**ABSTRACT**

College teachers may find it difficult to manage the wide range of research resources, such as papers, patents, publications, and journals, in the quickly changing academic environment. Using the power of contemporary technologies, this abstract proposes a novel way to automate the management of these research resources. Our approach makes use of a number of technologies, such as HTML, CSS, JavaScript, ReactJS, MongoDB, and ExpressJS, to expedite the organisation, retrieval, and finding of material.

Our programme is made to provide researchers and college instructors with quick access to research papers that are most pertinent to their chosen topic of study. The ease with which users can use keywords to find relevant content greatly improves the effectiveness of the research process. The days of searching through mountains of paperwork or having trouble finding a certain document in a jumbled digital environment are long gone.

The integration of machine learning (ML) and natural language processing (NLP) algorithms forms the core of our system. Research materials can be retrieved from a database using these state-of-the-art technologies in response to the user's search queries. Our system is able to identify and rank the resources that most closely correspond to the user's demands since it comprehends the context and content of every document.

To sum up, research material management automation has the potential to completely change how academic resources are accessed and used by college professors and researchers. Our solution leverages the capabilities of HTML, CSS, JavaScript, ReactJS, MongoDB, ExpressJS, NLP, and ML to improve the efficiency and effectiveness of the research process. An overview of the bright future of research material management, where technology streamlines and expedites knowledge acquisition, is given in this abstract.

**Table of Contents**

|  |  |
| --- | --- |
| *Acknowledgements* | *i* |
| *Abstract* | *iii* |
| *Abbrevations* | *v* |
| *List of Figures* | *vi* |
| *List of Tables* | *vii* |
|  |  |
| Chapter-1: Introduction | 1 |
| 1.1 Definition | 3 |
| 1.2 Scope of the Work | 4 |
| 1.3 Existing Methods | 9 |
| 1.4 Objective  1.5 Project Introduction | 13 |
| 1.6 Thesis Organization | 16 |
| Chapter-2: Literature Survey | 17 |
| 2.1 Classification of Models in the Literatures |  |
| 2.2 Studies using approach-1 |  |
| 2.3 Studies using approach-2 |  |
| Chapter-3: Methodology |  |
| 3.1 Introduction |  |
| 3.2 Existing System |  |
| 3.3 Requirements |  |
| 3.4 Proposed System |  |
| 3.5 Workflow Diagram |  |
| Chapter-4: Design |  |
| 4.1 Data Flow Diagrams/UML Diagrams |  |
| Chapter-5: Implementation |  |
| 5.1 Datasets Used |  |
| Chapter-6: Conclusions |  |
| Chapter-7: Future Scope |  |
| References |  |
| APPENDIX A: Dataset Descriptions |  |
| APPENDIX B: Source Code |  |
| APPENDIX C: Result Graphs which are not included in main document. |  |

**CHAPTER 1**

**INTRODUCTION**

The academic world is characterised by a never-ending search for knowledge, and at the centre of this endeavour is a wealth of research resources, including articles, patents, books, and journals. Academic development is fundamentally driven by college instructors and researchers' ability to efficiently handle and access this plethora of knowledge. It is not just an issue of convenience. Seeing this need, we set out to create a state-of-the-art application that will revolutionise the handling and availability of research resources.

In a time when technology permeates every part of our existence, we must take advantage of its potential to efficiently automate the management of research materials. Modern technologies like HTML, CSS, JavaScript, ReactJS, MongoDB, and ExpressJS will be integrated to make this project a reality. These technologies will collectively serve as the foundation for our creative solution. Through the simplification of the process of organising, finding, and retrieving materials, this application aims to empower researchers and faculty members.

A strong search function is one of our application's main features. With the use of pertinent terms that are directly related to their job, users will be able to search for research papers and journals with this functionality. The days of searching through innumerable documents or having trouble locating particular study resources in a jumbled digital environment are long gone. Rather, customers may easily peruse an extensive library of scholarly materials, greatly improving the effectiveness of their research procedure.

Our application provides an advanced search option in addition to keyword-based search. With the use of this function, individuals can look up research articles by using the information contained in their own study abstracts. Finding resources that support a user's research goals adds a personalised touch to the academic process and increases focus and productivity.

Furthermore, by utilising data analytics, our application seeks to provide a comprehensive perspective on faculty research activities. Comprehending the academic research environment is essential for well-informed decision-making, resource distribution, and long-term planning. Our system will generate a culture of research excellence by gathering and evaluating data on research activities, giving important insights into the faculty's research areas of strength and improvement.

Our solution's core leverages the power of Natural Language Processing (NLP) and Machine Learning (ML) algorithms, going beyond simple technological integration. These state-of-the-art tools are essential for efficiently retrieving research records from our database. The system is able to suggest and rank items that most closely match the user's demands since the algorithms are able to understand the context and content of each document. All things considered, this technology-based strategy greatly improves both the user experience and the calibre of search results.

In summary, our project to create an automated research material management programme sits at the nexus of academics and technology. It has the potential to completely change how academic resources are accessed and used by college professors and scholars. Our approach, which combines HTML, CSS, JavaScript, ReactJS, MongoDB, ExpressJS, NLP, and ML, has the potential to significantly increase research efficiency and effectiveness by making important academic materials easily accessible.

**Problem Statement**

* For college professors and researchers, managing research materials—which includes papers, patents, publications, and journals—is a difficult and time-consuming endeavour. The vast amount of academic resources available and the dynamic nature of knowledge acquisition make it difficult to effectively arrange, retrieve, and make use of these resources. This causes a number of related issues for the academic community, including:
* Information Overload: of the digital era, there is an exponential increase of research materials, leading to an overload of information. It is difficult for college professors and scholars to browse through enormous data banks and locate pertinent items fast.
* Ineffective Resource Discovery: Present research material discovery techniques are frequently ineffective. In order to locate pertinent materials, researchers frequently rely on their personal network or suggestions as traditional keyword searches may not produce thorough or accurate results.
* Fragmented Research Management: The organisation of research materials is dispersed among multiple platforms, ranging from digital databases and personal folders to physical libraries. Disorganisation, duplication, and the loss of important research materials can result from this fragmentation.
* Lack of Personalization: Many times, users are unable to locate materials that closely correspond with their unique study aims in the current systems, wasting time and energy on unrelated content.
* Data-Driven Decision-Making: To make well-informed decisions, researchers and academic institutions need to know about faculty research activities. Institutions could find it difficult to identify areas of research excellence or to distribute resources efficiently in the absence of a uniform data analytics system.
* User Experience: The way in which research resources are accessed and managed by users is frequently not at their best. Users' productivity can be negatively impacted by slow search engines, difficult-to-use interfaces, and restricted access to sophisticated search options.

A creative application that can automate research material management, expedite the discovery process, provide personalised search capabilities, and offer data analytics to support better informed decision-making within academic institutions is desperately needed in light of these difficulties. A system like this would enable academics and researchers at colleges to more easily traverse the wide academic terrain, increasing their productivity and efficacy in their quest for information and superior research.

**1.3 Objective of the project:**

Manage research materials automatically, offer data analytics, boost discovery through tailored search, and enhance the user experience for academics and researchers at universities.

**1.4 Scope:**

The main goal of this project is to create an application that makes it easier for researchers and college instructors to handle their research materials. Using technologies like HTML, CSS, JavaScript, ReactJS, MongoDB, and ExpressJS, the scope includes developing an integrated platform that automates the organisation and retrieval of research resources, including papers, patents, publications, and journals. The programme will improve the effectiveness of material discovery by offering a user-friendly search system with keyword-based and sophisticated search options based on research abstracts. In addition, it will leverage machine learning (ML) and natural language processing (NLP) algorithms to provide tailored recommendations for study materials, guaranteeing that users can obtain materials that support their research goals. In order to help academic institutions make data-driven decisions, the initiative will also feature a data analytics component that will offer insights on faculty research activity. The programme will place a high priority on a responsive and user-friendly interface that is available across several platforms and protects user privacy and data security. Essential elements of the project include a strategy for long-term sustainability, ongoing user training, and support. The initiative intends to optimise the experience of conducting academic research by revolutionising the management and accessibility of research materials.

**1.5 Project Introduction:**

An essential component of success in the academic world, where information acquisition is valued highly, is effective research material management. With the proliferation of research papers, patents, publications, and journals that have come with the digital age, college instructors and researchers are faced with an enormous task. Our goal is to create a cutting-edge application that tackles these issues and transforms the academic environment. Our project attempts to use current technology to manage and access research materials more efficiently by offering a dynamic platform that combines multiple technologies, including HTML, CSS, JavaScript, ReactJS, MongoDB, and ExpressJS.

This project's main goal is to develop a user-centered application that enables academic researchers and instructors to effectively manage their research materials. The main purpose of the application is to provide a thorough and intuitive search mechanism that makes it easier to find pertinent study resources. Users will be able to utilise keyword-based searches, which will make it easy for them to find research articles and publications that are relevant to their job. In addition, our sophisticated search feature based on research abstracts makes the study experience more customised by helping users locate materials that are exactly in line with their research goals. We hope to lessen the amount of human labour required in the process by automating research material handling, freeing up time for the most important thing—research.

Our application will leverage the capabilities of machine learning (ML) and natural language processing (NLP) algorithms in addition to enhancing material discovery. By offering users tailored recommendations for study materials, these state-of-the-art technologies will make sure that they have easy access to resources that are directly related to their individual research interests and goals. The initiative will also include a data analytics component that provides faculty research efforts with insights, giving academic institutions the data-driven tools they need to make well-informed decisions. The project's focus goes beyond protecting sensitive research materials with strong data security measures, to include improving the user experience with a responsive and user-friendly interface that is available on several devices. The project's scope is completed by user training and support, a long-term sustainability strategy, and other elements that guarantee this cutting-edge application will adapt to the demands of academia and improve college instructors and researchers' entire research experience, productivity, and efficiency.

**1.6 Thesis of Organization:**

In this report chapter 1 describes about introduction of the project which includes overall process in a brief way. Chapter 2 describes about literature survey of 3 different papers which their contribution on their papers using NLP, keyphrases. Chapter 3 describes about the methodology which includes the difference between existing systems and proposed systems and the work flow diagram. Chapter 4 describes about the design which includes the objectives of design where the output is readability, relevance, visual appeal, custamizaion and uml diagram. Chapter 5 describes about the implementation of dataset which includes the dataset of patents, journals, conferences and book chapter and also experimental setup. Chapter 6 describes about the conclusion of the project which includes the overall information about research papers of our college faculty. Chapter 7 describes about the future scope of the project which includes the combination of NLP and machine learning algorithms.

**Chapter 2**

**LITERATURE SURVEY**

**2.1 Related Work:**

Ronan R. K. Ando and T. Zhang [1] The paper "Natural Language Processing (almost) from Scratch" authored by Ronan R. K. Ando and T. Zhang and published in the Journal of Machine Learning Research (JMLR) in November 2005, is a pivotal work that has significantly shaped the field of Natural Language Processing (NLP). This research introduces a groundbreaking framework that enables the development of NLP systems with minimal reliance on prior knowledge or extensive feature engineering. One of the key innovations in this paper is its emphasis on leveraging unlabeled data, a fundamental concept in unsupervised learning, to construct NLP models. This approach revolutionizes the conventional paradigm by allowing the system to draw knowledge from vast pools of unannotated text data, reducing the need for manual feature engineering and human domain expertise.

Furthermore, the paper introduces the concept of transfer learning to NLP, demonstrating how knowledge acquired from one task can be effectively transferred to another related task. This idea has profound implications for NLP, as it enables the pre-training of models on large, general-purpose datasets and fine-tuning them for specific NLP tasks, resulting in significant performance improvements. The paper's contributions extend well beyond its initial publication, as the framework and ideas it presents have inspired subsequent research, leading to the development of highly effective NLP models and systems.

Published in the esteemed JMLR, this paper has had a lasting impact on the machine learning and NLP communities, serving as a cornerstone for research and development in the field. Its focus on reducing the barriers to entry for NLP system development and improving the use of unlabeled data and transfer learning has paved the way for the development of more sophisticated and accessible NLP solutions. goal of the Intrusion Detection System (IDS) is to automate the process of filtering and blocking abnormal traffic into our network. The anomaly detection process is considered as one of the NP-Hard problems. The process of discriminating the network traffic as good traffic or bad traffic is crucial because the detection system must be equipped with suitable knowledge with the help of a knowledge base. Unfortunately, the research contributions published so far, on anomaly detection are still incomplete because of the continuous challenges coining from rapid technological changes.

Section 2.2 of this chapter presents research contributions in the literature that motivated the present study, to carry out this research. Section 2.3 then, presents different machine learning techniques such as, supervised and unsupervised that are employed in the design of intrusion detection systems and pays attention to different measures that are employed. Section 2.4 of this chapter outlines different similarity measures used in the existing research literature for the identification of anomalies.

Ben Wellner [2] The literature survey titled "Automatically Identifying the Arguments in Discourse Connectives" by Ben Wellner, affiliated with The Mitre Corporation in Bedford, MA, USA, addresses the challenge of automatically identifying the arguments of discourse connectives in the Penn Discourse TreeBank (PDTB). Rather than attempting to identify the full extents of these arguments as annotated in the PDTB, the paper reframes the problem by focusing on identifying the argument heads, effectively circumventing the complexity of discourse segmentation. The authors showcase substantial improvements in their approach by incorporating features extracted from a dependency parse representation, which outperforms those derived from a constituent-based tree parse. Additionally, the paper highlights the significance of capturing inter-argument dependencies through a log-linear re-ranking model, achieving a high accuracy rate in correctly identifying both arguments for over 74% of the connectives on held-out test data with gold-standard parses.

This literature survey contributes to the field of computational linguistics and discourse analysis by presenting an innovative approach to handling the challenge of identifying arguments within discourse connectives. The paper's emphasis on dependency parsing and its effectiveness in improving accuracy is noteworthy. By sidestepping the intricacies of discourse segmentation, this work simplifies and enhances the identification of discourse connective arguments. It offers valuable insights into the practical applications of natural language processing and discourse analysis in various domains, including machine learning and language understanding.

The authors' approach, which combines the benefits of dependency parsing and a log-linear re-ranking model, presents a promising avenue for future research in discourse analysis and automated text understanding. This paper's findings and methodologies have the potential to impact the development of more sophisticated and accurate NLP systems and language understanding tools.

Hinton [3]The abstract describes a methodology that leverages "complementary priors" to address challenges in densely connected belief networks with multiple hidden layers. By using these priors, the paper introduces a fast and efficient algorithm for training deep, directed belief networks one layer at a time. This method is particularly effective when the top two layers of the network create an undirected associative memory. The fast, greedy algorithm serves as an initialization step for a slower learning process that refines the network's weights using a contrastive version of the wake-sleep algorithm.

The outcomes of this research are noteworthy. After the fine-tuning process, a network with three hidden layers can effectively model the joint distribution of handwritten digit images and their corresponding labels, serving as a powerful generative model. This generative model surpasses the performance of discriminative learning algorithms in digit classification tasks, highlighting the potential of generative models. The abstract also suggests that the network can represent low-dimensional patterns, akin to "ravines" in a free-energy landscape, within the top-level associative memory. This representation allows for the exploration of the patterns and insights contained in the associative memory, offering a unique perspective on the information encoded in the network.

**CHAPTER 3**

**METHODOLOGY**

**3.1 Existing System**

The inability of the current system to conduct precise keyword-based searches for research documents made it more difficult for users to find pertinent resources quickly. Furthermore, the lack of statistical analysis tools and data visualisations restricted the system's ability to draw insightful conclusions and identify patterns from the gathered data. The lack of these essential features in the system not only made for a worse than ideal user experience, but it also made it more difficult to manage and use academic resources effectively for research.

**Disadvantages of the Existing System:**

1. 1. Ineffective Keyword Searches: A time-consuming and frequently frustrating search process resulted from the inability to accurately search for research data using keywords.
2. 2. Lack of Data Visualisations: Users were unable to derive any useful conclusions from the research materials and their usage patterns due to the lack of data visualisations and statistical analysis.
3. 3. Decreased Productivity: Searching and organising research materials required additional work from users, which eventually led to a decrease in productivity.
4. Missed Opportunities: Without efficient search capabilities and data visualizations, the system missed opportunities for users to discover valuable research materials and trends in their academic activities.
5. Limited Decision Support: The system's inabili 4. Missed Opportunities: Without efficient search capabilities and data visualizations, the system missed opportunities for users to discover valuable research materials and trends in their academic activities.

5. Limited Decision Support: The system's inability to offer statistical analysis limited its potential to support data-driven decision-making in academic institutions.

**3.3 Proposed System**

Primarily focused on user efficiency and relevance, the suggested system aims to transform the way academic research resources are organised and accessible. In order to conduct highly targeted searches and make sure the publications they obtain closely match their particular research fields, users can now submit their research abstracts. To further improve the search experience, our programme has an advanced search tool that makes it simple for users to find research papers and journals relevant to their job using keyword-based searches. Using this method allows users to find materials that precisely align with their research aims, which maximises their academic endeavours and streamlines the process of conducting research, making it more efficient and customised.

**3.4 Advantages of the Proposed System:**

1. Improved Relevancy: Users can find resources that are closely related to their research goals by using the system's capacity to search research materials based on research abstracts. This saves time and makes the results more relevant.

2. Advanced Keyword Search: This feature makes it simpler for users to find relevant research papers and journals by streamlining the process of material discovery using keyword-based searches.

3. Personalization: By allowing users to examine resources that are relevant to their particular fields of study, research abstracts for searches provide a personalised touch.

4. Efficiency: The system increases productivity and efficiency in academic research activities by automating the management and search processes.

5. Streamlined Discovery: Users' entire research and knowledge acquisition experience is improved by the ease with which they may access and handle academic materials.

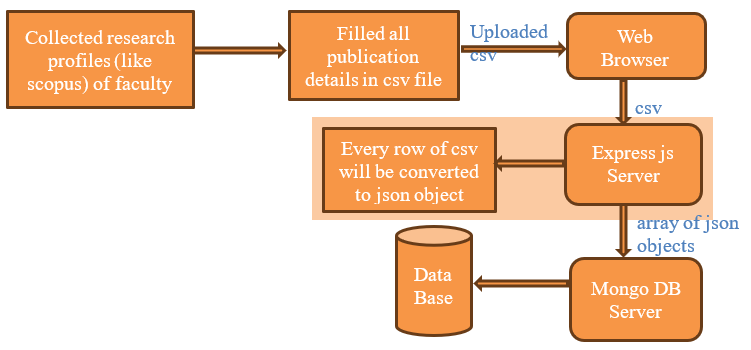
**3.5 Workflow of Proposed System** A diagram of a computer program

Description automatically generated

Fig.3.5.1 workflow

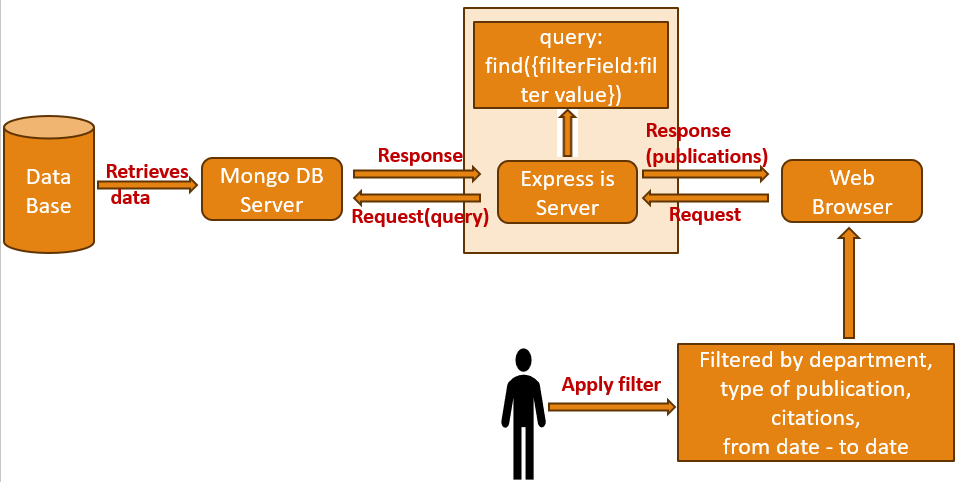
The suggested system's workflow is outlined in Figure 3.5.1. First, a client accesses the web browser and sends a request for the operations they need. The Express JSON server responds to the client by sending a request to the Mongo DB server, and the response is sent while the Express JSON server uses NLP and ML techniques.

**3.6 Methodology for uploading publications**

****

**3.7Methodology for Publication Retrieval**



****



**3.8 Software Requirements:**

Operating System : Windows 8GB RAM or above

Server Side Script : JavaScript

Programming Language : JavaScript

Libraries : Express,axios,mongoose,react-icons etc

IDE/WorkBench : Visual studio code

Technology : Mern Stack (MongoDB,ExpressJS,Reactjs,Nodejs)

Database : MongoDB

**3.9 Hardware Requirements:**

Operating system: Windows

Processor : intel i5

RAM :8GB or Above

**CHAPTER 4**

**DESIGN**

**4.1 Introduction of Input Design:**

Since it acts as the entry point for user interactions with the system, the input design is an essential part of every information system. In order for users to enter data and communicate with the system, forms, displays, and interfaces must be created. To guarantee that users can input data accurately, quickly, and in an intuitive way, an efficient input design is necessary. Within the framework of our academic research project, the input design is essential to allow users to easily search for research resources, enter study abstracts, and communicate with the system. It is a crucial component in guaranteeing the system's usability and directly affects the general efficacy and efficiency of tasks associated with research.

**4.1.1Objectives for Input Design:**

1. **User-Friendly Interface:** The main goal of input design is to provide an interface that is easy for users to use so they can enter data and communicate with the system. This entails creating user-friendly forms and interfaces that don't needlessly impede user input.
2. **Data Accuracy:** Ensuring data accuracy is another crucial goal. To reduce data entry errors and preserve data integrity, validation methods and error-checking procedures should be incorporated into the input architecture.
3. **Efficiency:** It is important to optimise the input design for efficiency. Inputting data should be quick and simple for users, saving time and effort when doing tasks like finding research resources or creating study abstracts.
4. **Compatibility:** To support a range of user preferences, input design should be interoperable with multiple platforms and devices, including desktop computers, mobile devices, and tablets.

**4.1.2Introduction to Output Design:**

The system's information and result presentation to users is the main emphasis of output design. It is essential for guaranteeing that the data is presented in a way that is clear, concise, and meaningful. Displaying search results, data analytics, and any other information obtained from the system's functions depends on the output design in the context of our project. A clear, aesthetically pleasing, and useful information that supports users' decision-making and research endeavours is the goal of effective output design.

**4.1.3Objectives for Output Design:**

1. **Readability and Clarity:** Presenting information in a legible and clear style is the main goal of output design. Users should find it simple to comprehend and analyse the data that the system presents.
2. **Relevance:** The presentation of pertinent information ought to be given top priority in output design. It seeks to guarantee that users obtain information and understandings that are directly pertinent to their goals and areas of interest in research.
3. **Visual Appeal:** To make the output visually appealing and make it easier for users to understand complex information, visual components like charts, graphs, and data visualisation should be employed.
4. **Customization:** Users should be able to arrange and visualise data in a way that best fits their preferences and research requirements by having some degree of control over the output from the system.

**4.2 UML Diagram:**

**4.2.1 Use Case Diagram:**

The functional requirements and interactions between various actors (users or external systems) and a system are represented visually in a use case diagram. It is beneficial to show the different scenarios or tasks that a system is capable of carrying out and how they connect to users or other external entities.

The use case diagram, as displayed in fig. 4.1, describes the interactions that take place between different actors and the system, specifying their interactions and the particular tasks they are able to carry out. These details provide further details about the characters and their roles in the context you indicated, which seems to be connected to a website:

**Users:** One of the main actors in the use case diagram of a website is the user. They stand for people who use the website for a variety of reasons. Depending on the type of website, users may assume the roles of students, researchers, readers, or clients, among others. A few examples of user-related use cases are "Browsing Content," "Searching for Information," "Registering for an Account," "Logging In," and "Providing Feedback."

**Authors:** Performers who write on the website are called authors. These people write content for the website, including posts, articles, and other items. An author's use cases could be "Creating Content," "Editing Posts," "Publishing Articles," and "Managing User Comments."

**Administrators:** They are essential to the upkeep and management of the website. They are able to manage user accounts, control content, guarantee security, and take care of technological matters. Administrator use cases could include tasks like "Content Moderation," "User Account Management," "Security Configuration," and "Database Maintenance."

**Website:** The website itself is depicted as an actor in certain use case diagrams. It is a point of reference that encompasses the functionality of the system, even though it might not actively interact with other actors. Use cases for the website would include essential features like "feedback collection," "user registration," "search functionality," and "content delivery."

Use case diagrams are a valuable tool for comprehending the functioning of a system and verifying that user demands and system requirements are clearly stated and fulfilled. They offer a clear visual depiction of the roles and interactions inside a system.

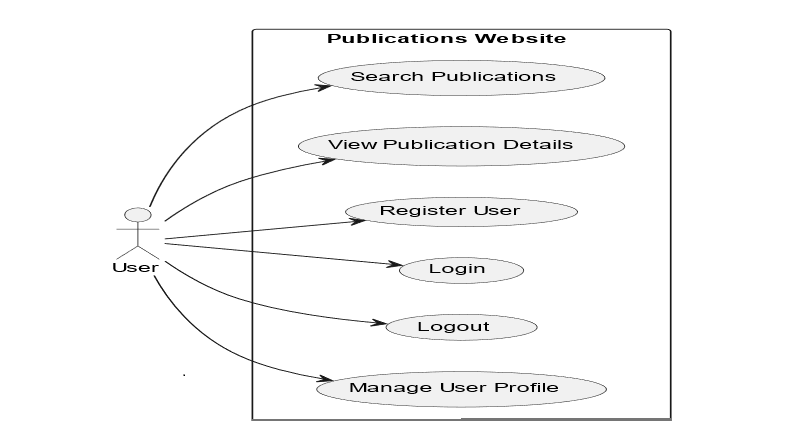
****

Fig 4.2.1.1 Use case Diagram for users

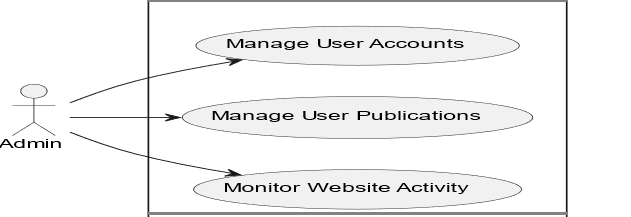


Fig 4.2.1.2 Use case Diagram for Admin

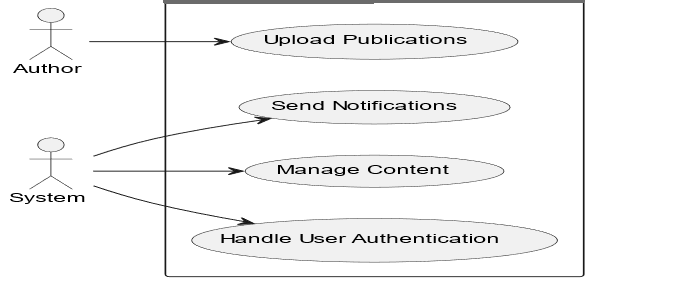


Fig 4.2.1.3 Use case Diagram for Authors

* + 1. **Class Diagram:**
* The class diagram, which is a sort of UML (Unified Modelling Language) diagram that depicts the structure and organisation of a system or software programme, is displayed in Figure 4.2 It is an essential tool for building and modelling object-oriented systems because it gives a visual representation of the classes, objects, connections, and attributes inside a system.
* A system's or software application's structure and organisation are represented by a class diagram.
* The classes or objects that make up a system, as well as their properties, functions, and connections to other classes, are represented using class diagrams.
* These are a few of the important courses.
* **User:**stands in for a website user.
* Properties: UserID, Username, Email;
* Techniques: Register(), Login(), Logout()
* **Publications**: Represents one of the website's publications.
  + PublicationID, Title, Content, Author are among the attribute.
  + Display() and CommentOn() are the methods.

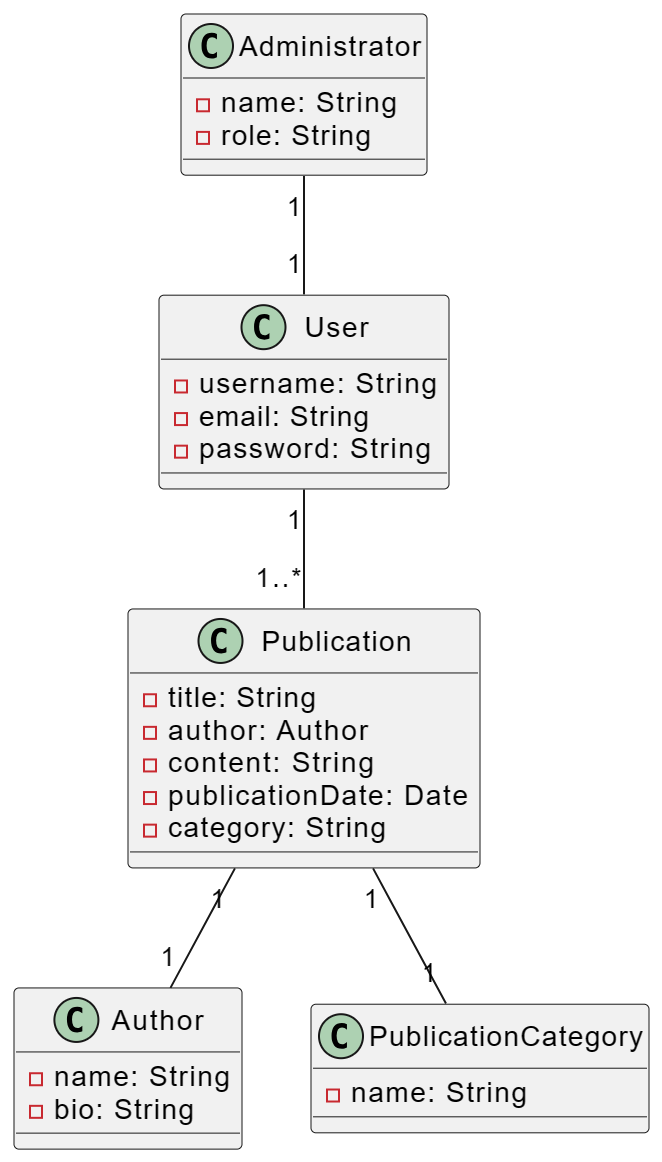


Fig 4.2.2 Class Diagram

**4.2.3 Sequence Diagram:**

One kind of Unified Modelling Language (UML) diagram that shows the relationships and message sequences between various objects or system components is a sequence diagram. It offers a dynamic picture of how several items work together within a system or software programme to complete a specific activity or reach a use case. Figure 4.2.3.1 illustrates how the diagram is depicted:

* A sequence diagram shows how several parts or objects interact with one another over a given amount of time in a system.
* A sequence diagram shows how several parts or objects interact with one another over a given amount of time in a system.**Participants** :
* User
* Website (Controller)
* Server
* A diagram of a website

  Description automatically generatedDatabase

Fig 4.2.3 Sequence Diagram

* + 1. **Activity Diagram:**

The workflow and activities inside a system, process, or business operation are modelled and visualised using an activity diagram, which is a form of Unified Modelling Language (UML) diagram. It is an effective tool for process modelling, analysis, and communication because it offers a high-level perspective of the phases, actions, and decision points involved in a particular process or activity.

* An activity diagram is employed to represent the dynamic elements of a process or system.
* A diagram of a company

  Description automatically generated with medium confidenceIt is especially helpful for visualising the workflow and processes within a system because it depicts the flow of actions, decisions, and activities within it.

Fig 4.2.4.1 Activity Diagram From Login to Logout

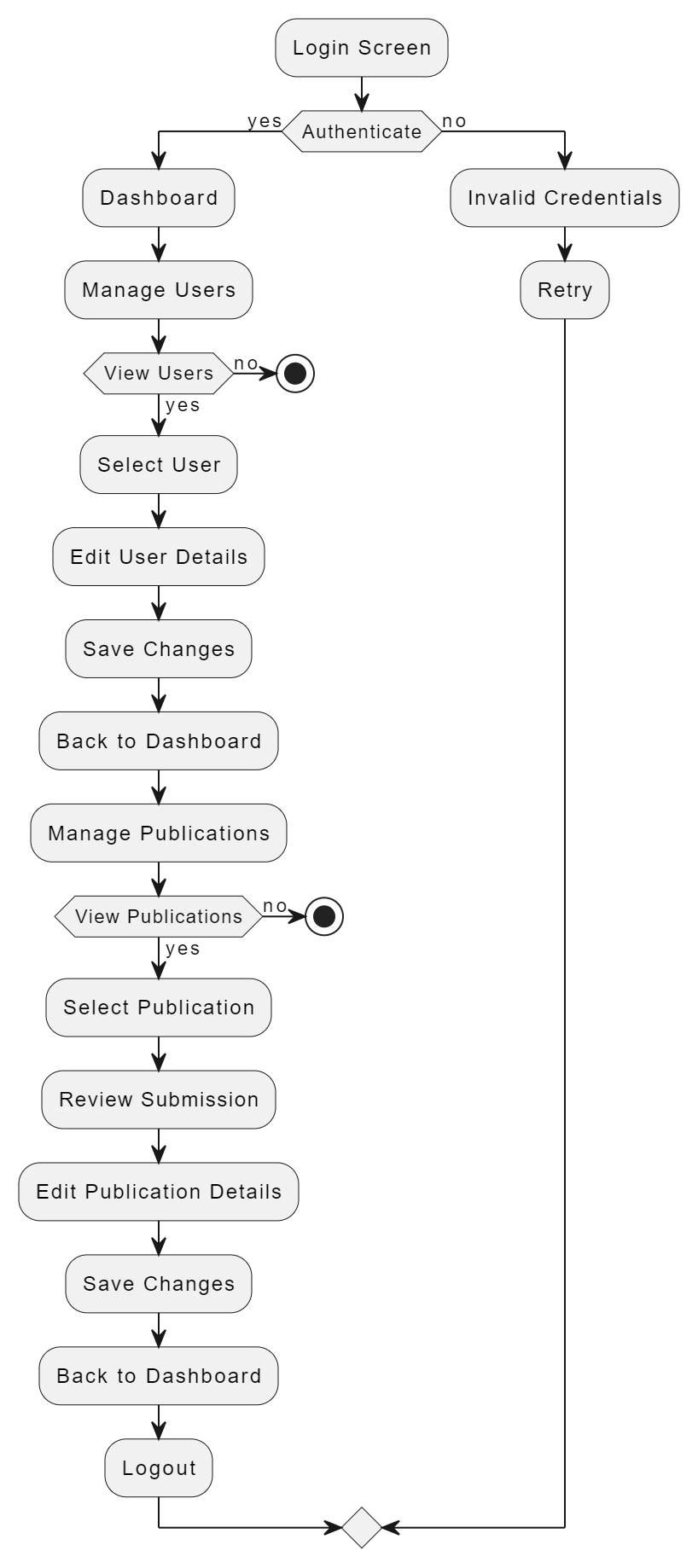


Fig 4.2.4.2Activity Diagram For Admin

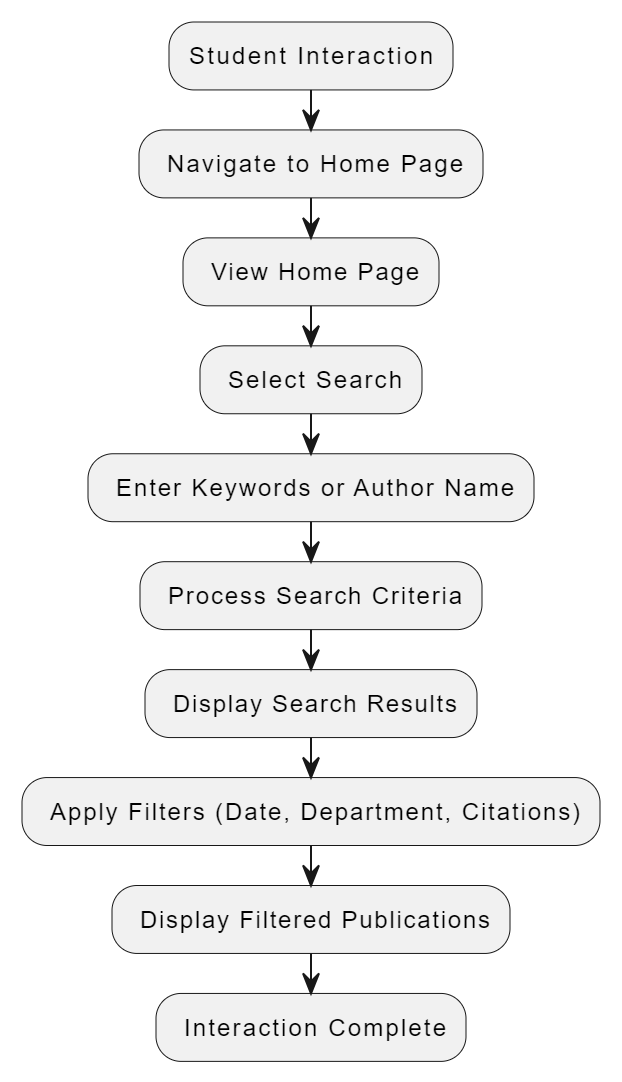


Fig 4.2.4.3 Activity Diagram For Student

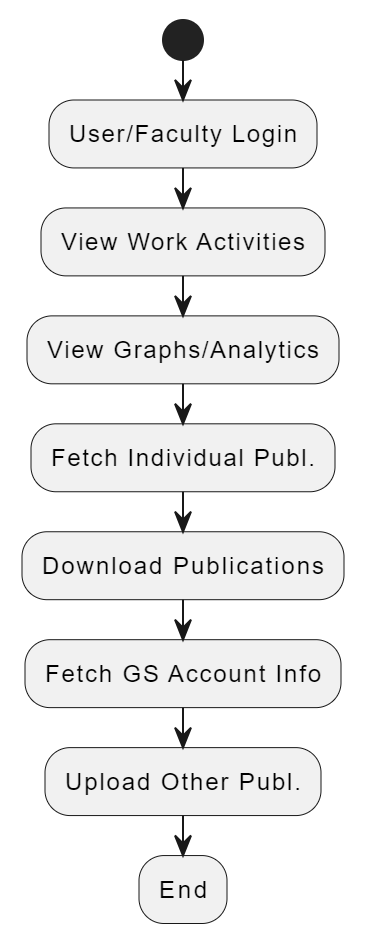


Fig 4.2.4.4 Activity Diagram For User

**CHAPTER 5**

**IMPLEMENTATION**

**5.1 Dataset used:**

A wide range of documents, including book chapters, journals, conferences, and patents, are included in our dataset.

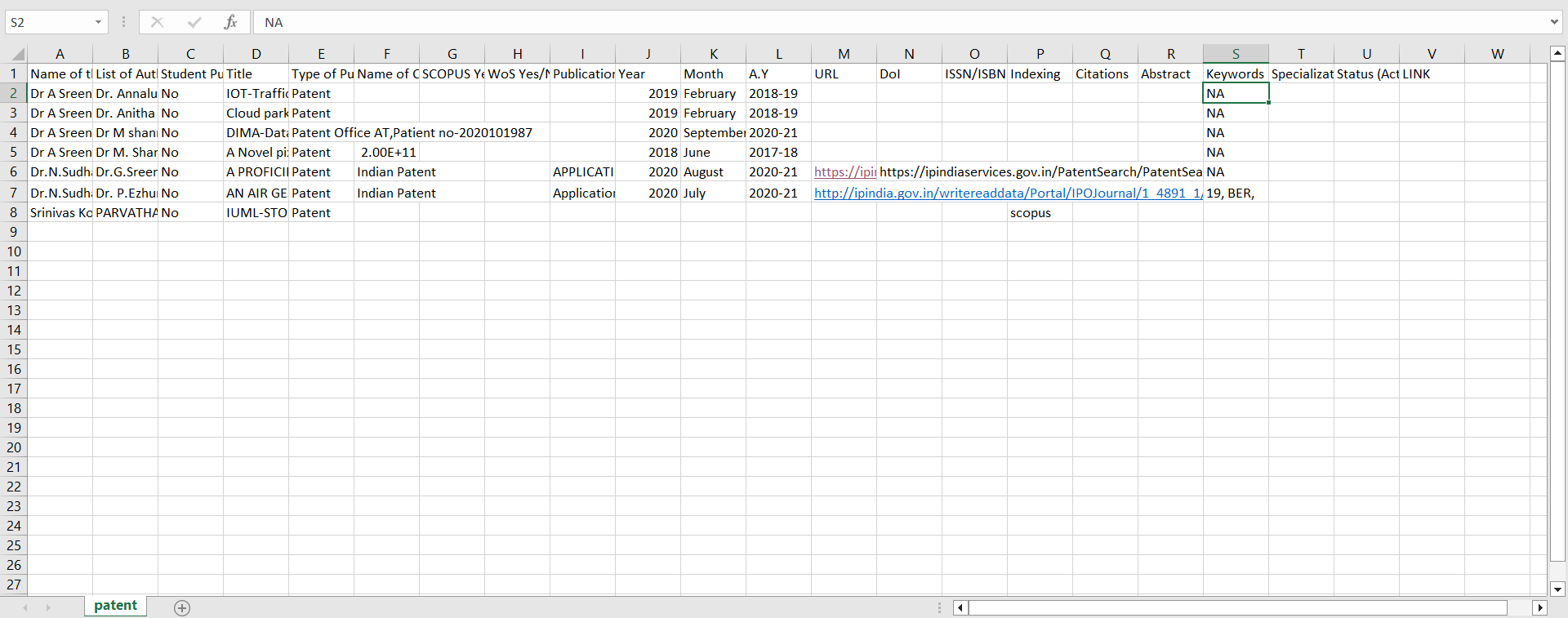
Patents: As seen in fig. 5.1.1, a government licence or authorization that grants a title or right for a predetermined amount of time, particularly the exclusive right to prevent third parties from creating, utilising, or commercialising creations.

Fig 5.1 Patents

**Journals:**

As shown in fig 5.1.2 newspaper or magazine that deals with a particular subject or professional activity.

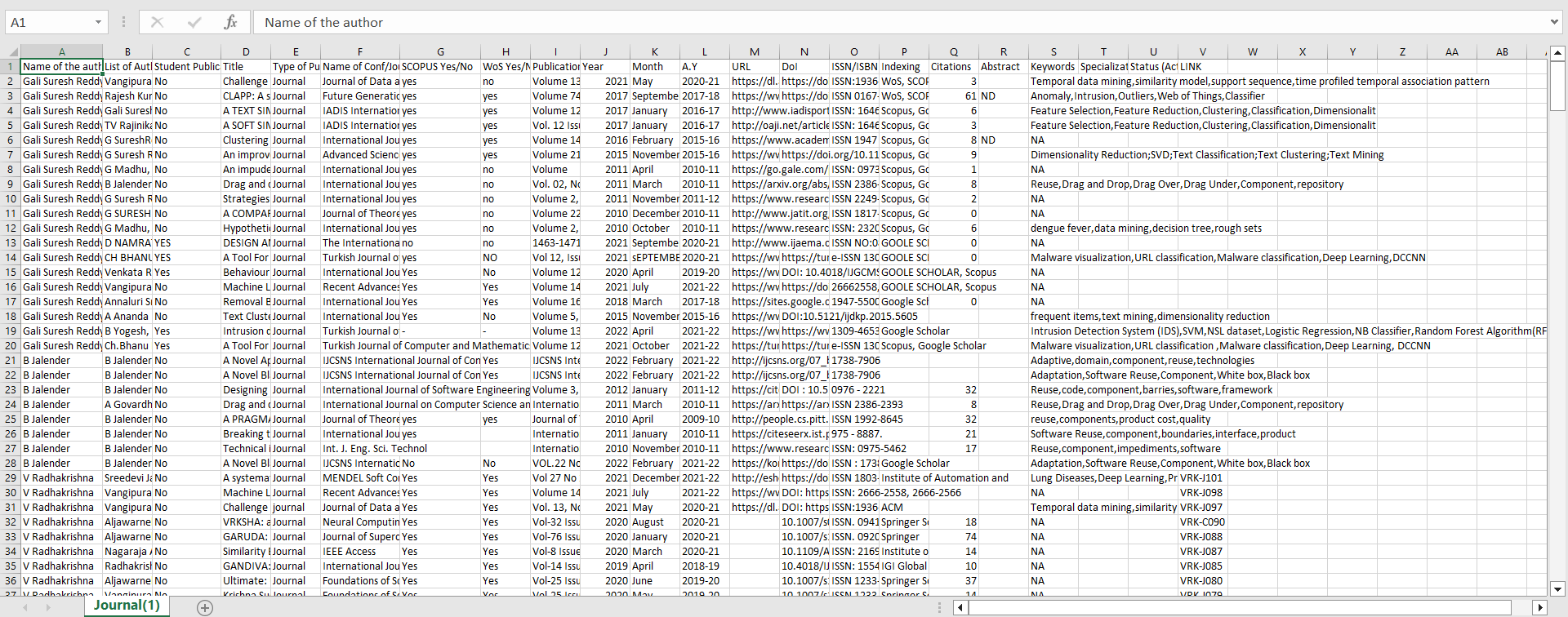


Fig 5.2 Journals

**Conferences:**

As shown in fig 5.1.3 formal meeting of people with a shared interest, typically one that takes place over several days.

A screenshot of a computer

Description automatically generated

Fig 5.3 Conferences

**Book Chapter:**

As Shown in fig 5.1.4 chapter is a chunk of a book that comes to a recognizable end, usually marked by a page break or by an extra space followed by a new numbered or titled chapter.

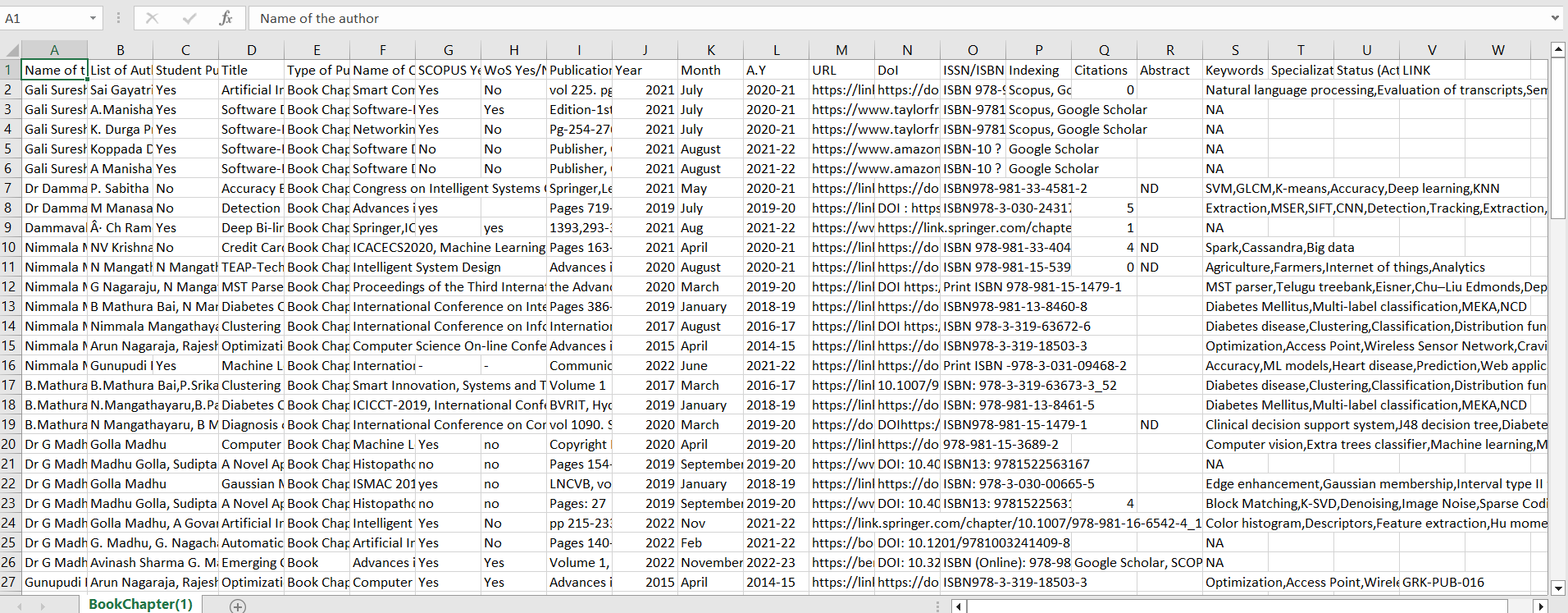


Fig 5.4 Book Chapters

* 1. **Experimental setup:**

Analytics:

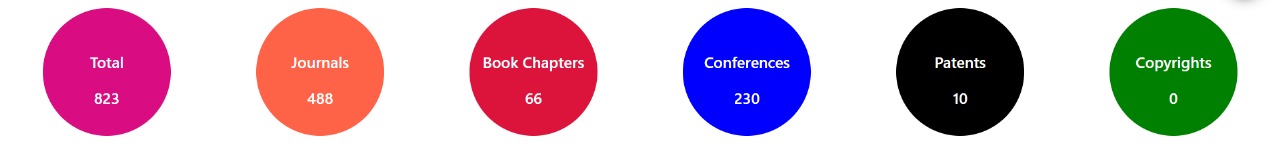


Fig 5.1

Fig 5.1 Shows the analytics about total number of publications and category of the publication and number of publications in each category

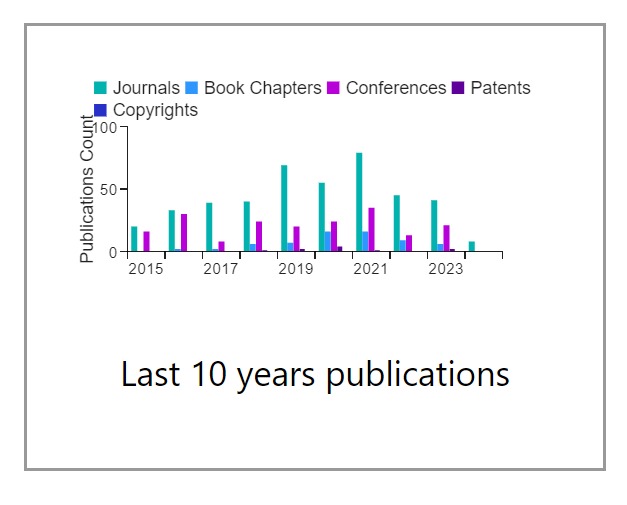
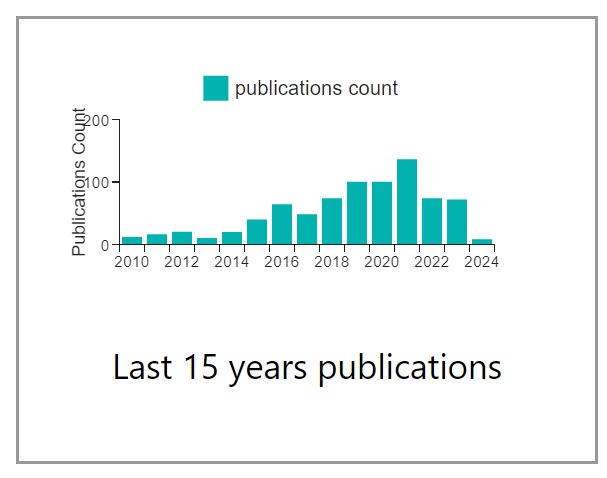


Fig 5.1.1 Fig 5.1.2

Figure 5.1.1 and figure 5.1.2 displays last 15 years and 10 years analytics based on total count and category respectively.

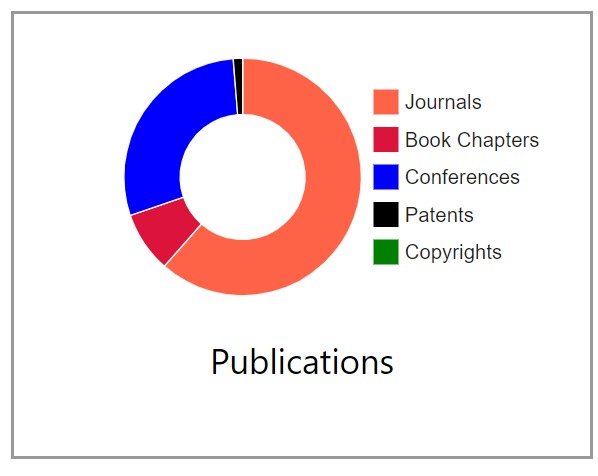
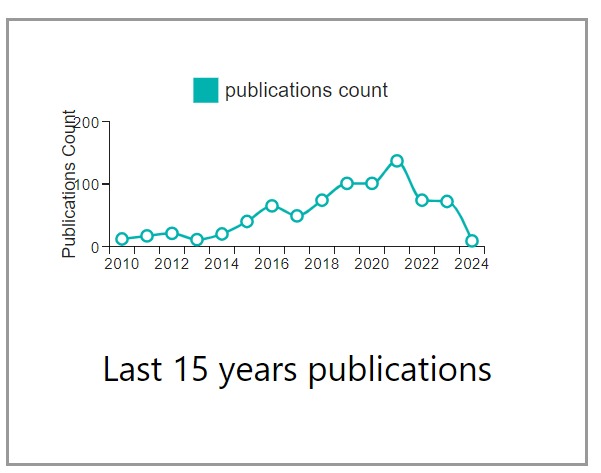


Fig 5.1.3 Fig 5.1.4

Figure 5.1.3 and figure 5.1.4 Illustrates the overall total number of publications and the publication count for the last 15 years using pie charts and line graph.

Home Page:

A screenshot of a computer

Description automatically generated

Fig 5.2 Home Page

Figure 5. 2 Illustrates the overall total number of publications and the publication count for the last 15 years using pie charts and bar graphs.

A screenshot of a computer

Description automatically generated

Fig 5.3 Home Page

Figure 5.3 showcases publications spanning the last 10 to 15 years and includes a navigation menu which provides user to navigate through the department. A screenshot of a computer

Description automatically generated

Fig 5.4 Home Page

In Figure 5.4, there are 20 publications displayed. If the user wants to view more, they can click on the 'ShowMore' button, it is redirected to login page.

A screenshot of a computer

Description automatically generated

Fig 5.5 Login Page

In Figure 5.5, users can access additional publications by verifying their email and password.

A screenshot of a computer

Description automatically generated

Fig 5.6 Admin Dashboard

Upon successful login with valid credentials, the Admin account will be authenticated, and a summary of the user's details and activities from the past 15 years will be displayed as showed in figure 5.6.

A screenshot of a computer

Description automatically generated

Fig 5.6.1 Upload new option

By Clicking on “Upload new “,admin can upload images to carousel ,can download and upload csv file figure 5.6.1.

A screenshot of a computer

Description automatically generated

Fig 5.6.2 Deleting publications

Admin can delete the publications all at once or selectively as shown in

figure 5.6.2.

A screenshot of a computer

Description automatically generated

Fig 5.6.3 User Profile

Alert box is shown to confirm delete as shown in figure 5.6.3.

A screenshot of a computer

Description automatically generated

Fig 5.6.4 Edit Page

Figure 5.6.4 shows edit page in which all the required fields are displayed and gives access to edit the content.

A screenshot of a computer

Description automatically generated

Fig 5.6.5 Edit Page

Figure 5.6.5 shows edit page in which all the required fields are displayed and gives access to edit the content.

A screenshot of a computer

Description automatically generated

Fig 5.7 User Profile

Upon successful login with valid credentials, the user account will be authenticated, and a summary of the user's details and activities from the past 15 years will be displayed as showed in figure 5.7.

A screenshot of a computer

Description automatically generated

Fig 5.8 User Activity for past 15 years

The chart in Figure 5.8 provides an overview of user details, including their total number of publications, citations, and a 15-year count.

A screenshot of a research activity

Description automatically generated

Fig 5.9 User Research Activity corresponding to sites Links.

Figure 5.9 displays various site links for users to access. By clicking on these links, users will be redirected to the respective sites. Additionally, the user's educational details are presented below.

A screenshot of a computer

Description automatically generated

Fig 5.10 Can update, download, and upload Publications.

The diagram labeled 5.10 shows that users have the option to update, download, and upload publications by clicking on the corresponding buttons. A screenshot of a computer

Description automatically generated

Fig 5.11 Can update, download and upload Publications.

As Shown in figure 5.11, By clicking the FetchGS button, the system updates the total number of publications through web scraping.

A screenshot of a computer

Description automatically generated

Fig 5.12 Can upload Publications.

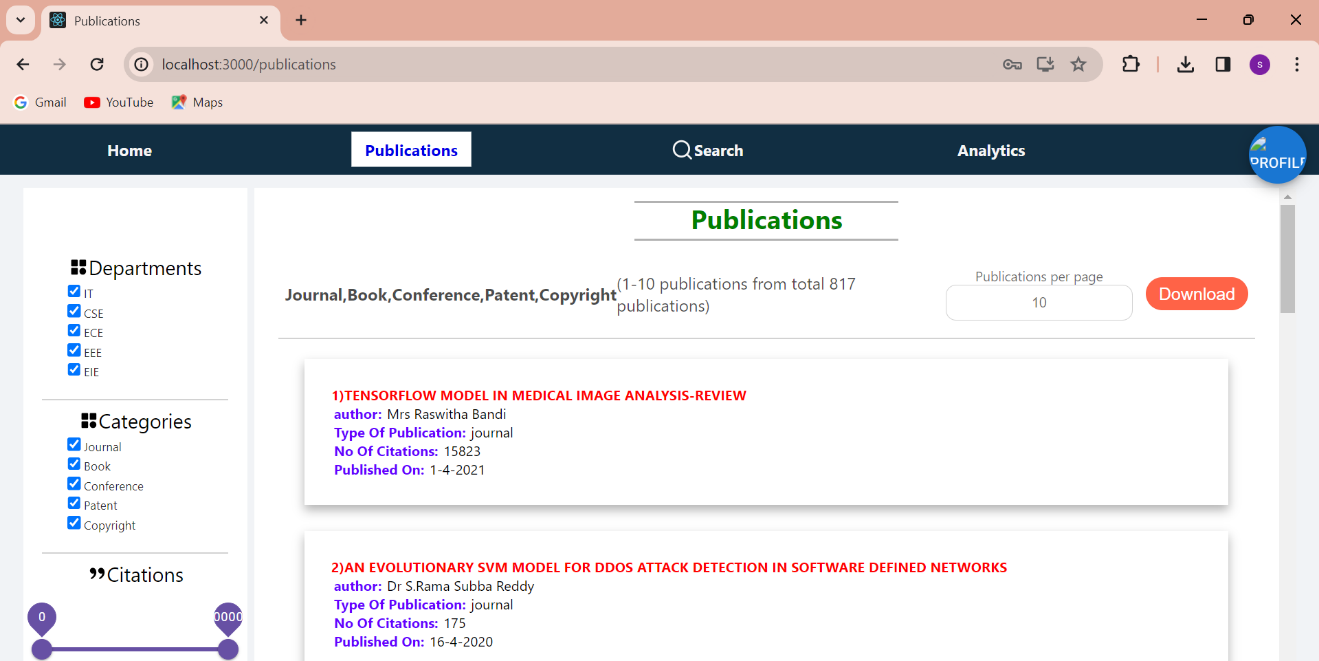
Users can upload publications by clicking the upload button. They should download the provided template CSV file, fill it out, and then upload the completed CSV file. +

Fig 5.13 Publications Page.

Figure 5.13 illustrates the total number of publications. Users have the option to download these publications.

A screenshot of a computer

Description automatically generated

Fig 5.14 Publications Page.

Figure 5.14 illustrates the total number of publications. Users have the option to download these publications.

A screenshot of a computer

Description automatically generated

Fig 5.15 Publications Page.

Figure 5.15 illustrates the total number of publications. Users have the option to download these publications

A screenshot of a computer

Description automatically generated

Fig 5.16 Publications Page.

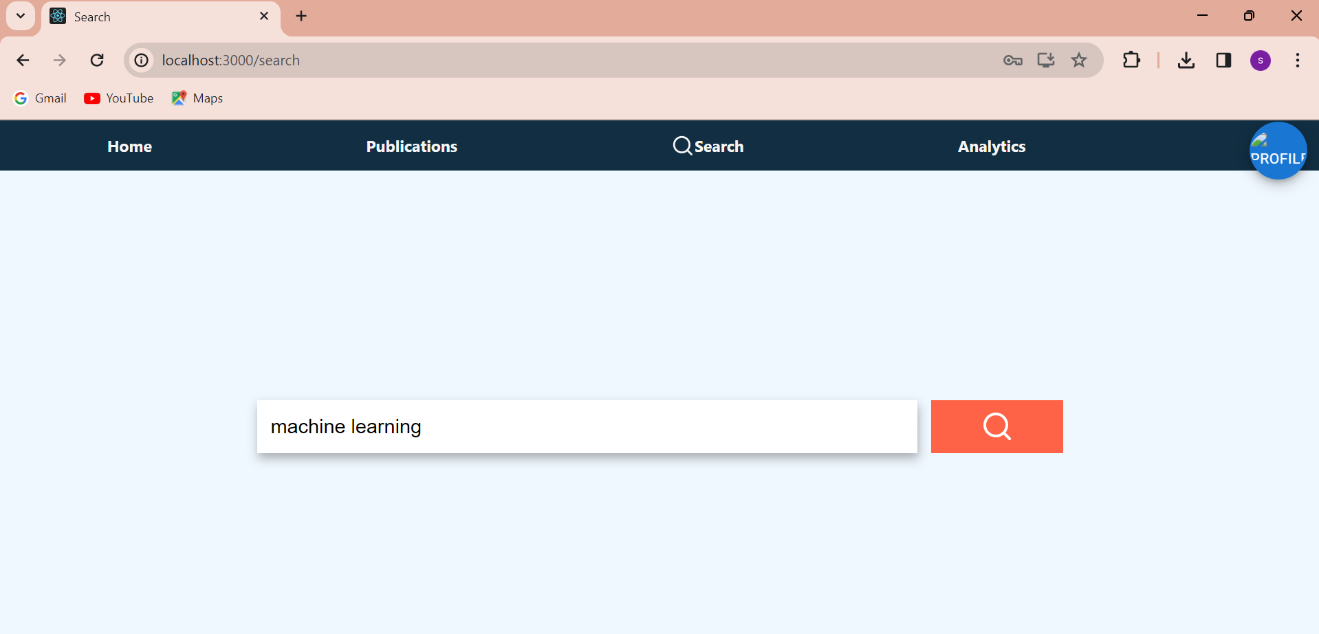
In Figure 5.16, users have the option to apply filters by interacting with checkboxes, drag bars, or date filters. When they click the apply button, the selected filter is applied, and relevant publications are shown. 

Fig 5.17 Search Page.

The user has the option to search by entering keywords or an abstract, as demonstrated in Figure 5.17.

**Document search using NLP:**

1. **Index Creation:**
   * Define which fields to index (e.g., keywords, abstract, title).
   * Specifies how data should be analyzed during search operations.
2. **Data Ingestion:**
   * Deploy documents onto the database after index creation.
3. **Query Parsing:**
   * User queries sent to the database using $search query operator.
   * Search parses the query to understand user intent, breaking it into tokens and identifying keywords.
4. **NLP Processing:**
   * Techniques include tokenization, lemmatization, part-of-speech tagging, NER, sentiment analysis, and semantic analysis.
5. **Query Execution:**
   * Executes search operation against indexed fields and documents.
   * Utilizes search algorithms like inverted indexes and TF-IDF.
6. **Presentation of Results:**
   * Results returned to the website's backend server and then to frontend for display.

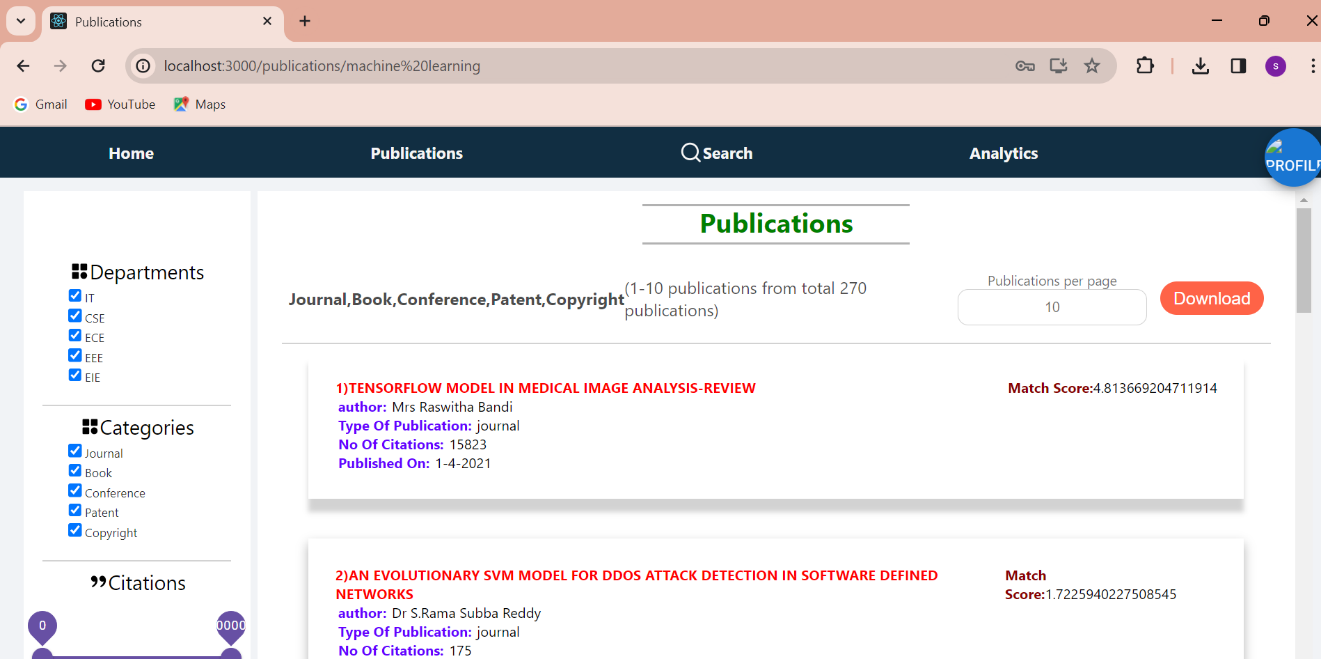


Fig 5.18 search page result

The results are generated and displayed on the screen based on the provided keywords or abstract, as shown in Figure 5.18. A screenshot of a computer

Description automatically generated

Fig 5.19 Account Logout.

When the user clicks on the logout button, they will be logged out of their account as shown in the figure 5.19.

**CHAPTER 6**

**CONCLUSION**

To sum up, our project marks a significant advancement in the field of academia as we set out on a revolutionary quest to improve the efficacy and efficiency of research material management for academic researchers and college instructors. This project is our answer to the issues presented by the unparalleled era of information dissemination that the modern era has ushered in. We are dedicated to using cutting-edge technology, such as HTML, CSS, JavaScript, ReactJS, MongoDB, and ExpressJS, since we want to provide academics with a cutting-edge tool that makes organising, finding, and retrieving research materials more efficient.

Fundamentally, the goal of this application is to completely transform the way that research resources are accessed by making them simply accessible and discoverable. Users may easily explore a multitude of academic resources with keyword-based and advanced search options, cutting down on search time and improving the effectiveness of their research projects. Additionally, the research experience is made more personalised by the integration of Natural Language Processing (NLP) and Machine Learning (ML) algorithms, which provide customised recommendations based on users' unique study goals.

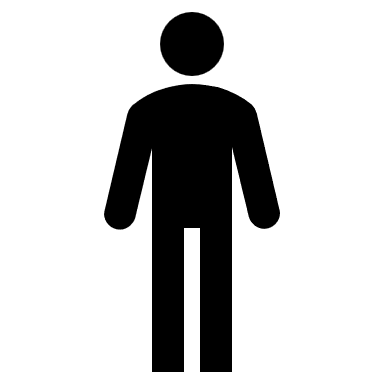
Beyond the efficiency benefits, our project aims to use data analytics to give academic institutions a complete picture of faculty research activity, enabling them to make well-informed decisions. We place a high priority on having an intuitive and responsive user interface so that users can easily access the programme on many platforms and that strong data security is maintained to protect confidential research materials.

In addition, we understand that continuous user education, assistance, and a sustainability strategy are critical to guaranteeing the application's durability and flexibility. This project is a major achievement at the nexus of technology and academia, not just a technological integration. It has the potential to increase research process productivity and efficacy, making it easier for academics and college instructors to move about the academic environment. As we go, our dedication to enabling academia with a technologically advanced, dynamic solution that maximises the potential of research materials and cultivates a culture of research excellence and knowledge progress does not waver.

**CHAPTER 7**

**FUTURE ENHANCEMENT**

We anticipate a state-of-the-art interface in the future that will seamlessly and effortlessly provide article searches, successfully addressing the shortcomings of the current system. By utilising the latest Machine Learning (ML) techniques along with Natural Language Processing (NLP) capabilities, our system will provide users with extremely precise and context-aware keyword-based and abstract-based searches, hence improving their overall experience. Future improvements will centre on sophisticated data visualisations, utilising machine learning algorithms to produce informative and interactive data representations, in addition to effective searching. These improvements will encourage a culture of research excellence and knowledge progress in the academic community by enabling users to easily locate research resources and to have a deeper grasp of the academic environment through dynamic and personalised data analytics.



Web Browser

**Request**

**Response**

**References**

1. Natural language processing (almost) from scratch ronan r. k. ando and t. zhang. a framework for learning predictive structures from multiple tasks and unlabeled data. jmlr, 6:1817–1953, 11 2005.
2. Ben Wellner, The Mitre Corporation 202 Burlington Road Bedford ,MA USA.
3. Hinton, G.E.; Osindero, S.; The, Y.W. A fast-learning algorithm for deep belief nets. *Neural Comput.* 2006, *18*, 1527–1554.