552         tpid = id['id_topic']
1553
1554         mycursor.execute("SELECT (id_thesis) FROM thesis_totopics WHERE id_topic = %s",(tpid,))
1555         thid = mycursor.fetchall()
1556
1557         srted = []
1558
1559         for x in res:
1560             for y in thid:
1561                 if y['id_thesis'] == x['id_thesis']:
1562                     srted.append(x)
1563
1564         fnsrt = []
1565
1566         for x in srted:
1567             mycursor.execute("SELECT * FROM thesis WHERE id_thesis = %s",(x['id_thesis'],))
1568             srthts = mycursor.fetchall()
1569             for y in srthts:
1570                 fnsrt.append(y)
1571
1572     return render_template("public_home_searchedsrt.html", fnsrt = fnsrt, s=s, t=t, stpd_list=stpd_list)
1573
1574 else:
1575     return redirect(url_for('home'))
1576

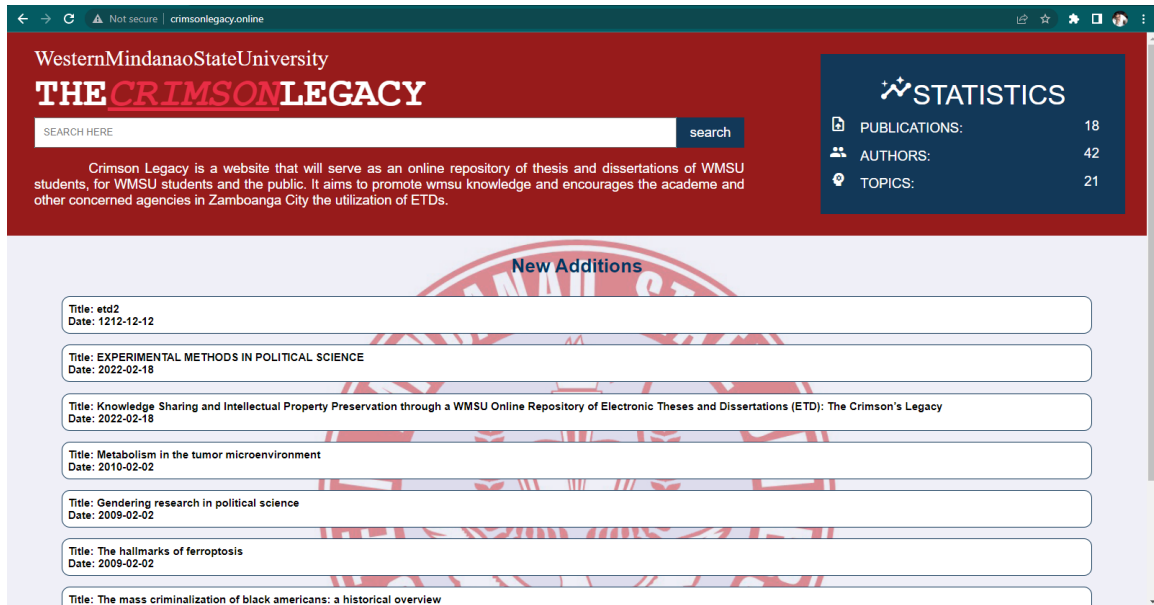
```

Search Recommendations Code Snippet:

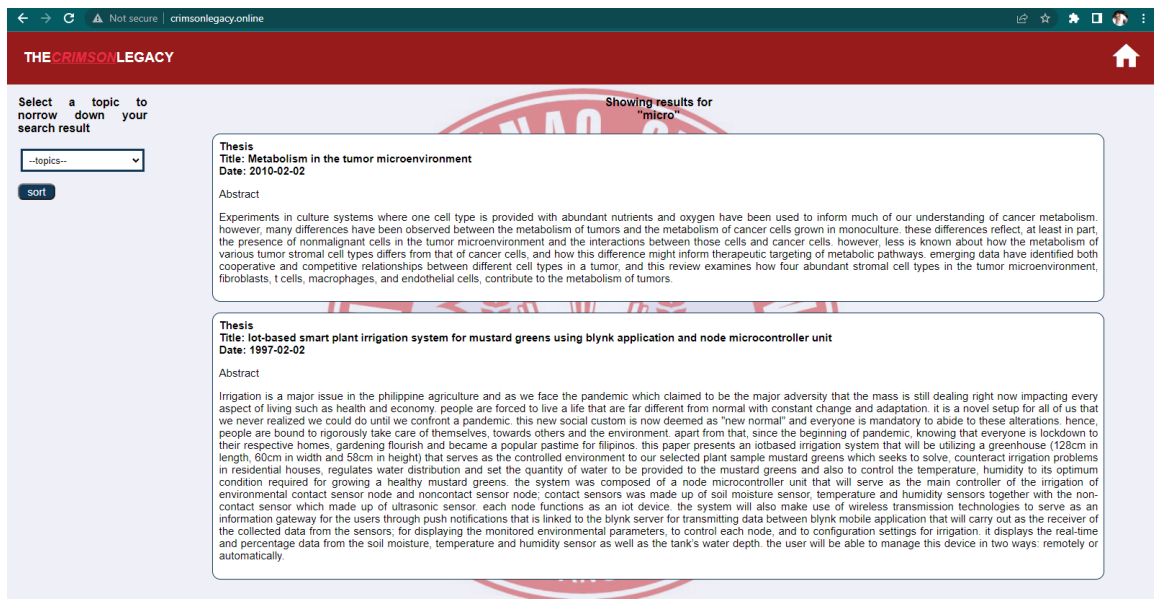
```
1579 mycursor.execute("SELECT *, topic.id_topic, topic.topic_name\  
1580                     FROM thesis_totopics\  
1581                     JOIN topic\  
1582                     ON thesis_totopics.id_topic = topic.id_topic\  
1583                     WHERE id_thesis = %s",(id,))  
1584 tpc = mycursor.fetchall()  
1585  
1586  
1587 mycursor.execute("SELECT *, authors.id_author, authors.author_fname, authors.author_lname, authors.course\  
1588                     FROM thesis_toauthors\  
1589                     JOIN authors\  
1590                     ON thesis_toauthors.id_author = authors.id_author\  
1591                     WHERE id_thesis = %s",(id,))  
1592 aut = mycursor.fetchall()  
1593  
1594 thsabs = natlangproc(ths['abstract'])  
1595  
1596 thsabs_syn = syno(thsabs)  
1597  
1598 fn1 = ' '.join(thsabs)  
1599 fn2 = ' '.join(thsabs_syn)  
1600  
1601 fn = fn1+fn2  
1602  
1603 # print(fn1)  
1604  
1605 mycursor.execute('SELECT * FROM thesis')  
1606 thesis = mycursor.fetchall()  
1607  
1608 mycursor.execute("DELETE FROM absrel")  
1609 mysql.connection.commit()  
1610  
1611 for x in thesis:  
1612     tmpabs = natlangproc(x['abstract'])  
  
1613     tmpabs_syn = syno(tmpabs)  
1614  
1615     fn12 = ' '.join(tmpabs)  
1616     fn2 = ' '.join(tmpabs_syn)  
1617  
1618     tmp = fn1+fn2  
1619  
1620     seq = difflib.SequenceMatcher(None,fn1,fn12)  
1621  
1622     rate = seq.ratio()*100  
1623  
1624     res = str("{:.4f}".format(rate))+ "%"   
1625  
1626     if rate >= 1:  
1627         mycursor.execute("INSERT INTO absrel (id_thesis,rate)\  
1628                             VALUES (%s,%s)",(x['id_thesis'],rate))  
1629         mysql.connection.commit()  
1630  
1631         # print(res)  
1632         # print(fn12)  
1633  
1634  
1635  
1636 mycursor.execute("SELECT * FROM absrel ORDER BY rate DESC LIMIT 3")  
1637 rel = mycursor.fetchall()  
1638  
1639 # print(rel)  
1640  
1641 relabs = []  
1642  
1643 for x in rel:  
1644     mycursor.execute("SELECT * FROM thesis WHERE id_thesis = %s",(x['id_thesis'],))  
1645     r = mycursor.fetchone()  
1646     relabs.append(r)  
1647  
1648 # print(relabs)  
1649 return render_template("public_home_openthesis.html", ths = ths, tpc = tpc, aut = aut, relabs = relabs, rel = rel)
```

Appendix D

Screenshot/Picture of the System Home page



Search Result



Clicked Thesis

THE **CRIMSON** LEGACY

HOME

THESIS TITLE:
Iot-based smart plant irrigation system for mustard greens using blynk application and node microcontroller unit

AUTHORS:
Ardlee Natividad
BS Electronic Engineering

Clyde Salaritan
BS Electronic Engineering

Raeric Karl Jalao
BS Electronic Engineering

DATES:
Date Proposed/Semester: 1997-02-02/1st
Date Accepted/Semester: 1997-02-02/2nd

ABSTRACT:
Irrigation is a major issue in the philippine agriculture and as we face the pandemic which claimed to be the major adversity that the mass is still dealing right now impacting every aspect of living such as health and economy. people are forced to live a life that are far different from normal with constant change and adaptation. it is a novel setup for all of us that we never realized we could do until we confront a pandemic. this new social custom is now deemed as "new normal" and everyone is mandatory to abide to these alterations. hence, people are bound to rigorously take care of themselves, towards others and the environment. apart from that, since the beginning of pandemic, knowing that everyone is lockdown to their respective homes, gardening flourish and became a popular pastime for filipinos. this paper presents an iot-based irrigation system that will be utilizing a greenhouse (128cm in length, 60cm in width and 58cm in height) that serves as the controlled environment to our selected plant sample mustard greens which seeks to solve, counteract irrigation problems in residential houses, regulates water distribution and set the quantity of water to be provided to the mustard greens and also to control the temperature, humidity to its optimum condition required for growing a healthy mustard greens. the system was composed of a node microcontroller unit that will serve as the main controller of the irrigation of environmental contact sensor node and noncontact sensor node, contact sensors was made up of soil moisture sensor, temperature and humidity sensors together with the non-contact sensor which made up of ultrasonic sensor. each node functions as an iot device. the system will also make use of wireless transmission technologies to serve as an information gateway for the users through push notifications that is linked to the blynk server for transmitting data between blynk mobile application that will carry out as the receiver of the collected data from the sensors; for displaying the monitored environmental parameters, to control each node, and to configuration settings for irrigation. it displays the real-time and percentage data from the soil moisture, temperature and humidity sensor as well as the tank's water depth. the user will be able to manage this device in two ways: remotely or automatically.

DOWNLOAD PDF:
[Revised-Research-Proposal-Paper.pdf](#)

TOPICS:

Related Abstracts

#1 Title: Iot-based smart plant irrigation system for mustard greens using blynk application and node microcontroller unit
Date: 1997-02-02

#2 Title: The symbolism and its meanings in the story of little prince
Date: 2002-02-02

#3 Title: Ideology of capitalist dystopia of suzanne collins' the hunger games
Date: 0996-02-02

#1 Rate: 100.0
#2 Rate: 2.97596
#3 Rate: 2.86267

Faculty Login Page

← → 🔒 Not secure | crimsonlegacy.online/login

THE **CRIMSON** LEGACY

User Login

Login

Admin Page

THE CRIMSON LEGACY
welcome Admin: Ronald Arcilla

STATISTICS

Number of Appointee	10
Number of Theses	18
Number of Topics:	21
Number of Authors	42
Date	2022-04-17
Department without Appointee	1

Navigation Links:

- Statistic
- Manage Users
- Manage College & Department
- Manage Thesis
- Manage Authors
- Manage Topics
- Account Settings
- Logout

Department Appointee Page

THE CRIMSON LEGACY
welcome Department Appointee: User1 Bernardino

Navigation Links:

- Manage Thesis
- Plagiarism Checker
- Logout

THESES

Title	Proposed Date	Accepted Date	Total Page	Actions
etd2	1212-12-12 2ndSemester	1212-12-12 1stSemester	123	Open
EXPERIMENTAL METHODS IN POLITICAL SCIENCE	2021-02-18 1stSemester	2022-02-18 2ndSemester	33	Open
Knowledge Sharing and Intellectual Property Preservation through a WMSU Online Repository of Electronic Theses and Dissertations (ETD): The Crimson's Legacy	1996-02-18 1stSemester	2022-02-18 2ndSemester	36	Open
Metabolism in the tumor microenvironment	2010-02-02 1stSemester	2010-02-02 2ndSemester	26	Open
Gendering research in political science	2009-02-02 1stSemester	2009-02-02 2ndSemester	25	Open
The hallmarks of ferroptosis	2009-02-02 1stSemester	2009-02-02 2ndSemester	22	Open
The mass criminalization of black americans: a historical overview	2008-02-02 1stSemester	2008-02-02 2ndSemester	29	Open
Integrating viral metagenomics into an ecological framework	2007-02-02 1stSemester	2007-02-02 2ndSemester	29	Open
Nyan nyan go: a basic mathematics android-based mobile game for grade 2 students of zamboanga city	2005-02-02 1stSemester	2005-02-02 2ndSemester	89	Open
Language analysis on bong joon-ho's snowpiercer and parasite in relation to conflict theory and marxism	2004-02-02 1stSemester	2004-02-02 2ndSemester	143	Open

Actions:

- [View Thesis](#)
- [Add Thesis](#)

Page Navigation:

- 1 (current)
- 2
- [Next](#)

Plagiarism Checker Page

← → ↻

Not secure | crimsonlegacy.online/copyleakswait

🔍 ⚙️ 🏠 🌐

THE **CRIMSON** LEGACY

welcome Department Appointee: User1 Bernardino

Manage Thesis

Plagiarism Checker

Logout

PLAGIARISM CHECKER

powered by

Copyleaks

Search username here search

Scan ID	Filename	User	Status	Actions
76345	final_hope.txt	user1	scanning	<div>wow</div>
29318	zampenproj.docx	user1	scanning	<div>wow</div>
67863	final_hope.txt	user1	scanning	<div>wow</div>
29471	final_hope.txt	user1	scanning	<div>wow</div>
97212	Arcilla_BSCS4A_RIZALAssign1.docx	user4	scanning	<div>wow</div>
52765	Arcilla_BSCS4A_RIZALrefpaper.docx	user4	scanning	<div>wow</div>
72348	Arcilla_BSCS4A_RIZALrefpaper.docx	user1	scanning	<div>wow</div>
97747	Arcilla_BSCS4A_RIZALAssign1.docx	user1	scanning	<div>wow</div>

View Scan

Submit Scans

Appendix E

Test Cases and Results

Project Name:	Western Mindanao State University - The Crimson Legacy								
Module Name:	Plagiarism Check								
Created by:	Ronald Arcilla								
Test Case ID	Test Scenario	Test Case Description	Steps	Pre Conditions	Test Data	Post Conditions	Expected Result	Actual Result	Status
TC_PC_001	Plagiarism Check	user uploads thesis for scan	1. Visit crimsonlegacy.online/login 2. login account 3. goto plagiarism tab 4.press submit scans 5. choose thesis to upload 6. press submit button	User logged in	1. User Account 2. Thesis file	redirect to user page, scanned thesis output in table	alert:"Scan successful, please wait for the result"	same as expected	P
TC_PC_002	Plagiarism Check	user does not upload thesis for scan and presses submit	1. Visit crimsonlegacy.online/login 2. login account 3. goto plagiarism tab 4.press submit scans 5. choose thesis to upload 6. press submit button	User logged in	1. User Account 2. Thesis file	redirect to user page, scanned thesis output in table	alert: "No thesis uploaded"	same a s expected	P

Project Name:	Western Mindanao State University - The Crimson Legacy								
Module Name:	Plagiarism Check								
Created by:	Ronald Arcilla								
Test Case ID	Test Scenario	Test Case Description	Steps	Pre Conditions	Test Data	Post Conditions	Expected Result	Actual Result	Status
TC_PC_001	Plagiarism Check	user uploads thesis for scan	1. Visit crimsonlegacy.online/login 2. login account 3. goto plagiarism tab 4.press submit scans 5. choose thesis to upload 6. press submit button	User logged in	1. User Account 2. Thesis file	redirect to user page, scanned thesis output in table	alert:"Scan successful, please wait for the result"	same as expected	P
TC_PC_002	Plagiarism Check	user does not upload thesis for scan and presses submit	1. Visit crimsonlegacy.online/login 2. login account 3. goto plagiarism tab 4.press submit scans 5. choose thesis to upload 6. press submit button	User logged in	1. User Account 2. Thesis file	redirect to user page, scanned thesis output in table	alert: "No thesis uploaded"	same a s expected	P

Project Name:	Western Mindanao State University - The Crimson Legacy								
Module Name:	Login Module								
Created by:	Ronald Arcilla								
Test Case ID	Test Scenario	Test Case Description	Steps	Pre Conditions	Test Data	Post Conditions	Expected Result	Actual Result	Status
TC_login_001	User logging in	user inputs correct credential in the login fields	1. Visit crimsonlegacy.online/login 2. Input credentials 3. Press login button	User registered	Users on database	redirect to user page	redirect to user page	same as expected	P
TC_login_002	User logging in	user inputs wrong credential in the login fields	1. Visit crimsonlegacy.online/login 2. Input credentials 3. Press login button	User registered	Users on database	redirect to user page	alert: "Wrong username or password"	same a s expected	P
TC_login_003	User logging in	user doesn't input any credentials and presses login button	1. Visit crimsonlegacy.online/login 2. Input credentials 3. Press login button	User registered	Users on database	redirect to user page	alert: "Login fields empty"	same a s expected	P

Project Name:	Western Mindanao State University - The Crimson Legacy								
Module Name:	Public module								
Created by:	Ronald Arcilla								
Test Case ID	Test Scenario	Test Case Description	Steps	Pre Conditions	Test Data	Post Conditions	Expected Result	Actual Result	Status
TC_Pb_001	User inputs search query	The user presses the search button without an input	1. Visit crimsonlegacy.online. 2. Input search query. 3. Press search button	None	Thesis on database	search results and recommendation outputs	Alert : search field empty	same as expected	P
TC_Pb_002	User inputs search query	The user inputs search A and presses enter	1. Visit crimsonlegacy.online. 2. Input search query. 3. Press search button	None	Thesis on database	search results and recommendation outputs	search results and recommendation outputs	same a s expected	P

Appendix F

Curriculum Vitae

Ronald M. Arcilla
09567194157
arcillarnldmcrhn@gmail.com
BAT Compound, Governor Camins, Canelar, Zamboanga City

I. PERSONAL PROFILE



A creative Bachelor of Science in Computer Science Graduate of Western Mindanao State University who is always eager to enhance his skills or learn new ones. Aims to help people solve their problem through utilization of art and technology.

II. PERSONAL DETAILS

Gender: Male
Date of Birth: February 18, 1996
Nationality: Filipino
Marital Status: Single
Present Address: BAT Compound, Governor Camins, Canelar, Zamboanga City

III. RELATED EXPERIENCE

Udemy Online Course:
Front End Web Development Ultimate Course 2021
Instructor: Learn Tech Plus
Cert no: UC-499a870b-2dba-48bc-b9e7-6ed61014fb3b
Programming Commissions:
-Python programming with Tkinter GUI

IV. EDUCATIONAL BACKGROUND

COLLEGE
Bachelor of Science in Computer Science
Western Mindanao State University
2021 - 2022
HIGHSCHOOL
Our Lady of the Abandoned Catholic Schools
Muntinlupa City
2008-2013
ELEMENTARY
Our Lady of the Abandoned Catholic School
Muntinlupa City
2003 - 2008

V. SKILLS

TECHNICAL SKILLS
a. Platform
Windows 10
b. Other Tools/Software
Proficient in Adobe Photoshop
Microsoft Family
Proficient in Davinci Resolve
c. Expertise
Fullstack Web Development
-Frontend(HTML, JS, & CSS)
-Backend(Python, Flask, PHP, sqlite3, MySQL)
Digital Art

SKILLS WITH DATA

a. Documentation
Proficient in MS Word

SKILLS WITH PEOPLE

a. Interpersonal Communication
Clear Communicator
Fluent in English and Tagalog
b. Project Management
Time Management
Critical Decision Making
Capable of managing a Team
Equipped with leadership skills

VI. CLUBS/ORGANIZATION

The University Digest
Western Mindanao State University
Senior Cartoonist
2020 - 2022

VIII. INTEREST/HOBBIES

ART
-Sketching
-Painting
-Digital Art
-Character Creation
-Concept Art
-Illustration
-Photoshop (Photo Manipulation)

Emmanuel L. Toledo

(+63) 927 203 9897

toledoemoem23@gmail.com

Salcedo Drive, Lustre st., Brgy. Sta. Catalina, Zamboanga City, Philippines

I. PERSONAL PROFILE



A hardworking, organized, resourceful Bachelor of Science in Computer Science graduated in Western Mindanao State University with experience working across software development, equipped with programming skill and computer ethnics.

II. PERSONAL DETAILS

Gender: Male
Date of Birth: May 23, 1999
Place of Birth: Zamboanga City, Philippines
Nationality: Filipino
Marital Status: Single
Present Address: Salcedo Drive, Lustre st. Brgy. Sta. Catalina, Zamboanga City, Philippines

I. RELATED EXPERIENCE

- Personal Assistant at GMA Networks
- Summer job at Dunkin Donut

II. EDUCATIONAL BACKGROUND

TERTIARY

Bachelor of Science in Computer Science
Western Mindanao State University
August 2018 - June 2022

SECONDARY

Senior High School
STEM - Engineering
Western Mindanao State University
With Honor (90%)
June 2016 - March 2018

III. SKILLS .

TECHNICAL SKILLS

- Proficient in the following Programming Languages including C++, Java, HTML,

SKILLS WITH DATA

- Proficient with MS Office Suite (Word, Powerpoint, Excel)

SKILLS WITH PEOPLE

- Able to express ideas and thoughts both comprehensively and concisely.

LANGUAGE

- English
- Filipino
- Chavacano
- Tausug

IV. CERTIFICATIONS

- Front end web development Bootcamp 2021
- The Complete Introduction to C++ Programming
- Java Collections Framework + Generics, Lambdas & Stream API

V. INTEREST/HOBBIES

- Programming
- Gaming
- Animation
- Photography