ENHANCEMENT FOR THE ACCESS AND UTILIZATION OF LIBRARY RESOURCES USING MACHINE LEARNING TECHNIQUES

Agrey Kato

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Information and Communication Science and Engineering of the Nelson Mandela African Institution of Science and Technology

Arusha, Tanzania

June, 2024

ABSTRACT

The growing demands for online information have motivated researchers to explore the most effectively use of digital library (DL) resource tools. The main challenges of online DL are information search and retrieval attributes related to label relevance and feature correlation segments. Previous research mainly relied on unbalanced multi-label data and therefore could not develop a reliable tool to access online information. To improve availability and usefulness of online DL, this work uses machine learning techniques to enhance the access and utilization of library resources. The research data were collected at The Nelson Mandela African Institution of Science and Technology (NM-AIST), Mzumbe University (MU), and the University of Dar es Salaam (UDSM) through questionnaire and purposeful sampling technique were then analysed with python and MAXQDA tools respectively. The survey found that 1,217 (73%) of respondents were aware of electronic information resources (EIRs) but faced accessibility limitations due to social and technical issues. Then, the proposed ensemble model (PEM) for machine learning (ML) methods was used to develop a resource discovery tool (RDT). The effectiveness of the PEM was then evaluated by comparing the accuracy of the PEM, logistic regression (LR), support vector machine (SVM), and knearest neighbor (kNN) algorithms. The experimental results reveal that PEM offers the highest precision of 95%, as compared to LR's 84%, SVM's 65%, and kNN's 57%. The Web Content Accessibility Guidelines (WCAG) 2.1 standards had been successfully used to test the four digital library tools, the developed RDT, NM-AIST, MU, and UDSM to see how well the developed system performs. The developed RDT had the highest established compliance score for online content accessibility, which is 90% with only one violation, compared to NM-AIST's 80% with 16 violations, MU's 55% with 12 violations, and UDSM's inability to be evaluated because of the excessive number of infractions. Therefore, the results of this study show the need to regularly check the accessibility of an online resources as well as optimization of the digital libraries.

DECLARATION

I, Agrey Kato do hereby declare to the Senate of Nelson Mandela African Institution of Science and Technology that this thesis is my own original work and that it has neither been submitted nor being concurrently submitted for degree award in anyother institution.

Agrey Kato Akato 10/12/2023

Candidate name and Signature **Date**

The above declaration is confirmed

Prof. Michael Kisangiri 10/01/2024

Name and Signature of Supervisor **Date**

Prof. Shubi Kaijage 12-12-2023

Name and Signature of Supervisor **Date**

COPYRIGHT

CERTIFICATION

The undersigned certify that they have read and hereby recommend for submission to the Nelson Mandela Institution of Science and Technology (NM-AIST) a thesis titled Enhancing the access and utilization of library resources using machine learning techniques, in fulfillment of the requirements for the degree of Doctor of Philosophy in Information and Communication Science and Engineering of the Nelson Mandela African Institution of Science and Technology.

Prof. Michael Kisangiri  10/09/2024

Name and Signature of supervisor **Date**

Prof. Shubi Kaijage  12-12-2023

Name and Signature of Supervisor **Date**

ACKNOWLEDGEMENTS

First of all, I thank Almighty God who gave me guidance and wisdom to complete this thesis; It is my great pleasure to thank many people who have supported me during my PhD. Words will never be enough to tell my supervisor, Professor Michael Kisangiri to express my gratitude for his constant and valuable guidance during my PhD program, for the time he spent discussing and carefully reviewing all of the research work in this thesis. In a friendly working environment, I learned much more from him than just scientific concepts and his enthusiasm was essential to overcome any difficulties that arose during these five years. My sincere thanks go to my supervisor Professor Shubi Kaijage for his support and valuable suggestions. I am sincerely gratitude to Dr. Elizabeth Sylvester Mkoba, Dr. Judith Leo, Dr. Bonny Mgawe, for their untiring support and guidance in the making of this study. Their encouragement, know-how and mentorship have tremendously helped in moulding me into a research scientist. I want to express my gratitude to the dean of the CoCSE School, Dr. Mussa Ally Dida, Prof. Anael Sam, Prof. Dmitry Kuznetsov, Prof. Verdiana Massanja, Dr. Ramadhan Sinde, Dr. Neema Mduma, Dr. Devotha Godfrey Nyambo, Dr. Geminpeter Lyakurwa, Dr. Maranya M. Maranya and Dr. Silas Steven Mirau for their insightful remarks and supportive criticism during graduate seminars. To all my colleagues, staff and students at the Nelson Mandela African Institution of Science and Technology, you have made my time during my studies much more enjoyable, I am honored to be a part of this institution. Many thanks also to Fernuniversität in Hagen, Germany, for the opportunity gave to me to work with Professor Eric W. Steinhauer. Special thanks also go to all the people who made my two weeks at Mzumbe University and University of Dar es Salaam possible, research delivery partners, research module leaders, research support staff, research funders and everyone who contributed in any way. Thank you also for all the opportunities to meet important figures in the field of research; I would like to thank the internal and external reviewers of this thesis for their careful reading and valuable comments and suggestions. Unreserved thanks to the German Academic Exchange Service (DAAD) for granting me a three-year scholarship to continue my studies at NM-AIST. Special thanks also go to the DAAD representative of the NM-AIST, Dr. Michael Haule. Our sincere thanks go to Professor Steinmann Stephanie from the Fernuniversität Hagen, who has made a name for himself with this work and my skills, he is a special person; I like his friendly nature, which makes me feel comfortable in every communication. I appreciated the opportunity to work with the library administrators from Mzumbe University, Nelson Mandela African Institution of Science and Technology and the

University of Dar es Salaam. Last but not least, I am very grateful to the DAAD branch office in Nairobi for providing information and advice on study, research and funding opportunities. Thank you for your financial contributions to attend and present the conferences and workshops. I couldn't have done it without all of you I am truly thankful and may God bless you.

DEDICATION

I dedicate this research work to God Almighty. I also dedicate this work to my mother, my family, colleagues and friends for their love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND SYMBOLS

API	Application Programming Interface
DAAD	The German Academic Exchange Service
DDC	Dewey Decimal Classification
DOI	Digital Object Identifier
EAD	Encoded Archival Definition
EIRs	Electronic Information Resources
FRBR	Functional Requirements for Bibliographic Records
HTML	Hypertext Markup Language
ICT	Information and Communication Technology
ILS	Integrated Library System
IR	Information Retrieval
KOS	Knowledge Organization System
LAMP	Linux, Apache, MySQL, and PHP
MARC	Machine-Readable Cataloging
MU	Mzumbe University
NCIP	NISO Circulation Interchange Protocol
NISO	The National Information Standards Organization
NM-AIST	The Nelson Mandela African Institution of Science & Technology
OA	Open Access

OAI	Open Archives Initiative
OPAC	Online Public Access Catalog
PDF	Portable Document Format
Ph.D	Doctor of Philosophy
PHP	Hypertext Pre-Processor
PMH	Metadata Harvesting Protocol
QoS	Quality of Service
RDA	Resource Description and Access
RDM	Research Data Management
RDT	Resource Discovery Tool
SDI	Selective Information Dissemination
SIP	Standard Interchange Protocol
UDSM	University of Dar es Salaam
URL	Uniform Resource Locator
WCAG	Web Content Accessibility Guidelines
WWW	World Wide Web
XC	eXtensible Catalog
XML	eXtensible Markup Language

CHAPTER ONE

INTRODUCTION

1.1 Background of the Problem

The resources of the Digital Library are an essential part of the knowledge pit in daily life and are of great use to all educational institutions today. With the advent of information and communications technology (ICT), libraries have begun to use electronic resources, although student and general public usage of these resources is very low. Changes in the ICT worldwide have shifted the content of library resources from printed materials to online information resources [1], [2]. Digital library to researchers is defined as the network that provides ready access to the world's recorded knowledge, wherever it is held while to professional librarians, digital library is institution that arrange for the preservation, collection and organization of materials as well as for access to it [3]. The ease of access to electronic materials provides a new dimension of learning and hence has a significant impact on education. In addition, libraries use ICT to enable users to access and use online materials and other services, as a result, electronic library resources are widely used in many higher education institutions around the world, and many users can access them anytime and from anywhere, making their use extremely convenient [4]. Previous research has shown that university libraries use a variety of electronic resources and participate in a variety of local initiatives to increase awareness and use of these resources. Scholars, on the other hand, are unaware of the electronic library resources available, and students who are aware use library resources for teaching and research despite lacking the advanced search and retrieval skills required for proper use of library resources [5]**.** In Tanzania researchers investigated some of the challenges that university students face while trying to access electronic information resources, their outcomes indicate that surveyed participants were moderately aware of the library-subscribed e-resources, and the major barriers that caused under utilization were poor infrastructures, unstable network connectivity and lack of searching skills [6]–[8]. Inline with above findings data was gathered in three institutions; Nelson Mandela African Institution of Science and Technology (NM-AIST), Mzumbe University (MU), and the University of Dar es Salaam (UDSM) by assessing student usage patterns of e-library resources and initiatives established. Therefore, this work determined digital library accessibility obstacles in terms of information searching and retrieving aspects for label relevance and feature correlation as a

research gap. Hence, solutions were developed in this work to prevent library users from becoming victims of electronic library resources.

In order to develop user-friendly digital library collections this work proposed a Machine Learning Techniques to provide new insight into how digital library technology can be enhanced to deliver value-added innovative library services to achieve digital library transformation. It is also encourage library managers, stakeholders and information professionals to adopt ML to complement effective service delivery [9]. According to the literature review, infrastructure and technological tools are needed to assess the system's digital library's support services and instruct users on how to use them as well as to increase knowledge and improve use of electronic library resources.

1.2 Statement of the Problem

Higher education institutions are investing in electronic library resources to assist scholars and college members in their academic activities when they are looking for references and relevant materials in a timely manner. These higher learning institutions pay significant sums of money to library management each year to subscribe to electronic library resources in order to meet the information needs of their users as well as the teaching, learning, and research needs of their universities. It is only acceptable and economical that these resources are optimally used to contribute to scholars' educational achievements and thus get value for money spent by universities on electronic library resources. Despite the significance of electronic library resources and attempts to ensure that library users may access them, not all students and staff members are aware of them, and those that are aware do not make as much use of them as they deserve. The factors that affecting user satisfaction in a DL context can be categorized as ease of access, few download exceptions, and limitations, simplicity of the DL interface design, quality of interaction process, internet performance, quality assurance service, and ease in communication provided for by a social network. As a result, promoting student knowledge of the use of DL tools is no longer sufficient; researchers are presently working to increase the comprehension, availability, and ease of use of DL in this context. Therefore, a research study is required to find the solution for making the best utilization of electronic library resources in terms of information searching and retrieval attributes associated to label relevance and feature correlation segments.

1.3 Rationale of the Study

Academic institutions, via their library departments, invest heavily in e-library resources, including subscription fees, information management systems, awareness, and e-library resource sales, to ensure that users adapt to and use the resources. Electronic library users should think about and optimize their use of digital library resources to meet their information needs, so that the library department can reap the benefits of investing in these resources [10]–[12]. The efforts of educational institutions to increase student awareness and usage of digital library resources at the university level are insufficient; in this case, their researchers should work to increase digital library users' knowledge of digital library resources. Many studies have been conducted around the world regarding the awareness and use of digital library resources, some of which are cited in this work, which have identified many barriers that prevent students from directly accessing digital library resources. Although many university barriers are similar, many more problems are found in developing countries than in developed countries, primarily due to the inability to obtain or lack of technology required to access digital library resources. Significant barriers documented in the literature include a lack of awareness, a lack of information and digital skills, scholars' negative attitudes toward electronic resources, a lack of/poor internet connections, a lack of ICT infrastructure, information overload with a large amount of irrelevant information, and licensing of digital collections for educational purposes, as well as access restrictions in digital libraries.

1.4 Objectives

1.4.1 Main objective

The main objective of this research is to enhance the access and utilization of library resources using machine learning techniques.

1.4.2 Specific objectives

The specific objectives are as follows in order to achieve the overall goal of the study:

(i) To identify the requirements for enhancing the access and utilization of electronic information resources by the distance learners

(ii) To develop an accessible and user-friendly digital library system

(iii) To validate the performance of the developed digital library system

1.5 Research Questions

The research work was guided by the following research questions:

(i) What data required for enhancing the access and utilization of electronic information resources by the distance learners?
(ii) What technology and simplified features should be included for the development of accessible and user-friendly digital library system?
(iii) What accessibility performance level should be achieved by the developed digital library system?

1.6 Significance of the Study

Libraries are currently involved in digital content presentation, the organization of virtual academic libraries and open access, educational film production and management, providing online access to resources, and mass digitizing of print resources [13]. The presentation of books in e-book, audiobook, and text creation in PDF, HTML, and EPUB alternatives is one of the leading activities in the digital library environment [14]. To promote learning, research, and other events, all these programs related to the tradition of gathering and organizing library information should be conducted in organizations. In addition, the focus of user engagement with libraries is on library search tools as libraries continue to embrace technology. User interactions in library have been expanded from basic card catalogs to web-based catalogs, then tools for meta-searching, and finally to web-scale resource discovery tools (so-called RDTs) [15]. RDTs are considered as modern library catalogs that provide users with a Google-like search interface, give libraries statistics of their collections, and provide content providers with an alternative platform to encourage the use of resources. Due to its ability to use a single search box to search in-house and remote databases in a manner that is suitable for even experienced users, the RDT has become the preference of library. The developed RDT holds an index that represents the full text content which may be searched, and also offers a linking facility to open the full text content. This research work focuses on using the suggested ML techniques to facilitate the development of digital library resource discovery. Additionally, this work offers the following advantages: it offers a single point of entry for searching and accessing multiple sources of information, enabling the creation of a rapidly accessible and user-friendly digital library system, enabling readers to

obtain materials on demand, facilitating instant resource discovery for users, enabling customization, and preventing resource deterioration.

1.7 The Study area and Scope of the Research

1.7.1 Geographical Scope

The study reviewed literature globally and gathers data on awareness, accessibility, technology, and various library activities in one place which promote the use of digital library resources. But also, three universities were involved in the research fields; the University of Dar es Salaam, Mzumbe University, and the Nelson Mandela African Institution of Science and Technology. The researchers investigated the students' knowledge and use of electronic library resources at these three universities. Because these three universities use technology for e-library resources, it was reasonable to expect their students to participate fully in this study.

1.7.2 Content Scope

The content scope of this study expands knowledge of electronic library resource awareness and usage by developing a practical electronic library resource system using machine learning techniques to promote the effective use of e-library resources.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

The literature review chapter cover in detail the following subsections; theoretical framework, the quality of service in digital libraries, the satisfaction of users with digital library systems, design aspects of digital library resources, library discovery systems, machine learning text classification approaches and research gap.

2.2 Theoretical Framework

2.2.1 Unified Theory of Acceptance and Use of Technology (UTAUT)

It examines the acceptance of technology which is determined by the effects of four main factors, performance expectancy, effort expectancy, social influence and facilitating conditions. The original UTAUT framework was developed to explain and predict the acceptance of technology in an organisational context although, later it was tested in non-organisational settings too. Over the years, UTAUT showed wide application, which enhanced the generalisability of the theory. Given the variance of information communication technologies and the advances in the sector, a number of scholars extended UTAUT to adapt it to the context or improve its predictive power[16]. This study adopted this theory, which is useful for understanding system development, changes in operations, and system performance behavior.

2.2.2 Technology Acceptance Model (TAM)

It defines an information systems theory that describes how users come to accept and use technology. TAM is designed to measure the adoption of new technology based on customer attitudes. This was at a time when computers were being introduced into the workplace and Davis was looking for a way to predict and explain system use both for vendors and IT managers[17]. This theory, which enables users to consider system circumstances and environmental aspects holistically, was adopted by the researchers to help participants' better grasp how to embrace and use technology.

2.2.3 Information Processing Theory (IPT)

This theory is an approach to cognitive development studies that aims to explain how information is encoded into memory. It is based on the idea that humans do not merely respond to stimuli from the environment. Instead, humans process the information they receive. While experts believe that the brain's mechanisms and functions are relatively simple, the magnitude and scope of neural networks and their behaviors are quite powerful as a whole[18]. The theory was adopted as a result of the crucial cognitive process it provides for the encoding, storing, and retrieval of information.

2.3 Empirical Literature Review

The term electronic library resources define the information processed and digitally-driven using hardware and software that offer information that can be accessed by digital electronic users through remote information provider networks or mounted locally by digital library managers. Digital libraries are mainly designed to solve specific library problems, such as e-books, e-newsletters, e-references, theses, and dissertations [19]. In terms of teaching, training, and studying, online academic libraries are important components of any university's information infrastructure [20]. The development of DL resources is as a result of the need for people to keep their history, discoveries, and achievements records and collections. Academic institutions, through their library departments, make major investments for DL resources in the area of subscription fees, information management systems, awareness, and, marketing of DL resources to ensure clients are aware of the available tools and are actively using them [21]. Therefore, library users must know their information needs and make maximum use of the provided DL resources to appreciate the value of investments made by their libraries. It is not enough for academic institutions to raise students' awareness about the use of DL tools; it is now in this context that researchers are working to improve the understanding, availability, and use of DL. The factors affecting user satisfaction in a DL context can be categorized as ease of access, few download exceptions, and limitations, simplicity of the DL interface design, quality of interaction process, internet performance, quality assurance service, and ease in communication provided for by a social network [22], [23]. According to some studies the knowledge and use of DL around the world have revealed a range of barriers preventing learners from using online library resources [24]–[26]. Libraries worldwide are quickly transforming due to the ongoing growth and application of ICT [3]. It is important to note that while the barriers to accessing DL resources may be similar in different universities, there are

more in developing countries than in developed ones. The notable obstacles in literature are poor information and digital literacy skills, students' negative attitudes towards electronic tools, poor internet connectivity, poor ICT infrastructure, information overload, vast amounts of irrelevant information, licensing limitations on access to the DL collection, lack of generic e-resource portal interfaces, preference for print assets over electronic resources, discouraging e-resource use by academic staff, user authentication, download delay, lack of comprehensive ICT and searching skills among library staff, high cost of affordable online access and, low organizational budget for library departments [27]–[30].

Nonetheless, the absence of training for staff and other library users is a limiting factor in the accessibility and use of e-resources as users face difficulties in accessing services and the staff may lack adequate expertise to assist the users [31]. Among the e-library manager's obstacles is that the development of a DL requires substantial investment in capital, technology, and manpower to satisfy users. Therefore, library culture has a history of resource sharing due to budget constraints that most libraries face. A library organization embraces a common goal of resource sharing and cooperation, commonly referred to as a consortium.

Various studies have been done regarding the utilization of EIRs, most of these studies, however, have not focused on a particular group, especially in the case of Tanzania at the University of Dar es Salaam, the Nelson Mandela African Institution of Science and Technology, and Mzumbe University. The findings of this work are in line with Isibika who revealed that surveyed participants at MU were moderately aware of the library-subscribed e-resources, and the major barriers that caused underutilisation were unstable network connectivity and lack of searching skills. It was recommended that library managers should equip library users with intensive training on information searching skills to increase utilisation of the subscribed e-resources. Furthermore, it was strongly recommended that the library should market its subscribed e-resources to attract more users [6].

This study is an attempt to assess the awareness, utilization, and development of electronic resources in the library by the academic community of universities. Therefore, this research work represents a critical review of DL resources taking into account the practical aspects of library services. The work defines the criteria used in creating a DL for the university-level research community. As a result, the purpose of this research is to review and evaluate some research works in order to identify the usability of DL services and to develop an accessible and user-friendly DL discovery system to alleviate technological difficulties encountered by

online information learners.

The Quality of Service in Digital Libraries - The quality of service, it can be defined as the ability of an organization to provide a service and assist in its utilization[32]. With the advent of DL services, the quality of service (QoS) has become crucial in assessing the efficacy of service delivery. Diverse models and structures have been suggested to analyze the QoS in DL systems nevertheless, the concentration of reviewed work in QoS for the DL is about the extent of the user experience perspective which address different dimensions that are from the digital service providers. The models show that the level of QoS provided by digital service providers directly affects the perception and satisfaction of end-users.

The Satisfaction of Users with Digital Library Systems - User satisfaction assessment and preservation is an essential part of library quality management. Within the DL background, satisfaction refers to "feeling happy with the DL in helping to complete a task" [33]. For online library consumer, DL satisfaction is closely linked to their experience with DL services. By identifying the typical standards of DL providers the following components will be involved, (i) comprehensiveness–including everything, (ii) accessibility – everything available immediately, (iii) instant-gratification – response speed, (iv) software capacity–seamless, (v) user-friendliness – single interface, and (vi) various formats – wording, images, and audio [34]. On the other hand, when assessing DL, user satisfaction focusing on the interface and functionality offered by the DL [3]. See Fig. 1 showing users' satisfaction with digital libraries.

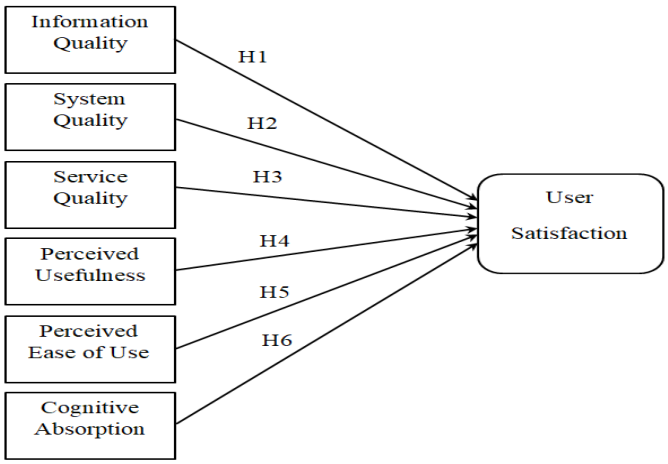

Figure 1: Users' Satisfaction with Digital Libraries [3]

(i) **Information quality** - Several academicians and researchers have recognized the importance of data reliability as a key component in the analysis of an effective computer-based data system. Information quality is essential in DL's point of view to support information needs. As a result, its attributes are usually associated with consistency, design, timeliness, currency, reliability, completeness, accuracy, and significance [35]. Therefore, the quality of information significantly affects the satisfaction of DL users.

(ii) **System quality**- System quality affects the perception of users of the performance of a DL in knowledge assortment and delivery [36]. In the development of information systems, the quality cycle of the systems is a strong determinant for user satisfaction in various contexts. Accessibility, accuracy, reliability, and quality are the key attributes of DLs performance measurement systems [36]–[39]. Quality, accessibility, and consistency can enable remote access of DL infrastructure to access information wherever and whenever, it also assures that the DL is accurate and functionally usable over time. When using the DL to search for information, effectiveness is associated with the quality, accuracy, and completeness of users. Considering the findings of Mesrek and Gaskin, the quality performance of the DL's software dramatically affects satisfaction, hence Fig. 1 is formulated to show that the quality of the system predicts users' satisfaction with the DL significantly [3].

(iii) **Service quality -** The User perception of the performance of a DL in the processing and distribution of information is characterized by service quality. As with information quality and applications, earlier studies show strong support for the contribution of the QoS to user satisfaction with DL. Different service quality models have been developed by scholars to test for the DL functionality such as DigiQUAL's and LibEval models [36], [40]. One of the prominent qualities of digital service performance is accession, reliability, accessibility, and responsiveness. Digital Library's service quality factor is very important as there is no face-to-face interaction in digital communication. For an open, healthy, sensitive, and well-integrated DL, there should be a favorable user satisfaction score and it can be concluded that the service quality has a big impact on user satisfaction with DL.

(iv) **Perceived usefulness -** Perceived usefulness is described as the degree to which a person archieved when using DL to improves academic performance [41]. Some

authors have adapted products used by researchers to measure the perceived value by operationalizing the individual net benefit of DL services and incorporating information system success model (ISSM) into the Technology Acceptance Model (TAM) [38], [42]. Several studies have shown that perceived usefulness influences the acceptance of DLs by users [43]–[46]. Other studies have however shown that the perceived value is a strong indicator of satisfaction [47], [48]. Users of DL need information and are searching for it, however, their satisfaction levels can be lowered if their experience of using DL is not useful or beneficial.

(v) **Perceived ease of use -** This is defined as "the degree to which an individual believes that it would be effortless to use a particular system" [41]. In this context, user-friendliness indicates a belief that using DL would require minimal effort. Literature also indicates that accessibility is sometimes related to ease of use [49]. When it is difficult to access a DL, users tend to consider it difficult to use and on the other hand, users are more likely to see it as user- friendly when it is readily-available [50]. Other library researchers found that online public access catalog (OPAC) satisfaction was related to ease of use [51]. Based on these findings, the perceived ease of use is thought to have a positive effect on satisfaction. To this end, it is assumed that user-friendliness significantly increases user satisfaction with the DL.

(vi) **Cognitive absorption** - The concept of cognitive absorption denotes the deep involvement or holistic interaction with the DL of a person. This concept was first introduced into the information system (IS) studies, to help users evaluate such systems. The first concept incorporated the facility of use as one of the building's dimensions [52]. Nevertheless, mental retention was also used in the modified and improved model of cognitive absorption as determinants of user-friendliness [3]. No one has ever tried to examine the effect of cognitive focus in the sense of DL on satisfaction. Since the DL is a subset of computer-based IS, its functionality is somewhat close to other forms of IS. Therefore, results on the effect of cognitive integration on the satisfaction of users from previous research should also be important for DLs. With this rationale, the cognitive absorption greatly increases user satisfaction with DL.

2.4 Design Aspects of Digital Library Resources

Technological aspects of designing DL resources concentrate on the following elements; ease of access, interface design, communication level, device reliability, and QoS [53]. Digital library resources variables are identified as the technical tools required for developing a DL for the research community and digital services based on the theoretical model, updated documents, and DL initiatives around the world [54].

Digital library resources structure - The digitization of documents created by the library and information studies of the research community is preferred since the goal is to integrate some of the features of the semantics DL, a room is given for sharing and collaboration in the writing of user-generated content. DL resources involve researchers' works as well as online records including; monographs, educational event promotional posters, online images, workshops, symposia, or colloquia, and visiting interviews with researchers. In librarianship and information research, DL systems have a direct relation with digital publishing. Services suggested by users and library managers are mobile information services, resource-increasing programs, group annotations, tagging, and the highest number of consultations. The mobile information services should also include Simple Syndication (RSS) and electronic media notification services [55]. The other component that is recommended is social network communication, which enables users to communicate with their library [56], see the proposed DL structure and services outlined in Fig. 2.

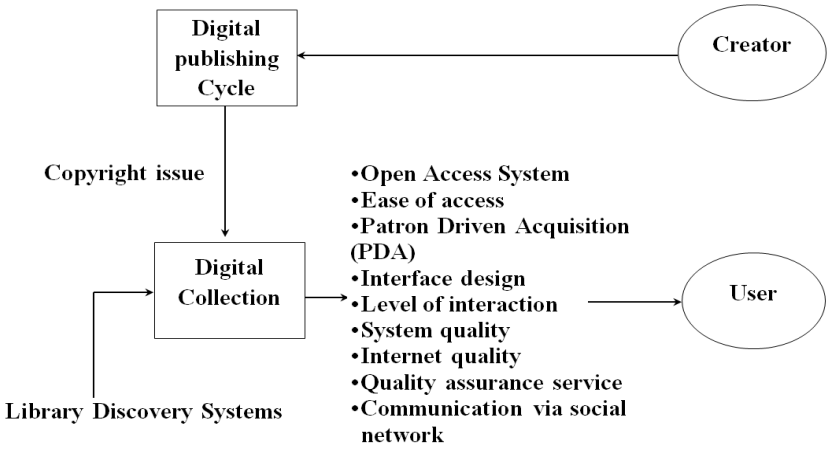

Figure 2: Digital Library Services Organization

Copyright Issues - Academic librarians have complex barriers and problems in managing the copyright process of the labyrinth. Some of the issues are evident, but others, especially in a digital environment, are much more subtle. Issues that cause confusion and concern about copyright law's applicability include e-reserves, registration, delivery, and equal use of information. The role of academic libraries is to offer access to information to their constituents anywhere and at any time [57].

Digital library creators and their materials, i.e. digital-born, digital turned, and acquired digital, should preserve copyright issues in the production and management of digital libraries in a legal framework to address copyright issues and protect the rights of content creators and owners. It would be a good idea to consider technical protection measures such as e-watermarking, digital signatures, authentication, etc. since DL are in the process of creating digitally born content to regulate infringement in the digital environment. In the current era of open access (OA) where information is made freely available to users, content producers should be worried about the infringement of copyrights. In addition to uploading digitized content on websites, libraries are digitizing to protect old, fragile documents and posterity. The DL is under contract or has been given licenses to use this material for digital content. Based on the license terms and conditions, the DL must sensitize and notify users, so that there is no total copyright infringement and a strong obligation to comply with the contract terms and conditions. Many owners of digital media copyright are found to rely on technological mechanisms to regulate access to digital files and recycle works. Since technology can provide protection systems, it also has the potential to create systems that can bypass this kind of protection [58]. Some of the rights management functions of computer systems could include tracking, identifying, and authenticating users, providing the copyright status of each digital object and limiting its use or charges, handling user transactions by only allowing access to not so many copies, or charging for copies, or transmitting requests to publishers (Cleveland, 2008).

Open access initiatives and adoption - Open Access to information is a generic term used for information tools made available for large-scale commercial or corporate use in the public domain, regardless of subscription fees or access fees [60]. Open access facilitators and users need to create an online link for the dissemination of information. The OA movement uses the term Gold OA for OA provided by publishers, Green OA for OA provided by libraries while work that is not OA or accessible only at a fee is called Toll

Access [61]. Thus, all kinds of OA remove price barriers and are termed as Gratis OA, as opposed to Libre OA if at least some permission barrier is removed as well.

According to research conducted at private universities in Nigeria, by Bolarinwa and Utulu, academic librarians have shown a positive view of the OA information tools [62]. The work done by Edda et al., revealed that about 75% of librarians had strongly supported OA topics on campus. The study suggested that the librarians can further support OA through i) providing enhanced access to OA works by linking OA works via library catalogs, including researchers' OA resources and Selective Information Dissemination (SDI) services, making them part of a federated search, ii) publishing OA works; many research organizations have in-house publications that can be called OA, which host the journal through Open Journal's systems and libraries can take steps to publish and identify potential outlets for publishing activities in different campus environments, ranging from newsletters to special collections and resources for students, iii) digitize OA versions of non-copyrighted works, and iv) initiate the October OA Week each year to mark the International OA Week which can include a variety of activities such as setting up talks, awarding researchers, and directing researchers on new developments in the field [63], [64].

2.5 Digital Library Discovery Systems Development Aspects

Development of DL discovery systems are created using technologies that enable library users to simultaneously search for multiple web-based resources and produce accessible search results. The present developments are completed through an open-source software with features that include organization and indexing, advanced searching, ranking and sorting, enhanced services, enhanced displayed, personal information environment, and support/integration with social networking tools [65]. The advantage of these open source software i.e blacklight is that are used by libraries to create discovery layers or institutional repositories to improve and update available technical frameworks to suit new digital materials [66]. Blacklight was first developed at the University of Virginia based on a Ruby on Rail programming framework and Apache SOLR indexing, search, and retrieval technology [67]. It also provides a flexible toolkit for a broad variety of record types and is the principal search interface to the hydra project digital asset management. The technological tools used for the creation of DL discovery system contents are i) Ruby, ii) Rails, and iii) Solr open source search platform, written in Java. Blacklight has several plugins. For instance, Blacklight MARC is library catalog enhancements, this is a spotlight

that enables librarians, and curators to create attractive and feature-rich. In the present development, the DL discovery system features are grouped into three categories, namely i) large search space, a search can find the data from various data sources, ii) intuitive usability, like a search engine, simple form with advanced search functions, iii) ranking of the results by relevance, the "best" hit is shown first, not necessarily the newest hit. One approach to support the development is the use of machine learning text classification to archive the aforementioned features.

Machine Learning Text classification (MLTC) is an effective and efficient technology for information organization and management during assigning of labels. MLTC is one of the latest digital transformation technological trends the university library can use to provide library users with alternative library services. MLTC can foster intelligent decisions for retrieving and sharing information for learning and research. However, extant literature confirms a low adoption rate by the university libraries in using MLTC to provide innovative alternative services, as this is missing in their strategic plan. In the literature, many feature selection methods and text classification algorithms have been proposed. However, the dramatic increase in the availability of massive text data from various sources is creating a number of issues and challenges for text classification such as scalability issues. As the amount of information available on the internet grows at a rapid pace, text classification becomes critical. This data is in an unstructured state and will need to be digitized. Due to the digital nature of these documents, data must be organized by automatically assigning a collection of documents to predefined labels based on their content. To mitigate the growing impact of new text classification, keyword detection approaches based on mostly supervised classification methods have been proposed. Therefore in this supervised classification method, statistical results classification rate, precision, recall and F-score for articles text classification should be validated to show that this approach significantly improve the performance [68]. Based on different machine learning algorithms used in the different works, the system of text classification is divided into four sections namely text pre-treatment, text representation, implementation of the classifier and classification. Authors, K. Shah, H. Patel, D. Sanghvi, and M. Shah designed BBC text classification system. In the classifier implementation section, the authors separately chose and compared logistic regression, random forest and K-nearest neighbour as classification algorithms. Then, these classifiers were tested, analysed and compared with each other and finally got a conclusion. The experimental conclusion shows that BBC text classification model gets satisfying results

on the basis of algorithms tested on the data set. The authors decided to show the comparison based on five parameters namely precision, accuracy, F1-score, support and confusion matrix. The classifier which gets the highest among all these parameters is termed as the best machine learning algorithm for the BBC text data set. Hence, the result shows that logistic regression classifier with the TF-IDF Vectorizer feature attains the highest accuracy of 97% for the data set [69].

Based on the text information processing, other study on the application of support vector machine in text categorization was taken. Through introducing the basic principle of SVM, a description on the process of text classification was performed and further proposed a SVM-based classification model. Finally, experimental data show that F1 value of SVM classifier has reached more than 86.26%, and the classification results comparing to other classification methods have greatly improved, and that proves SVM is an effective text classification machine learning method[70]. Other work proposes Like a Good Nearest Neighbor (LAGONN), an inexpensive modification to SetFit that requires no additional parameters or hyperparameters but modifies input with information about its nearest neighbor, for example, the label and text, in the training data, making novel data appear similar to an instance on which the model was optimized. LAGONN is effective at the task of detecting harmful content and generally improves SetFit's performance. To demonstrate LAGONN's value, authors conducted a thorough study of text classification systems in the context of content moderation under four label distributions [71].

Recent advancements in natural language processing (NLP) have reshaped the industry, with powerful language models such as GPT-3 achieving superhuman performance on various tasks. However, the increasing complexity of such models turns them into "black boxes", creating uncertainty about their internal operation and decision-making. Tsetlin Machine (TM) employs human-interpretable conjunctive clauses in propositional logic to solve complex pattern recognition problems and has demonstrated competitive performance in various NLP tasks. In the other study, ConvTextTM was proposed as a novel convolutional TM architecture for text classification. While legacy TM solutions treat the whole text as a corpus-specific set-of-words (SOW), ConvTextTM breaks down the text into a sequence of text fragments. The convolution over the text fragments opens up for local position-aware analysis. Further, ConvTextTM eliminates the dependency on a corpus-specific vocabulary. Instead, it employs a generic SOW formed by the tokenization scheme of the Bidirectional

Encoder Representations from Transformers (BERT). The convolution binds together the tokens, allowing ConvTextTM to address the out-of-vocabulary problem as well as spelling errors. This effort investigates the local explainability of proposed method using clause-based features. Extensive experiments are conducted on seven datasets, to demonstrate that the accuracy of ConvTextTM is either superior or comparable to state-of-the-art baselines[72].

Table 1: Table Matrix Comparison of Different Authors Addressing Solutions on Text Classification Using Machine Learning Techniques

Author	*Problem Adressed*	*Solution*	*Limitations*
[73]	Low character detection rate, difficulty in small character detection, and various character detection categories	An improved stacking model combining random forest and logistic regression	Accuracy training set 99% and Test set 92%, Not adressed feature recognition,
[74]	Imbalanced multilabel data	• Multilabel k-nearest neighbor (ML-kNN) • For the label of latent information carried by the nearest neighbors ✓ A value calculation method was proposed and used to directly classify majority labels • To address the misclassification problem caused by a lack of nearest neighbor information for minority labels ✓ weight calculation was proposed	kNN is not suitable for datasets with lower imbalance
[75]	Clear representation of the relationship	• Term Frequency – Inverse Document Frequency (TF-IDF) feature. A Label-Oriented (LO) approach	Accuracy average

	from a text to its assigned label	✓ This approach takes account of the relationship between a text and its assigned label by introducing a new feature label-oriented score. This score represents the level of the importance of the term regarding all terms and texts assigned to the label compared to all terms and texts unassigned to the label	test set of 97.80%, Not addressed multi-label classification
[76]	Problem as multi-class classification	• Hierarchical Deep Learning for Text classification (HDLTex)	Accuracies of 94% for the first level, 92% for the second level and 86% overall.

2.6 Research Gap

Using ICT has revolutionized how the worldwide community uses information. Library members are no longer limited to online tools designed to meet their information needs despite the challenges faced by scholars and the research communities when accessing online information. Some of these problems include difficulty in access, download exceptions and restrictions, complex interface design of information systems, low quality of the communication process, and poor quality of the internet. Taking into account the need for information, it is now that electronic library resources with multitasking information accessible, simple interface design, quality assurance service, and ease of communication via a social network should be developed and implemented to satisfy DL users. The case study was conducted in three universities; the Nelson Mandela African Institution of Science and Technology (NM-AIST), Mzumbe University (MU), and the University of Dar es Salaam (UDSM) by examining how students use e-library resources and novel initiatives. The aim to determine awareness of EIRs by library users at universities, to examine the accessibility of electronic library resources and services offered in the university libraries, to identify the problems encountered by the community users in accessing the e-library resources, and to draw the solutions regarding ease-of-use towards improving the accessibility of e-library resources. As a result, this work identified a research gap in digital library accessibility barriers related to information searching and retrieval attributes in terms of label relevance

and feature correlation segments. Therefore, in this work, solutions have been developed to prevent library users from becoming victims of electronic library resources.

2.7 Summary of the Chapter

This chapter has adopted three theoretical frameworks UTAUT, TAM and IPT, provides empirical literature review on digital library awareness, design, development and implementation. This review provides a systematic analysis of DL information accessibility in the following aspects: i) analyses the barriers preventing learners from using online library resources, ii) examines users' satisfaction with digital libraries, iii) studies essential variables related to DL adoption and usage, and iv) describes the next generation in DL services. Therefore, this review intended to provide a systematic analysis of DL information accessibility and the development of a user-friendly DL discovery system. The analysis shows that the library user's dimension plays the most significant role in DL systems improvement. This allows researchers to identify what aspects could be inherited from the other technologies specifically for the DL system enhancement. This analysis describes the different DL system models and tools, and the proposed design of library digital services that considers how to provide their patrons with QoS.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Introduction

This part presents the research philosophies, research techniques, and strategy that figure out data collection and analysis, followed by the methodology for constructing a digital library system with its evaluation and a chapter summary.

3.2 Research Philosophies

Acceptable research philosophies were selected using a suitable selection guide for research approaches; positivism research philosophy was adopted that uses quantitative research technique and interpretivism research philosophy was taken which uses qualitative based methodology to control this research.

3.3 Research Methods and Strategy

3.3.1 A Synopsis of the Research Process

The research roadmap incorporates the following main phases; i) data collection and analysis ii) digital library system development and iii) digital library system performance evaluation, see Fig. 3 an overview of the research workflow.

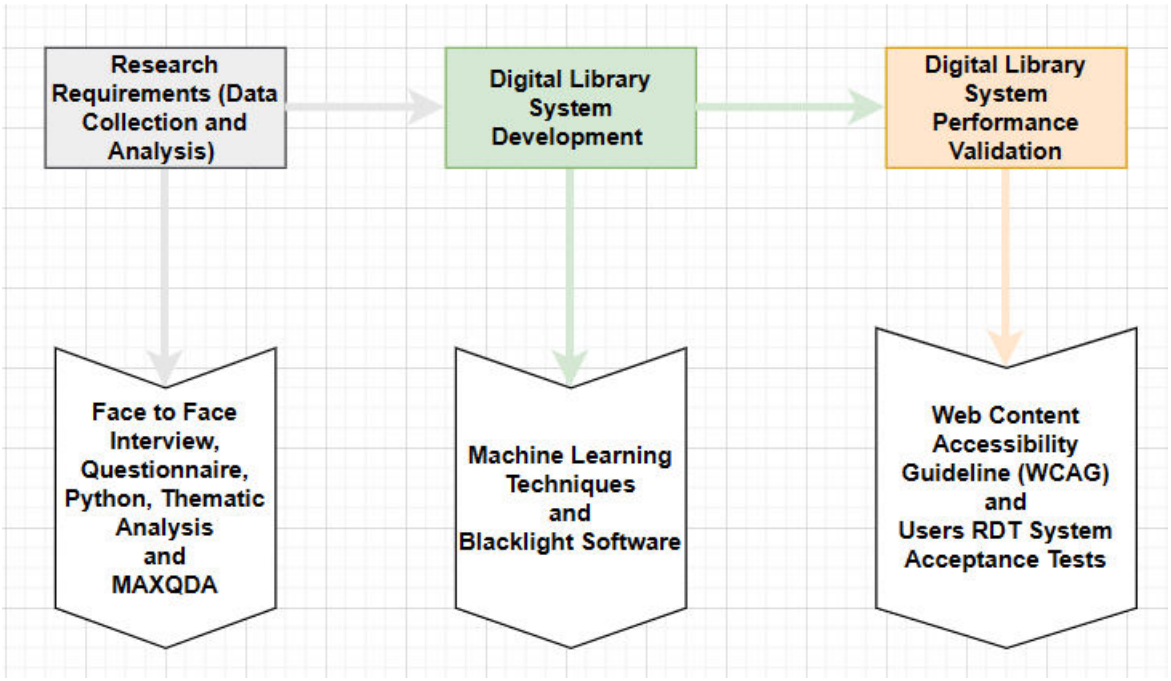

Figure 3: An Overview of the Research Workflow

The first phase's goal was achieved by using a classification analysis method that included qualitative as well as quantitative aspects. Quantitative data was gathered through issues concerning access to personal computer, internet, e-library resources, learning opportunities, technical infrastructure, and student marketing of e-library resources. In-depth research and analysis on the subjective and personal experiences of people who use digital library resources were conducted using a qualitative style. An online questionnaire was used to collect research data then analyzed with Python tool, the number of library users was 5258 from three universities; 4888 from the University of Dar es Salaam, 250 from Mzumbe University, and 120 from NM-AIST. Professionals were included in the study to examine local initiatives of library departments to increase student and department staff awareness and use of e-library resources.

The design and development of accessible and user-friendly digital library system as a second goal was accomplished through Agile Scrum methodology using Proposed an Ensemble Model (PEM) machine learning technique which integrated logistic regression, support vector machine and knearest neighbor machine learning techniques. The comparison was made with other platforms to uniquely show the new features incorporated and the merits of the work.

In order to find out the performance of the developed system the four digital libraries (Developed RDT, NM-AIST, MU and UDSM) have been successfully tested for the compliance using Web Content Accessibility Guidelines (WCAG) 2.1 standards thereafter users RDT system acceptance tests were conducted.

3.3.2 Data Collection and Analysis

The study reviewed literature globally and gathers data on awareness, accessibility, technology, and various library activities in one place which promote the use of digital library resources**.** Therefore, the research work implemented a case study strategy to gather data at the Nelson Mandela African Institution of Science & Technology (NM-AIST), the University of Dar es Salaam (UDSM), and Mzumbe University (MU). The survey aimed to assess the degree to which the academic community uses electronic information resources (EIRs). The present study employed both quantitative and qualitative approaches. The quantitative approach involved use of questionnaire while the qualitative approach involved unstructured interviews, casual talks, and in-depth discussions. A semi-structured questionnaire was designed and distributed as a data collection tool. The questionnaire was divided into three main categories. The primary part was an awareness of EIRs, the second on how one gets to access electronic library resources, and the third was about the issues and challenges faced when using the online searching platform to access electronic library materials and references. A 19-questions survey was used to investigate the extent of use of EIRs by distance learners. The library users that responded were 1667 drawn from the three universities UDSM with 1350 respondents, NM-AIST with 100 respondents, and MU with 217 respondents. Then thematic and MAXQDA analysis were performed on the qualitative data while python tool was used for the quantitative data analysis.

(i) Population, Sampling Techniques and Sample Size

The entire population in selected institutions was very large about 41589 faculty members from the University of Dar es Salaam (UDSM) with around 32589 faculty members, the Nelson Mandela African Institution of Science and Technology (NM-AIST) having around 800 faculty members, and Mzumbe University (MU) accommodates around 8200 faculty members. A reasonable, convenient, and manageable sample size of 5258 library users was randomly selected from these institutions for the study. The sample included 4888 users from UDSM, 120 users from NM-AIST, and 250 users from MU.

The minimum sample size was decided following the Kish formula (Equation 1) for cross-sectional studies [77]. The calculation was based on the library users' prevalence of 32% and the whole population of the universities of about 41589 people at the precision of 50, and a confidence level of 95%. With this method, the minimal sample size was determined to be 332 people, and 1667 faculty members participated in the survey which was a sufficient number to produce the desired results.

$$n' = \frac{NZ^2P(1-P)}{d^2(N-1)+Z^2P(1-P)} \quad \text{Equation 1}$$

Where n´ = Sample size with finite population correction,

N = Population size,

Z = Statistic for a level of confidence,

P = Expected proportion (prevalence),

d= Absolute error (Precision)

3.3.3 Status of Respondents

Table 2: Universities Participated Within the Survey Questionnaire, the Population, the Sample Size, Number of Library Users Responded, and Percentage of Respondents

No.	Name of the University	Population	Sample Selected	Respondents (res)	Percentage of Respondents ((No. of res/ Total res)*100)
1	University of Dar Es Salaam	32589	4888	1350	81
2	The Nelson Mandela African Institution of Science and Technology	800	120	100	6
3	Mzumbe University	8200	250	217	13
	Total	41589	5258	1667	100

A total number of 1667 library users responded to the questionnaire; the number of library users responded from each university by percentage is as shown in (Table 2). UDSM was 1350 (81%), NM-AIST 100 (6%), and MU 217 (13%).

(i) Gender

A total of 1184 (71%) of the respondents were males while 483 (29%) were females (Table 3).

Table 3: Number of Male and Female Library Users Who Responded from the three Universities; UDSM, NM-AIST, and MU

		Name of the University				
Rank	Gender	UDSM	NM-AIST	MU	Total	Percentage
1	Male	960	71	153	1184	71
2	Female	390	29	64	483	29
	Total	1350	100	217	1667	100

(ii) Qualifications

Regarding the distribution of the status of the respondents, the analysis showed a majority 1100 (66%) were at Master's level, followed by Bachelor's at 300 (18%), Ph.D. holders at 250 (15%), and Diploma 17 (1%) see Table 4.

Table 4: The Distribution of the Status of Respondents at UDSM, NM-AIST, and MU

		Name of the University				
Rank	Membership Category	UDSM	NM-AIST	MU	Total	Percentage
1	Ph.D	202	17	31	250	15
2	Master's	966	41	93	1100	66
3	Bachelor	167	42	91	300	18
4	Diploma	15	0	2	17	1
5	Certificate	0	0	0	0	0
	Total	1350	100	217	1667	100

(iii) Group Title

A huge majority of the respondents, 1384 (83%) were students, 217 (13%) teaching staff, and the rest 66(4%) were non-teaching staff (Table 5).

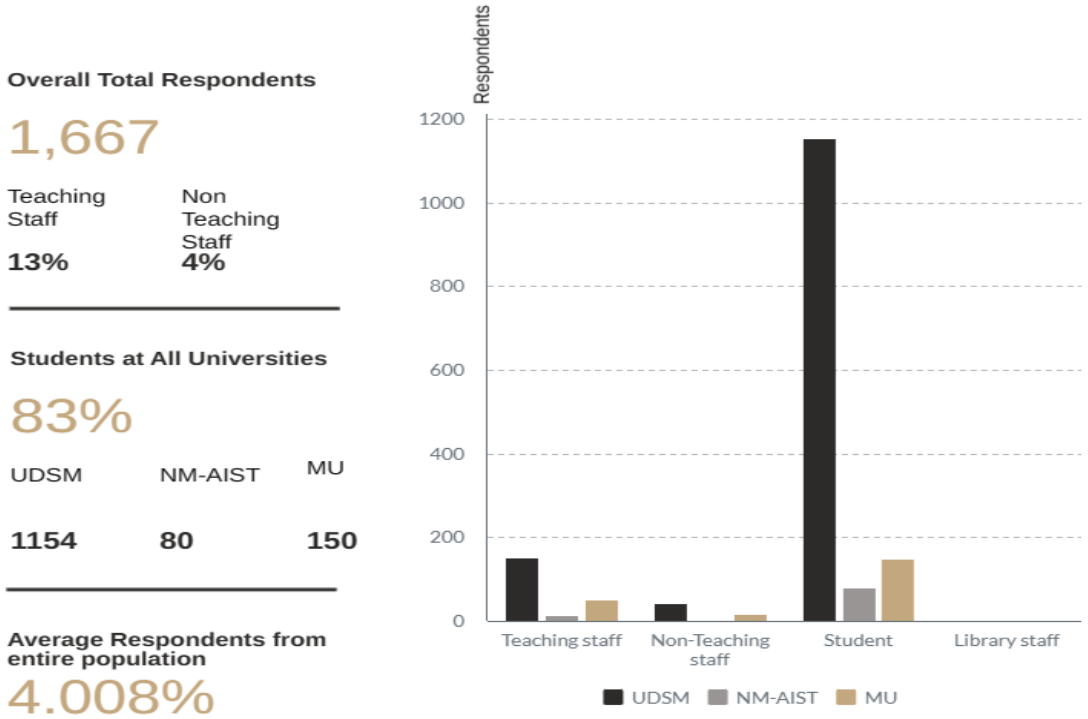

Table 5: Group Title Percentages of Respondents at UDSM, NM-AIST, and MU

	Name of the University				
Group title	UDSM	NM-AIST	MU	Total	Percentage
Teaching staff	153	13	51	217	13
Non-teaching staff	43	7	16	66	4
Student	1154	80	150	1384	83
Library staff	0	0	0	0	0
Total	1350	100	217	1667	100

(iv) Faculty

The faculty distribution of the respondents in the study is shown in Table 6. The analysis showed that 350 (21%) respondents were affiliated with Science and Engineering. Other respondents 317 (19%) were from the faculty of Engineering and technology followed by the respondents from Faculties of Science and Technology 300 (18%). Natural and applied science

had 217 (13%), Health and Allied Science133 (8%), Social Sciences 117 (7), Agricultural Science and Fisheries 83 (5%), Law 83 (5%), Education 33 (2%), Business and Development Science 17 (1%), and Resource Assessment 17 (1%).

Table 6: Faculty Percentages of Respondents at UDSM, NM-AIST, and MU

Rank	Discipline area	Total respondents	Respondents Percentage
1	Science and technology	300	18
2	Engineering and technology	317	19
3	Science and Engineering	350	21
4	Natural and Applied science	217	13
5	Humanities	0	0
6	Social Sciences	117	7
7	Agricultural Science and Fisheries Technology	83	5
8	Health and Allied Science	133	8
9	Business and Development Science	17	1
10	Education	33	2
11	Law	83	5
12	Journalism and Mass Communication	0	0
13	Kiswahili Studies	0	0
14	Resource Assessment	17	1
15	Marine Studies	0	0
Total		1667	100

As previously stated, the initial phase (data gathered above) was the collection and analysis of research requirements, and then followed by the design and development of a digital library system.

3.3.4 Methodology for Developing a Digital Library System

Agile Scrum methodology was used to design and construct an accessible and user-friendly digital library system. In order to provide higher quality methods, this study adopted an agile scrum methodology for framework development with regard to proper engineering

procedures for system development. Agile Scrum methodology is powerful framework that can help team improve agility, collaboration, and product quality. Therefore, this work used supervised machine learning which involves an output label associated with each instance in the data-set refers as text classification. Hence the proposed ensemble model (PEM) for machine learning approaches was developed that described label relevance and feature correlation. The proposed ensemble model (PEM) was created by combining three models: logistic regression (LR), support vector machine (SVM), and kNN algorithms. Finally, PEM ML technique was integrated with Blacklight software to develop resource discovery tool (RDT) that holds an index which may be searched, and also offers a linking facility to open the full text content.

(i) Proposed Ensemble Model (PEM) Development for Text Classification Using Machine Learning

In order to archieve the model development for text classification using machine learning the following components were used during digital library development see Table 7.

Table 7: Components of digital library development

Target	Dataset & Tools	Process & Purpose
Framework for automated subject indexing	Python Virtual Environment (Python 3.8.13 version)	Requires to install and configure Python virtual environment with Python (3.8+) including components like TensorFlow and Gensim.
Vocabulary dataset preparation	Linked Open Dataset from Web of Science (WOS-46985 Dataset)	The SKOS-compliant LCSH in TTL format is deployed to develop the backend KOS for the framework.
Accessibility performance	Accuracy, precision, recall, and F1 tests	Digital library test for the compliance score

(ii) Preparing the Training Dataset

After the vocabulary feeds inside the backend KOS, it requires training to ensure efficient prediction of subject descriptors against a text (usually a combination of the title of a

document and its abstract or summary note, separated by space or any other character. The backend KOS requires a training dataset as a TSV file with the first column containing a text corpus and the second column containing the URIs of the subject descriptors from LCSH for MARC creator. The MARC files obtained are then split into smaller MARC files (.mrc) for the sake of easy handling, by utilising the MARCSplit option in the tool MARCEdit. Each of these MARC files is then exported into an OpenRefine-compatible format (.tsv) from MARCEdit for further processing. OpenRefine, is an open-source data wrangling software, allows selecting only the rows having certain tags. This sort of selective display is needed to reconcile human-indexed subject descriptors.

Therfore, in this work data were sourced from Web of Science (WOS) dataset; this WOS is a document classification dataset that contains 46,985 documents with 134 categories which include 7 parents' categories. The Web of Science dataset is a collection of data that encompasses published papers available from the Web of Science. WOS has been released in three versions: WOS-46985, WOS-11967 and WOS-5736. WOS-46985 is the full dataset while WOS-11967 and WOS-5736 are two subsets of WOS-46985. Dataset was imported into Python using Pandas; data cleaning was perfomed through applymap function technique and feature selection using fisher score. Before training and testing, all special characters should be cleaned out of data sets so as to increase the performance of developed machine learning model. Then the data was splitted into train and test data model, see Fig. 4 representing conceptual framework for proposed an ensemble model flow text classification.

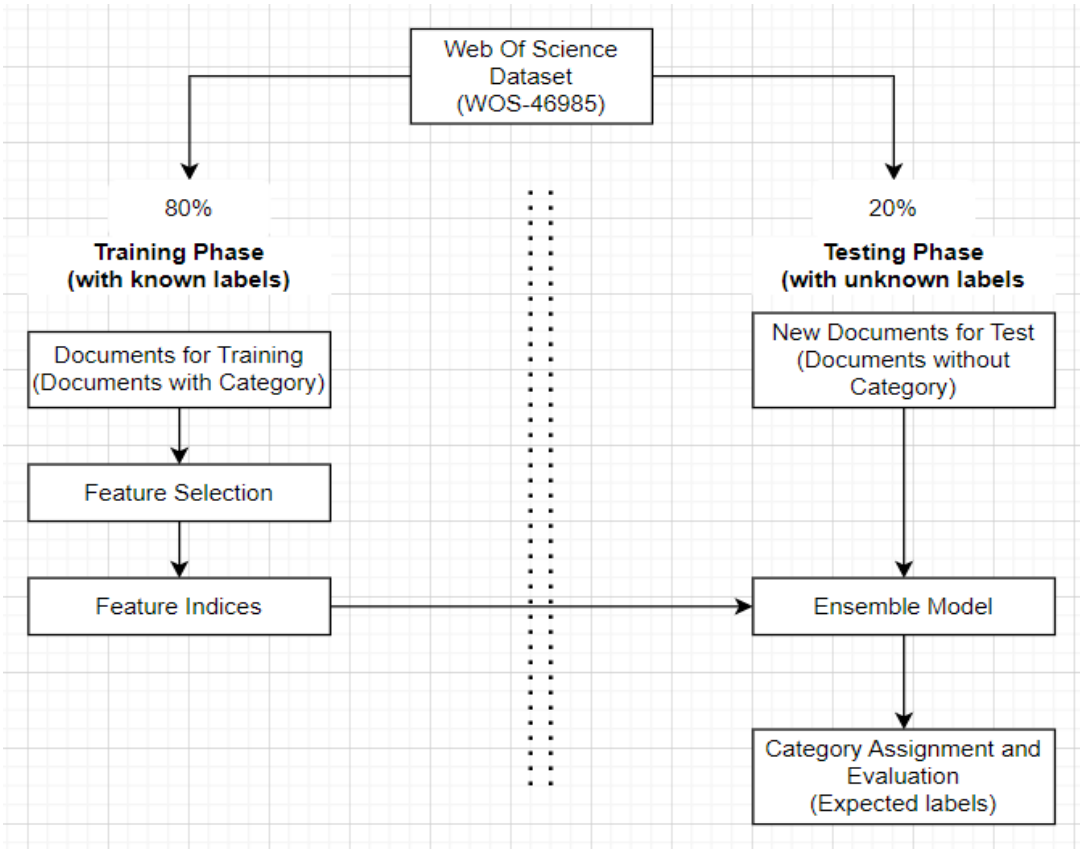

Figure 4: Conceptual Framework for Proposed an Ensemble Model Flow Text Classification

As explained earlier WOS has three datasets which include WOS-11967, WOS-46985, and WOS-5736 folders with 7 parents' categories. Each folder contains: X.txt, Y.txt, YL1.txt, and YL2.txt. X is input data that include text sequences, Y is target value, YL1 is target value of level one (parent label), YL2 is target value of level one (child label). Additional attribute in WOS-46985 are Domain, area, keywords and abstract. Abstract is an input data that include text sequences of 46,985 published papers that are free accessible from the Web of Science see Fig. 5 the sinipped file WOS-46985.

	Y1	Y2	Y	Domain	area	keywords
2	0	12	12	CS	Symbolic computation	(2+1)-dimensional non-linear optical waves; erbium-doped optical fibre; symbolic computation; soliton solution; soliton interaction
3	5	2	74	Medical	Alzheimer's Disease	Aging; Tau; Amyloid; PET; Alzheimer's disease; Polypathology
4	4	7	68	Civil	Green Building	LED lighting system; PV system; Distributed lighting control; Energy efficiency; Green building; Daylight responsive dimming system
5	1	10	26	ECE	Electric motor	NdFeB magnets; Electric motor; Electric vehicle; Hybrid electric vehicle; Recycling; Rare earth elements
6	5	43	115	Medical	Parkinson's Disease	Parkinson's disease; dyskinesia; adenosine A(2A) receptor; striatum; glutamic acid decarboxylase
7	0	0	0	CS, ECE	Computer vision	Tea category identification; computer vision; color histogram; wavelet packet entropy; winner-takes-all; radial basis function; artificial r
8	6	0	125	biochemis	Molecular biology	Brain-machine interfaces; Electrical stimulation of the brain; Transcranial magnetic stimulation; Trends; Psychiatry
9	1	10	27	ECE	Satellite radio	Software-defined radio; AIS receiver; VHF data exchange services; Satellite radio applications; Flexible RF front end
10	3	4	56	MAE	Fluid mechanics	plumes/thermals; turbulence modelling; turbulence simulation
11	2	13	46	Psycholog	Prenatal development	Stomach of human fetus; Development; Endocrine cells; Immunohistochemistry
12	5	51	123	Medical	Sprains and Strains	National Electronic Injury Surveillance System; water sports; trauma
13	6	5	130	biochemis	Enzymology	enzymology; factor IXa; hemophilia; structure; substrate; thrombosis
14	6	8	133	biochemis	Southern blotting	bioinformatics approaches; causal missense mutation; gene F8; genetic counselling; haemophilia A; HAMSTeRS database
15	6	8	133	biochemis	Southern blotting	age-related macular degeneration; genetics; retina; telomere
16	1	10	27	ECE, CS	Satellite radio	statistical tomography; high-latitude ionosphere; kilometer-scale irregularities
17	5	38	110	Medical	Cancer	Fluorocholine; Positron emission tomography; Hepatocellular carcinoma; Phosphatidylcholine; Fatty acids
18	6	5	130	biochemis	Enzymology	[FeFe]-hydrogenase; kinetics; enzymology; electron bifurcation; time-resolved infrared spectroscopy
19	6	7	132	biochemis	Northern blotting	nonalcoholic fatty liver; Chinese herbal; uncoupling protein 2
20	5	50	122	Medical	Sports Injuries	Elastic therapeutic tape; Material properties; Tension test
21	5	50	122	Medical	Sports Injuries	Sports injury; Athletes; Postural stability
22	5	48	120	Medical	Senior Health	Ankle range of motion; Balance; TETRAX

: : : : : :

46984	5	38	110	Medical	Cancer	Antifouling biosensor; Peptide; Electrochemistry; PEDOT; Cancer biomarker; DNA
46985	0	10	10	CS	Distributed computing	High Performance Computing; Parallel Computing; Distributed Computing; Automation; Image Matching; DSM Extraction
46986	6	0	125	biochemis	Molecular biology	gene expression; MPK3; plant-host interactions; septoria; wheat; Zymoseptoria tritici

Figure 5: The folder Web of Science Dataset WOS-46985

From Fig. 5; Y is target value, the target value has two levels, YL1 is target value of level one (parent label), YL1 = {1, .., 7} where YL1 = { Computer Science, Electrical and Communication Engineering, Psychology, Mechanical and Aerospace Engineering, Civil Engineering, Medical Science, Biochemistry} and YL2 is target value of level one (child label), which contain {17, 16, 19, 9, 11, 53, 9} specific topics belonging to YL1, respectively. To train and test the baseline and document classification methods, we collected data and meta-data on 46, 985 published papers available from the Web Of Science as elaborated above. We specifically extracted the abstract, domain, and keywords of this set of published papers. The text in the abstract is the input for classification while the domain name provides the label. The keywords provide the descriptors for the next level in the classification hierarchy. See Table 8 shows statistics for this collection, for example, Medical Sciences is one of the top-level domain classifications and there are 53 sub-classifications within this domain.

Table 8: Details of the Document Set Used in this Research

Domain	Number of Document	Number of Area
Biochemistry	5,687	9
Civil Engineering	4,237	11
Computer Science	6,514	17
Electrical Engineering	5,483	16
Medical Sciences	14,625	53
Mechanical Engineering	3,297	9
Psychology	7,142	19
Total	46,985	134

We divided the data set into three parts as shown in Table 9, data set WOS − 46985 is the full data set with 46,985 documents, and data sets WOS−11967 and WOS−5736 are subsets of this full data set with the number of training and testing documents shown as well as the number of labels or classes in each of the two levels. For dataset WOS − 11967, each of the seven level-1 classes has five sub-classes. For data set WOS − 5736, two of the three higher-level classes have four sub-classes and the last high-level class has three sub-classes. Before training and testing, all special characters were cleaned out of the three data sets.

Table 9: Details of three Data Sets used in this Research

Data Set	Training	Testing	Level 1	Level 2
WOS-11967	9574	2393	7	35
WOS-46985	37588	9397	7	134
WOS-5736	4589	1147	3	11

Text classification using machine learning, there a few drawbacks on how difficult it is to select the best ML models and time-consuming for tuning different model parameters to achieve better accuracy, to overcome this problem this work opted python library "Lazy Predict". This module helps us in determining the most accurate classification algorithm based on our data taken from web of science dataset. It offers a Lazy Classifier for classification issues; it has a high computational capacity and takes little time to process high dimensional data with numerous features see the results in Fig. 6.

Model	Accuracy	Balanced Accuracy	ROC AUC	F1 Score	Time Taken
LogisticRegression	0.998786	0.986219	0.981235	0.989477	0.0151712
SVM	0.988867	0.987783	0.987984	0.989454	0.0109444
KNearest Neighbor	0.985994	0.986878	0.986911	0.985878	0.671
Adaptive boosting regression	0.985973	0.984545	0.984767	0.985965	0.0120005
LinearSVC	0.985968	0.98269	0.98269	0.985934	0.0200432
LogisticRegressionCV	0.985965	0.98269	0.98269	0.985934	0.262997
SVC	0.982456	0.979942	0.979942	0.982437	0.0140011
CalibratedClassifierCV	0.982456	0.975728	0.975728	0.982357	0.0350015
PassiveAggressiveClassifier	0.975439	0.974448	0.974448	0.975464	0.0130005
LabelPropagation	0.975439	0.974448	0.974448	0.975464	0.0429988
LabelSpreading	0.975439	0.974448	0.974448	0.975464	0.0310006
RandomForestClassifier	0.97193	0.969594	0.969594	0.97193	0.033
GradientBoostingClassifier	0.97193	0.967486	0.967486	0.971869	0.166998
QuadraticDiscriminantAnalysis	0.964912	0.966206	0.966206	0.965052	0.0119994
HistGradientBoostingClassifier	0.968421	0.964739	0.964739	0.968387	0.682003
RidgeClassifierCV	0.97193	0.963272	0.963272	0.971736	0.0130029
RidgeClassifier	0.968421	0.960525	0.960525	0.968242	0.0119977
MLPClassifier	0.961404	0.959245	0.959245	0.961444	0.204998
ExtraTreesClassifier	0.961404	0.957138	0.957138	0.961362	0.0270066
SGDClassifier	0.961404	0.95503	0.95503	0.961276	0.0560005
BaggingClassifier	0.947368	0.954577	0.954577	0.947882	0.0559971
BernoulliNB	0.950877	0.951003	0.951003	0.951072	0.0169988
LinearDiscriminantAnalysis	0.961404	0.950816	0.950816	0.961089	0.0199995
GaussianNB	0.954386	0.949536	0.949536	0.954337	0.0139935
NuSVC	0.954386	0.943215	0.943215	0.954014	0.019989
DecisionTreeClassifier	0.936842	0.933693	0.933693	0.936971	0.0170023
NearestCentroid	0.947368	0.933506	0.933506	0.946801	0.0160074
ExtraTreeClassifier	0.922807	0.912168	0.912168	0.922462	0.0109999
CheckingClassifier	0.361404	0.5	0.5	0.191879	0.0170043
DummyClassifier	0.512281	0.489598	0.489598	0.518924	0.0119965

Figure 6: Results from LazyClassifier for ML Text Classification Model Performance

The best three models from the above LazyClassifier results were chosen and employed in this research, which are Logistic Regression, SVM, and KNearest Neighbor, as shown in Fig. 7; of the proposed an ensemble model (PEM) text classification. Boosting algorithm technique was employed since it is a powerful technique in machine learning that aims to improve the predictive accuracy of the model by combining multiple weak.

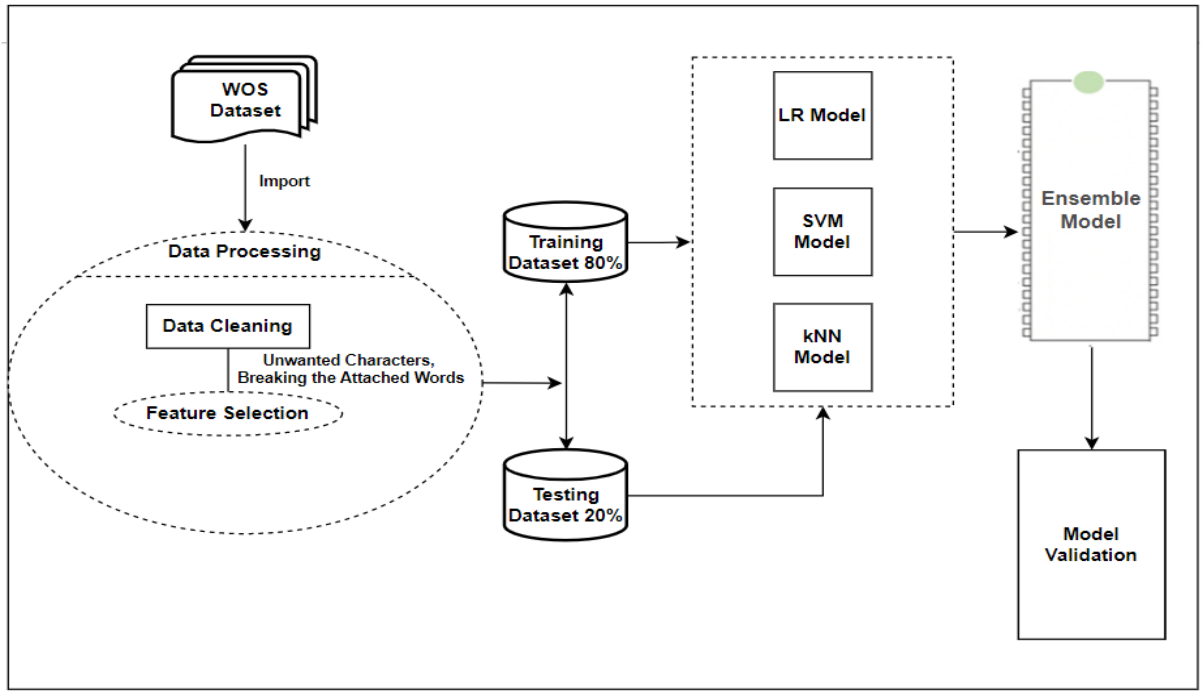

Figure 7: The Proposed an Ensemble Model (PEM) Text Classification

During the proposed an ensemble model (PEM) training set text classification, 80% (37588) of the WOS 46985 dataset was used, parent labels were classified into seven groups, as illustrated in Fig. 8; and the result was a category with 9303 entries classified into the aforementioned seven groups, as shown in Fig. 9.

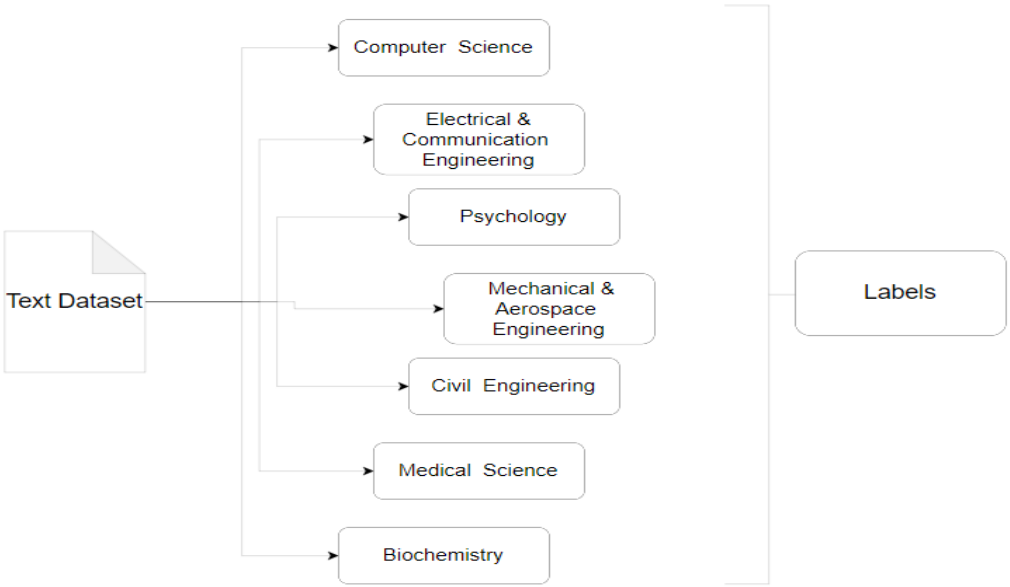

Figure 8: Seven parents' categories

Thus, WOS 46985 dataset was classified into the aforementioned 7 categories through the proposed an ensemble model (PEM). The outcome was a category with 9303 entries to two columns for text and label. See the resulting snippets (Fig. 9); which can be referred to this dataset as a dataframe (df).

	Content	Label
675	Learning Variational Word Masks to Improve the Interpretability	Computer Science
322	Can a Fruit Fly Learn Word Embeddings?	Computer Science, ECE
776	Electric distribution network multi	ECE
238	The biochemistry and physiology of S-nitrosothiol	Biochemistry
879	Biology and biochemistry of papillomaviruses	Computer Science
455	Observational research, randomised trials, and two views of medical.....	Medical Science
685	Neural networks in civil engineering. I: Principles and understanding.....	Civil Engineering

9303 rows x 2 columns

Figure 9: Showing Output category of 9303 Entries to two Columns for Text and Label Classification from WOS 46985 Dataset

Then the training set was carried out as well for the WOS-11967 and WOS-5736 data sets, and the results reveal inaccurate when compared to WOS 46985; this concludes that PEM performed better on larger data sets than smaller data sets. Additionally, testing was performed on these three data sets, and the evaluation model performance for WOS 46985 data set was remarked.

3.4 Integration of Proposed Ensemble Model (PEM) with Blacklight Software for Development of RDT

With different open-source software standards, a Digital Libraries Resource Discovery Tool (RDT) was developed. This case was focused on the library resource discovery procedure, which evaluates the methods in several configurations. It is essentially a method of discovery of resource exploration that can be incorporated with any online information retrieval (IR) framework [65]. This session shares the work experience of creating RDT using Blacklight software, an integrated library system (ILS), and a DSpace repository system framework as

well as other independent commercial and open-access external databases. Blacklight is a highly customizable open-source Web-scale Resource Discovery (WSRD) product that allows information repositories and integrates multiple services into a unified interface. It has powerful Apache Solr as search platform in order to showcase any digital content. Data (in Marc format) for the sample test was exported from the bibliographical databases (WOS into Koha Version 19.11.00) and then in Blacklight. At different layers and stages of implementation, a number of open-source software, protocols, and open standard technologies were used within the assembly. There are basically three layers: the greenstone LAMP architecture (Linux-Apache-MySQL-Ruby on Rails) is provided by Layer-I (core layer), MySQL is included like an electronic database (release 5.7), an Apache web server (version 2.4), a Ruby on Rails programming/scripting language (Ruby version 2.5.8 and Rails version 6.0.0) and utilizes Linux Ubuntu-18.04 LTS as it's an operating system. Layer-II (full-text layer) uses Apache Tika as a full-text extractor (version 1.18) and Apache Solr (version 6.1.0) to index the harvested metadata from various subscribed or external sources. Layer-III (front layer) is just an interface adapted to give users value-added services with Blacklight. See Fig. 10; illustrates the scheme of obtaining bibliographic data from WOS external sources of discovery. As stated in this work, the Discovery system model focuses on the collection of Koha formatted records from machine-readable cataloging (MARC 21) by setting up the ILS's with Open Archives Initiative (OAI) compliant function.

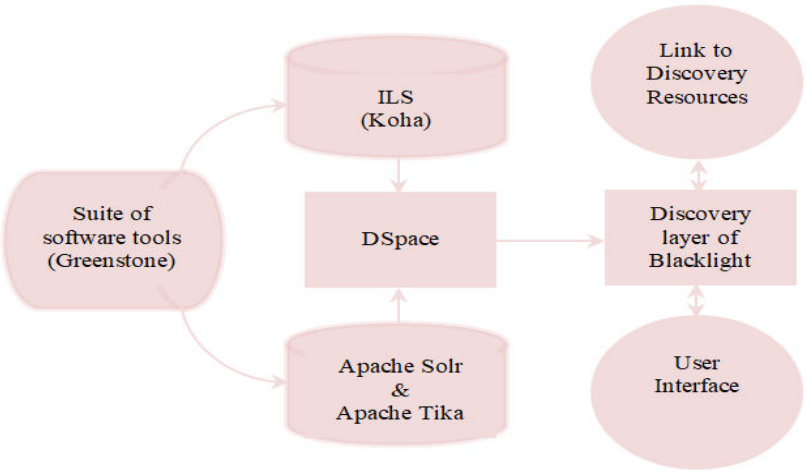

Figure 10: Functions of the discovery system model

The seamless integration of a discovery system with online IR Z39.50 system as well as its interoperability was important for its success, and protocols such as the Standard Interchange

Protocol (SIP) and NISO Circulation Interchange Protocol (NCIP) (i.e. NISO stands for The National Information Standards Organization) needed to be established. Koha OAI-PMH (i.e. PMH stands for the Metadata Harvesting Protocol) would be the prime logical method of integrating blacklight with any ILS. The file *Koha.ini* and *config.ini* must be configured both to sync the Koha driver with Blacklight in the next process [78]. The following logical step was engaged in making the changes to Koha during a *MARC_local.properties* file to import MARC documents from Koha to Blacklight. By creating a separate section named that source, *oai.ini* files had to be configured from any OAI-PMH source to collect and index Blacklight data.

The harvest was followed by the batch import of collected records into the invention sheet's indexing process. In Koha, Tag should be loaded in for full-text indexing by uploading full-text documents in any form and scale. Apache Tika must be installed and configured properly for this purpose and the link must be generated within the *fulltext.ini* file. On top of Koha, the Blacklight layers serve as a major index, an interface, or as an optional portal. The following configuration requires more steps to import DSpace metadata into Blacklight, and to render OAI-PMH compatible with every other repository's framework, for which two files must be synchronized, such as *dspace.config* file, and the *oai.cfg* file. To line up an OAI-PMH harvester, a separate DSpace file was recognized in the *oai.ini* registry. When the machine is linked, is possible to collect documents into the Blacklight index from an OAI-PMH repository.

To better manage, represent, and usable data retrieval the incorporation of the Knowledge Organization System (KOS) into an online IR system is necessary [79]. The addition of the KOS framework would be important to improve the integration of Dewey Decimal Classification (DDC) into the Blacklight Interface config.ini (in the user interface to invoke alpha browsing), *facet.ini* (in the user interface to invoke DDC dependent browsing), and *MARC_local.properties* (in support of Dewey class number-based indexing). This configuration was able to create modules, thus, collects data from a broad range of relative external sources beyond the boundaries of the library concerned, including access and licensed access, which is not part of the Blacklight area. Users can enable resource search by setting the search API value in the *config.ini* section of the search tabs, and in the *searchbox.ini* file. A separate section must be generated in the *searches.ini* format after choosing the source. It also supports full-text indexing, and Apache Tika is used as a method

for full-text extraction. A setup was a Java-based approach that allows metadata and text from various file types to be extracted by default at a time full-text indexing of Blacklight is disabled. Then the trail was set in the *fulltext.ini* file by configuring Apache Tika and commenting on the line requested in the *MARC_local.properties* file. This evolves also by replacing old technology such as federated browsing, allowing the emerging worldwide adoption of discovery tools for the web-scale discovery services. In relation to its full-text level with advanced recoverability functions, the solution was able to index and search through local content and hosted systems on the level of metadata.

Hence before selecting blacklight application to be used in this study, the other two open-source platforms, namely VuFind and eXtensible Catalog, were shortlisted for comparison with blacklight having their differences in technical and architectural functionality variations. Comparisons of detailed criteria/parameters were given, not to demonstrate which instrument performs better than the other, instead to describe the tool's possible characteristics. As no broad variations are found in the operations or access to metadata or index information, as shown in appendix 3, since these tools are contrary to predefine data, therefore, Blacklight was essentially chosen for this model's front layering discovery applications. In different configurations, all features were tested and evaluated with separate search syntax to experience the applicability and utility of the specific software tool. In fact, the selection of any tools for exploration depends on their features/parameters because all tools do not support such parameter values at an advanced level. Blacklight advantages over other tools were the use of Ruby on Rails platform as a programming/scripting language in a place of PHP of that of VuFind, as Ruby is newer than PHP. The tool is providing strong search functionality to the convention-driven Rails environment and has made a great development platform for organizations requiring extensive customization. Blacklight has allowed many pleasant objectives to coexist and demonstrate different custom behaviors. Blacklight has been used for MARC-built Solr indexes, Encoded Archival Definition (EAD), Text Encoding Initiative (TEI), Public Broadcasting Core (PBCore), Dublin Core, digital content management (content DM) record exports, and a variety of locally defined metadata standards. This versatility makes Blacklight desirable not only as a front-end library catalog but also as an institutional repository search solution.

3.5 Summary of the Chapter

The research philosophies of positivism and interpretivism which employed quantitative and qualitative research techniques respectively have been selected for the current work. Then, this research used a case study methodology to collect data using both qualitative and quantitative methods. While the qualitative technique relied on in-depth discussions, informal conversations, and unstructured interviews, the quantitative approach used questionnaires, and the Kish formula for cross-sectional has been used to determine the minimum number of samples and finally Agile Scrum methodology was used to design and build the digital library system.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

This chapter presents the findings from the data analysis that carried out from the current digital library systems available at NM-AST, UDSM and MU, awareness of EIRs, ways for library users to access EIRs and satisfaction with EIRs. The session provides detailed evaluation of machine learning models performance and PEM model validation. The assembly also provides performance testing and assessment of the created digital library system.

4.2 Results of the Data Analysis

To learn about user behavior, it was required to be familiar with the digital library systems accessible at these three universities, which led to a decision on the types of strategies that can be used to increase access and utilization of library resources.

Table 10: The Current Digital Library Systems Available at NM-AST, UDSM and MU

Author	*Institution Name*	*Digital Library System Currently Used*
[80]	The Nelson Mandela African Institution of Science and Technology (NM-IST)	i. Library Catalog ii. E-Books iii. Institutional Repository iv. Databases i.e Journals v. My Library on Finger Tips (MyLOFT)
[81]	University of Dar es Salaam (UDSM)	i. Subscribed online databases such as EBSCOhost, Springer, Taylor and Francis bestselling ebook titles, Wiley Online Library, Research4Life ii. Free online databases are e-theses and dissertation, science/medicine, mathematics/computer science business/economics iii. Library repository and e-journals

[82]	Mzumbe University (MU).	i. ii. iii.	Electronic resources available are in the versions of e-books, and e-journals, online subscribed databases such as EBSCO Host, Wiley, and Emerald Free databases are CORE, DOAJ (Directory of Open Access Journals), Open Textbooks Library Schorlary repository and Catalogue (OPAC)

4.2.1 Awareness of EIRs

Regarding the awareness of EIRs by the library users, four questions were formulated to analyze the aim. Among the questions that were asked was; are you aware of electronic library resources available at your Institution? A total of 1217 (73%) of the respondents were aware while 450 (27%) were not aware (Fig. 11).

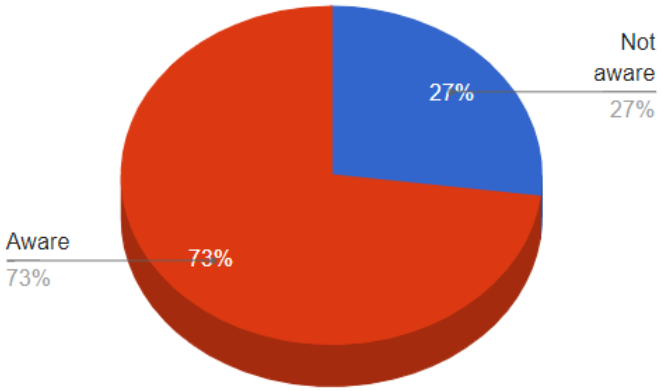

Figure 11 : UDSM, NM-AIST, and MU respondents' overall awareness

4.2.2 Ways for Library Users to Access EIRs

The second category of the survey was on how members of the academic community of the university get to access EIRs. A majority 503 (30.2%) of the respondents were found to be using two ways; physically by getting help from a librarian as well as remotely, 417 (25%) were using remote access way, 417 (25%) physically through the university library's print collections, 20 (15.6%) none of the choices provided and 70 (4.2%) physically through university library getting help from a librarian see Fig. 12.

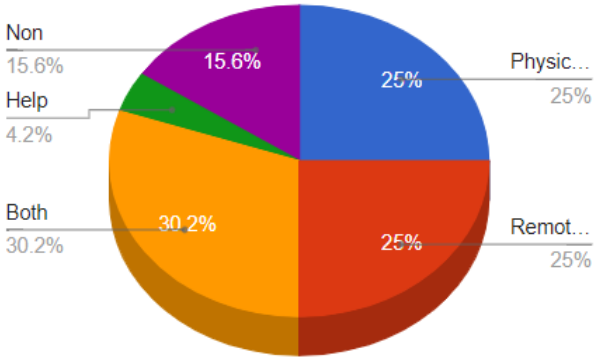

Figure 12: EIR Accessibility Percentages for all Respondents

4.2.3 Satisfaction with EIRs

Measuring users' satisfaction levels regarding electronic resources and services being offered by a library at UDSM, NM-AIST, and MU was the third aim of this survey. The question was formulated to understand the difficulties the library users face when using the online searching platform to access electronic library materials and references. A total of 834 (50%) of the respondents revealed that lack or poor internet connectivity was a major barrier to accessing electronic materials followed by 782 (46.9%) who attribute this to license restrictions to access digital collection, 628 (37.7%) to a limited number of downloads, 547 (32.8%) to the large mass of irrelevant information, 522 (31.3%) to the high cost of affordable online access, 495 (29.7%) to download delays, 338 (20.3%) to lack of in-depth ICT skills and knowledge, 300 (18%) to user authentication, 235 (14.1%) to information overload, 208 (12.5%) to poor ICT infrastructure, 130 (7.8%) to ack of standard interfaces on e-resources, 78 (4.7%) to the preference of print resources over e-resources, 27 (1.6%) to academic staffs discouragement, 27 (1.6%) to a lot of ads alerts and a final 27 (1.6%) did not use e-library resources (Fig. 13). Therefore each difficult was considered to be responded by 1667 number of respondents.

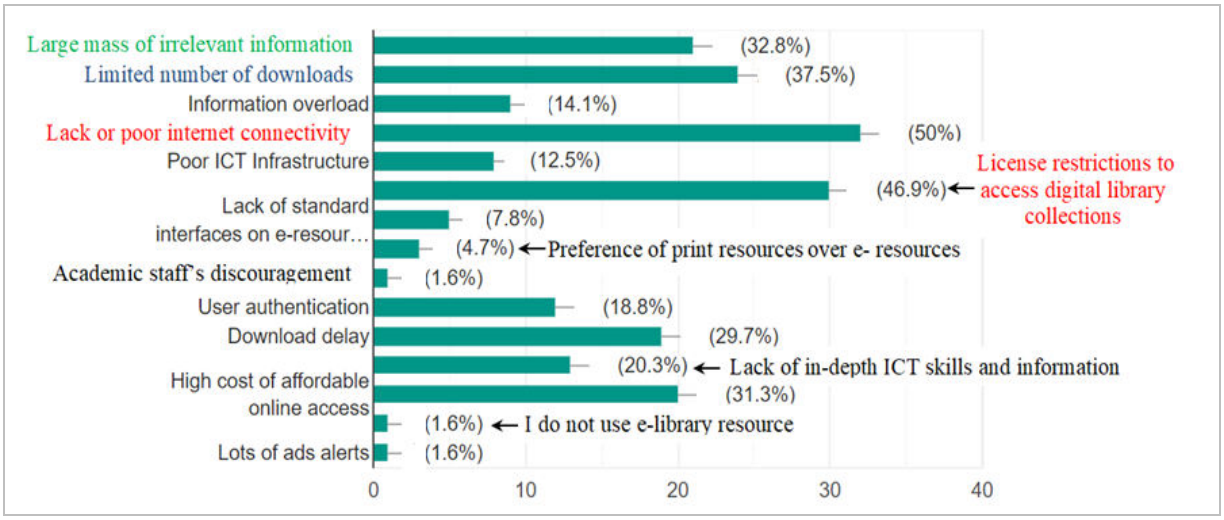

Figure 13 : Issues that Respondents Faced in Accessing EIRs, broken down by Percentage

Based on these results, the first phase of this research project was carried out to ascertain users' satisfaction with EIRs and challenges with digital library services in order to develop possible solutions. The general conclusion drawn from the survey questionnaire answers was that the two biggest problems faced by users of digital libraries were lack or poor internet connectivity and license restrictions to access digital library collections.

4.3 Evaluation for PEM, kNN, SVM and LR Models Performance

The performance of four models was evaluated in three stages: dataset training models accuracy with cleaning session, testing set models accuracy with cleaning session, and PEM Model training and testing without cleaning session for validation.

Table 11: The Output of WOS 46985 Dataset Training Model Accuracy with Cleaning Part

Category	Training Model Accuracy			
	PEM	kNN	SVM	LR
Computer Science	0.99	0.63	0.67	0.79
ECE	1.00	0.76	0.87	0.89
Civil Engineering	0.99	0.88	0.82	0.76
Medical Science	0.98	0.82	0.77	0.98
Biochemistry	0.99	0.98	0.92	0.82
Psychology	1.00	0.78	0.85	0.91
MAE	0.99	0.89	0.88	0.83
Average	**0.99**	**0.82**	**0.83**	**0.85**

We used the recorded history during training to get a plot of accuracy metrics. We picked up the training data acuracy (“acc”) for 7 trained labels during plotting see the Fig. 14.

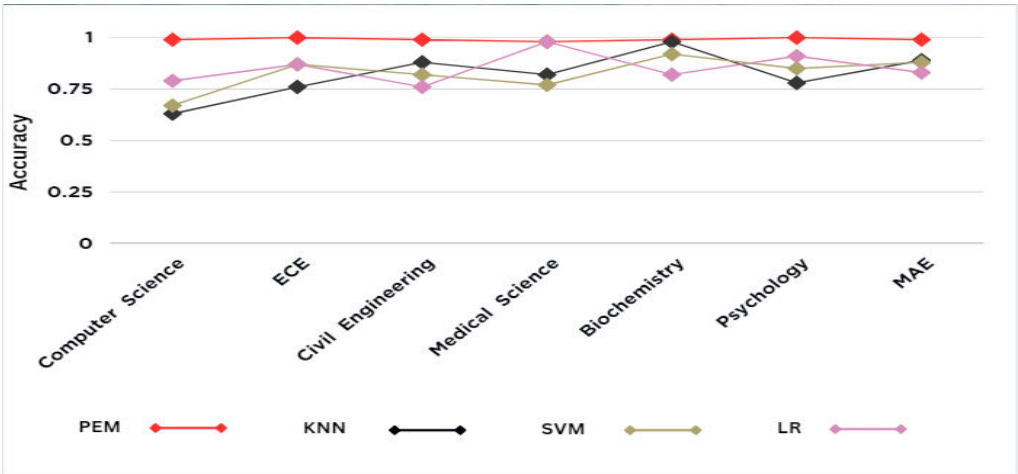

Figure 14: Graph of WOS 46985 Dataset Training Model Accuracy with Cleaning Part

The Fig. 14; shows that, the accuracy of proposed an ensembe model (PEM) increases rapidly in the first lebal which indicating that the network was learning fast. Afterwards, the curve flattens indicating that not too many epochs are required to train the model further, which means the model was starting to memorize the data. In order to calculate the Precision, Recall, and F-measure, we analyzed the four models—PEM, kNN, SVM, and LR on WOS 46985. The proposed an ensembe model outperforms other models with an average precision of 94% during the attempts determining the sequence prediction for the provided dataset. An F-score was the harmonic mean of a system's precision and recall values. It was calculated by the following formula: 2 x [(Precision x Recall) / (Precision + Recall)] (Table 12).

Table 12: Precision, Recall, and F-score are used to compare the performance of the PEM model with those of other models (kNN, SVM, and LR) on the WOS 46985 Dataset during the Training Set with Cleaning Part

Category	PEM			kNN			SVM			LR		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Computer Science	0.921	0.856	0.88	0.981	0.761	0.85	0.891	0.76	0.820	0.916	0.676	0.77
ECE	0.993	0.982	0.98	0.944	0.874	0.90	0.912	0.789	0.846	0.922	0.93	0.92
Civil Engineering	0.878	0.867	0.87	0.457	0.556	0.50	0.456	0.412	0.432	0.768	0.895	0.82
Medical Science	0.894	0.772	0.82	0.254	0.293	0.27	0.578	0.345	0.432	0.794	0.678	0.73
Biochemistry	0.926	0.907	0.91	0.689	0.587	0.63	0.634	0.533	0.579	0.987	0.87	0.92
Psychology	0.981	0.811	0.88	0.876	0.837	0.85	0.764	0.765	0.764	0.897	0.876	0.88
MAE	0.973	0.783	0.86	0.697	0.654	0.67	0.587	0.576	0.581	0.562	0.554	0.55
Average	**0.938**	**0.854**	**0.89**	**0.7**	**0.652**	**0.67**	**0.689**	**0.597**	**0.636**	**0.835**	**0.783**	**0.80**

P, R and F1 stand for Precision, Recall and F1-score respectively

Because all datasets contain predetermined train and test sets, the approach for the results in Table 12 used a random train-test split of 80% / 20%. The average precision, recall, and F1 scores are presented after the procedure is carried out five times.

Table 13: The Output of WOS 46985 Dataset Test Set Model Accuracy with Cleaning Part

Category	Testing Model Accuracy			
	PEM	kNN	SVM	LR
Computer Science	0.94	0.48	0.55	0.37
ECE	1.00	0.54	0.68	0.59
Civil Engineering	0.99	0.52	0.71	0.44
Medical Science	0.99	0.66	0.59	0.65
Biochemistry	0.99	0.69	0.73	0.43
Psychology	0.97	0.68	0.72	0.85
MAE	0.99	0.77	0.73	0.71
Average	**0.98**	**0.62**	**0.67**	**0.58**

When comparing the accuracy of the PEM model with alternative methods for text classification, WOS 46985 Dataset test set process; Table 13 shows that the PEM outperforms kNN, SVM, and LR by a maximum of 0.98. Additional in order to measure the effectiveness of the proposed system during test set a comparison of precision among selected algorithms, PEM, Logistic Regression, SVM and KNN also performed. The

experimental result with 37588 training documents and 9397 testing documents shows that PEM model provides the highest average precision of 95% compared to other algorithms (Tabe 14).

Table 14: Shows the precision, recall, and F1 results generated by applying the Four categorizers (PEM, LR, SVM and kNN) on WOS 46985 dataset during Test Set with Cleaning Part

Category	PEM			kNN			SVM			LR		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
Computer	0.987	0.973	0.98	0.681	0.692	0.686	0.893	0.697	0.783	0.768	0.686	0.725
ECE	0.991	0.833	0.905	0.788	0.806	0.797	0.784	0.623	0.694	0.834	0.739	0.783
Civil Engineering	0.875	0.938	0.905	0.345	0.467	0.397	0.487	0.373	0.422	0.892	0.795	0.841
Medical Science	0.976	0.857	0.913	0.321	0.352	0.336	0.467	0.267	0.34	0.789	0.676	0.728
Biochemistry	0.997	0.982	0.989	0.467	0.589	0.521	0.546	0.476	0.509	0.872	0.832	0.852
Psychology	0.917	0.879	0.898	0.783	0.578	0.665	0.689	0.718	0.703	0.937	0.567	0.706
MAE	0.929	0.997	0.962	0.578	0.474	0.521	0.674	0.556	0.609	0.753	0.783	0.768
Average	**0.953**	**0.923**	**0.936**	**0.566**	**0.565**	**0.560**	**0.649**	**0.53**	**0.58**	**0.835**	**0.725**	**0.772**

4.4 PEM Model Validation

There are several methods for validating a model, the two most well-known of which are cross validation and bootstrapping, although no one validation method works in all cases. Thus, it is important to understand the type of data that are dealing with. Therefore, the datasets for model validation were also pre-processed for this work. The primary goal of this second model validation procedure is to verify the strength of the developed model using the same WOS 46985 datasets without cleaning session as an effect of an imbalance accuracy. Hence, the datasets were just loaded into the codes in order to measure the outcomes in terms of performance, specifically to determine whether the model can be referred to and be adaptive enough to develop other models with new datasets. Table 15 shows the results of the training and testing set for PEM accuracy without the cleaning session.

Table 15: WOS 46985 Dataset Training and Test Set Model Accuracy without Cleaning Part

Category	Training set accuracy	Testing set accuracy
	PEM	**PEM**
Computer Science	0.83	0.59
ECE	0.95	0.72
Civil Engineering	0.87	0.69
Medical Science	0.90	0.78
Biochemistry	0.85	0.88
Psychology	0.82	0.91
MAE	0.96	0.76
Average	**0.88**	**0.76**

As a result, it was confirmed that the proposed ensemble model (PEM) without data cleaning performed better than the other three models (Logistic Regression, SVM, and kNN with data cleaning part), despite being less effective with an average accuracy of 0.88 compared to PEM with data cleaning part with an average accuracy of 0.99.

4.5 Developed an Accessible and User-Friendly Digital Library System

Increasingly large document collections require improved information processing methods for searching, retrieving, and organizing text. There are models that fall into different categories like supervised, unsupervised & reinforcement machine learning, in this work, supervised machine learning was selected which involves an output label associated with each instance in the data-set refers as text classification. Text classification is one of the important and typical tasks in supervised ML. This used technique allows machines to understand and then categorize text into known organized groups. Fig. 15; which shows the digital library resource discovery tool that has been developed.

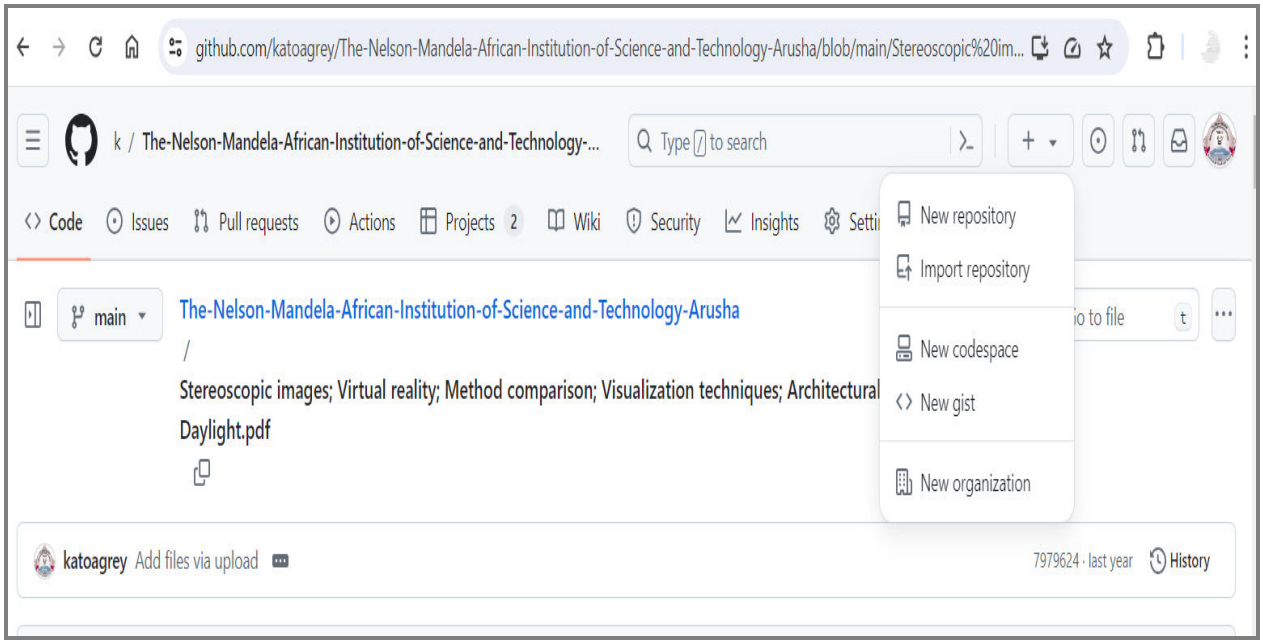

Figure 15: Developed an Accessible and User-Friendly Resource Discovery Tool (RDT)

RDTs are modern library catalogs that provide users with a Google-like search experience, provide libraries with collection information, and provide content suppliers with an alternative platform to stimulate resource use. The RDT has become the library's preferred search tool due to its capacity to search in-house and remote databases with a single search box in a way that is suitable for even experienced users. The created RDT contains an index that depicts the full text content that can be searched, as well as a linking facility to open the full text content.

4.5.1 Validation of the performance of developed digital library system

The digital library should adhere to the four principles of Web Content Accessibility Guidelines (WCAG) in order to be considered compliant. The four principles are outlined in the acronym “POUR” which stands for Perceivable, Operable, Understandable, and Robust. In order to find out website’s accessibility score the four digital library have been tested for compliance with WCAG 2.1 using the Webaccessibility.com service, the results are as shown in Fig. 16,17,18 and 19. WCAG is an internationally recognized set of standards and requirements for making web-based content accessible across all users including those with disabilities. Conformance with WCAG not only reduces legal risk (since many global accessibility laws are based on WCAG standards), but will also help an organization reach larger audiences since it reduces common technology barriers that users may face. The four library websites (Developed RDT, NM-AIST, MU and UDSM) have been successfully tested

for the compliance with the WCAG 2.1 standards using the Webaccessibility.com service. It has been established that only 1 websites out of 4 meet the accessibility standards of WCAG 2.1 AA with the compliance score of 90%. It has been observed that the libraries' current practices of providing web content accessibility are insufficient. The emphasis should be placed on the need to optimize and promote specialized library resources and to implement a continuous audit of web content accessibility by means of automatic testing and engaging users with poor ICT skills into these processes, see the snipped web accessibility health score results.

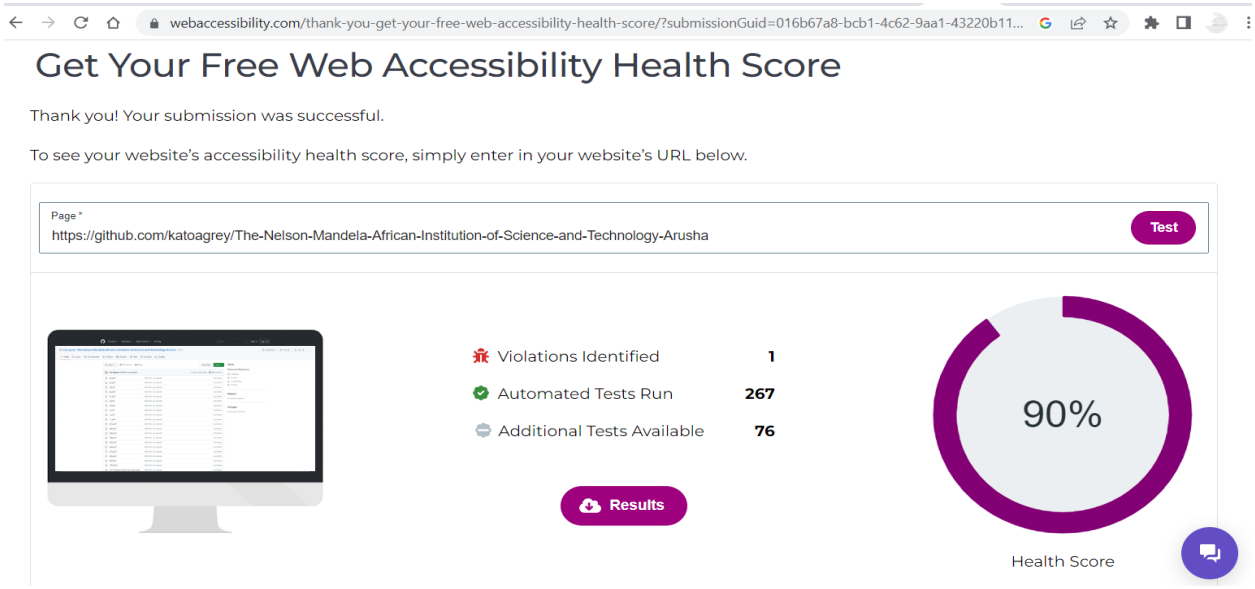

Figure 16: PEM Digital Library Web Accessibility Health Score

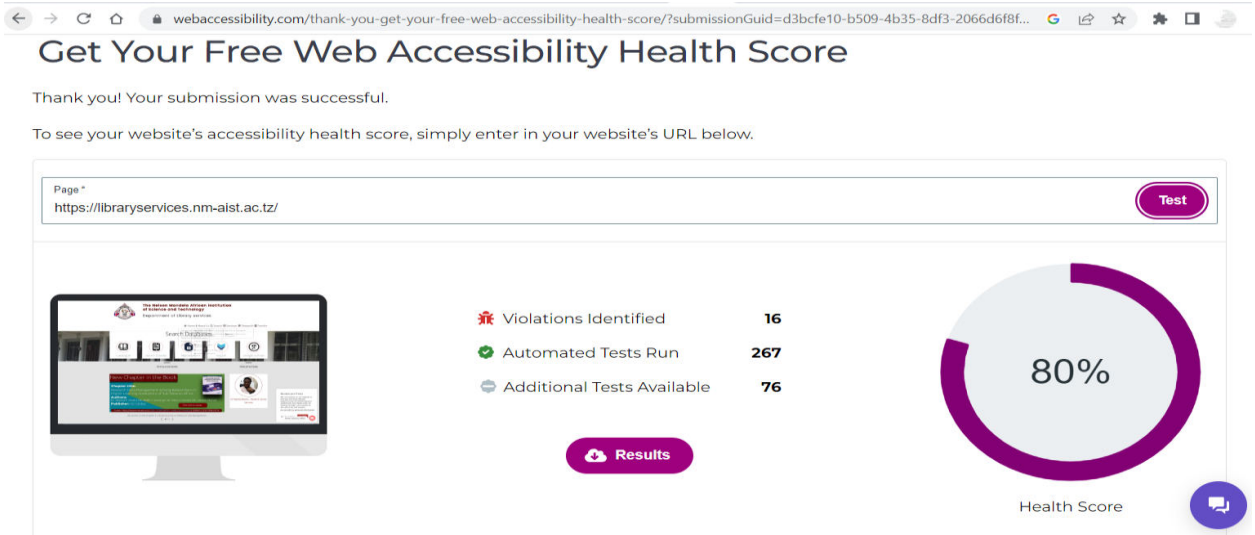

Figure 17: NM-AIST Web Accessibility Health Score

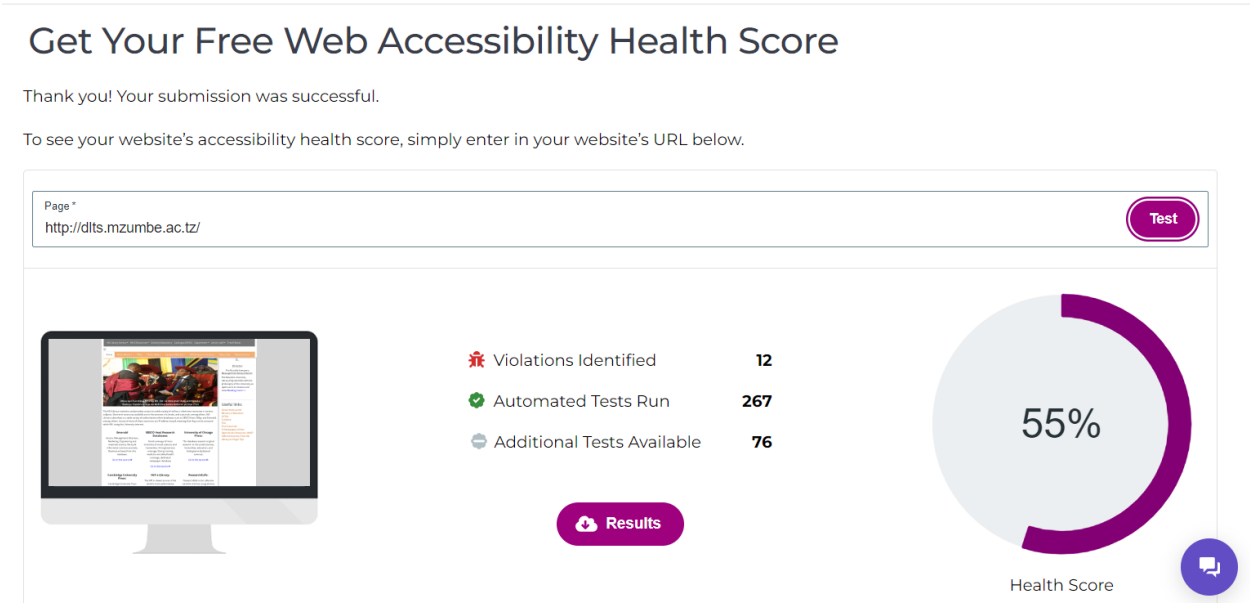

Figure 18: MU Web Accessibility Health Score

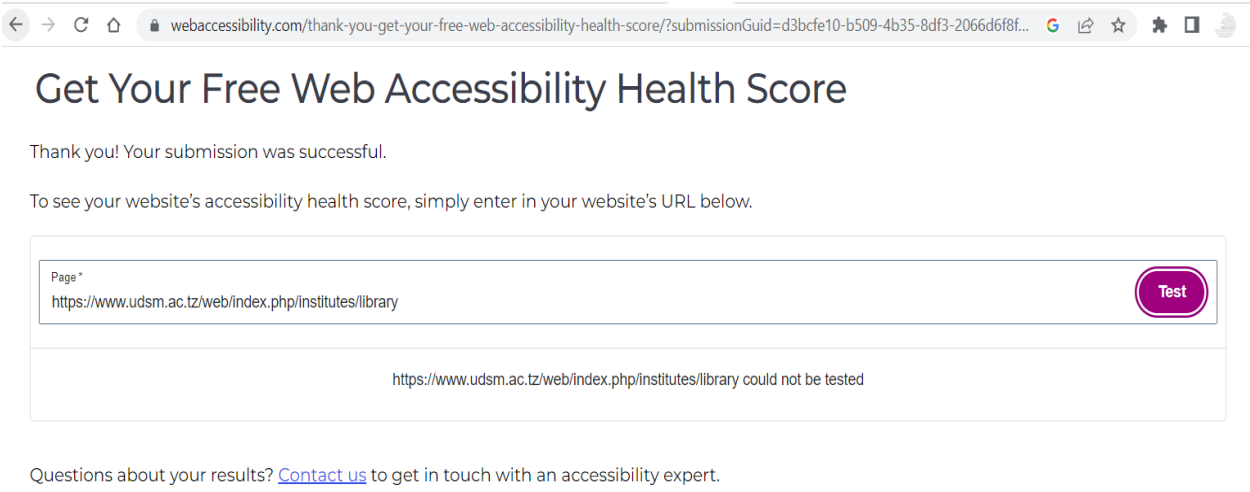

Figure 19: UDSM Web Accessibility Health Score

The resources of PEM developed digital library system have the highest established compliance score of web content accessibility of: 90% with 1 violation; NMAIST: 80% with 16 violations; MU: 55% with 12 violations and UDSM could not be tested due to exceeds number of violations. It should be emphasised that only 1 digital library out of 4 meets the accessibility standards with the compliance score of AA with one violation which requires to

be improved, 2 of these digital library (NM-AIST and MU) belong to the lower level A and the UDSM digital library doesn't meet lowest standard, it has accessibility barriers that need to be addressed, see Table 16 summarized compliance score of web content accessibility of four digital library tested.

Table 16: Compliance Score of Web Content Accessibility of Four Digital Library Tested with WCAG 2.1 standard Levels

Library Name	WCAG 2.1 compliance score, %	Number of violations	WCAG 2.1 Level
Developed RDT	90	1	AA
NM-AIST	80	16	A
MU	55	12	A
UDSM	-	-	-

WCAG levels are ranked on a scale from A to AAA. WCAG 2.1 A conformance is the minimum level of conformance, while WCAG 2.1 AA standards are midrange and widely considered legally acceptable. WCAG 2.1 AAA is the highest and most optimal accessibility level.

4.5.2 System Testing

(i) Functional Testing

Functional testing is a kind of software testing in which the system software is compared to functional specifications. It ensures that each application feature satisfies the functional requirements. The questionnaire (Appendix 2) helped to conduct functional testing, the questionnaire was distributed to library users to answer questions about system functionality, as indicated in Table 17.

Table 17: Functional Testing Results

Functional requirements	Results
Ease of use of the system	SUCCESSFUL
Availability of the information to multiple simultaneous users	SUCCESSFUL
Standard equipment/search tabs to provide access to the information	SUCCESSFUL
Auto technical guidance and instruction	SUCCESSFUL
Availability of the space needed to display the information	SUCCESSFUL
Ease of downloading in full text/ should be able to get the entire article immediately within the system	SUCCESFULLY
Meet high standards in literary, artistic, and aesthetic quality; technical aspects and format	SUCCESFULLY
Demonstrate quality appearance, and durability suitable to their intended use	SUCCESFULLY
Resources in a variety of formats accessible online and can be downloaded	SUCCESFULLY
The system should allow the enhancement and updating	SUCCESFULLY
The system should be accuracy, quality, and depth in searching online material	SUCCESFULLY
The system should be open source/open data sharing	SUCCESFULLY
The system should operate under balanced cost with need	SUCCESFULLY

4.5.3 System Validation

(i) System Satisfaction

The library users were generally satisfied with the system. A total of 22 library users from Nm-AIST and 13 library users from UDSM were interviewed to validate the system. To check their satisfaction with the system, out of 35 online library users, 94% (n=33) indicated being satisfied with the system as a tool for online library users as shown in Table 18. The p-value was 2e-8 which is less than 0.005, showing that the result was statistically significant.

Table 18: System Satisfaction

	Frequency	Percentage	Chi-squared test
How would you rate your overall satisfaction with this system as a tool for online library users?			$X^2 = 31.229$, df = 1, p = 2e-8
Satisfied	33	94	
Somehow satisfied	2	6	

Descriptive statistics are performed (frequency (n) and percentage (%)); Chi-square goodness of fit was used to compare the response answers of each question at a 5% confidence level.

(ii) Reasons for Satisfaction

Surveyed online library users were given the opportunity to describe why they were satisfied with the system. Out of 35 online library users, 46% (n =16) gave a reason that ease of use of the system was one of the factors that contribute to their satisfaction as shown in Table 19. While 31% considered ease of downloading in full text/ fewer of irrelevant information, 20% based on the system accuracy, quality, and depth in searching online material and 3% said was not sure. The P value became less than 0.001, which shows that the hypothesis was statistically significant.

Table 19: Reason for Satisfaction

	Frequency	Percentage	Chi-squared test
Give a reason for your answer on satisfaction/ dissatisfaction with this online library resource			$X^2 = 82.2$, df = 3, p < 0.001
Ease of use of the system	16	46	
The system accuracy, quality, and depth in searching online	7	20	

material		
Ease of downloading in full text/ fewer of irrelevant information	11	31
Not sure	1	3

Descriptive statistics are performed (frequency (n) and percentage (%)); Chi-square goodness of fit was used to compare the response answers of each question at a 5% confidence level.

4.5.4 System Recommendation

All (100% (n=35) of the respondents recommended this online library resource to be used in academic libraries for teaching, learning and research activities as shown in Table 20.

Table20: System Recommendation

	Frequency	**Percentage**	**Chi-squared test**
Will you recommend this platform to online library users to be used in academic libraries for teaching, learning and research activities?			
Yes	35	100	

Descriptive statistics are performed (frequency (n) and percentage (%)); Chi-square goodness of fit was used to compare the response answers of each question at a 5% confidence level.

As indicated in Table 21, of the 35 respondents who were asked why they would recommend a product, 71% (n=25) cited the system's ease of use as one of the reason for recommendation. The p-value became less than 0.001, which means its statistical significance to have a system that can function with no constraints.

Table21: Reason for the Recommendation

	Frequency	Percentage	Chi-squared test
Give the reason for your answer on or not recommending others to use this platform			$X^2 = 55.2$, df = 2, $p < 0.001$
System's ease of use	25	71	
Less irrelevant information during material search	9	26	
Not sure	1	3	

Descriptive statistics are performed (frequency (n) and percentage (%)); Chi-square goodness of fit was used to compare the response answers of each question at a 5% confidence level.

4.5.5 System Usability

The results showed that 94% (n=33) of online library users who were interviewed to check if the users required any external support to interact with the system said no, as shown in Table 22. The reason for this is that the digital skills required are relatively low, as shown in Table 23.

Table22: System Usability

	Frequency	Percentage	Chi-squared test
Will online library users require personal assistance to use this system?			$X^2 = 31.23$, df = 1, p = 2e-8
Yes	2	6	
No	33	94	

Descriptive statistics are performed (frequency (n) and percentage (%)); Chi-square goodness of fit was used to compare the response answers of each question at a 5% confidence level.

Table23: Reason for not Requiring Personal Support

	Frequency	Percentage	Chi-squared test
Specify your answer on how library users will need/ not need support			$X^2 = 33.06$, df = 1, p = 1e-8
The need of advanced ICT skills	1	3	
No need for advanced ICT skills	34	97	

Descriptive statistics are performed (frequency (n) and percentage (%)); Chi-square goodness of fit was used to compare the response answers of each question at a 5% confidence level.

4.6 Summary of the Chapter

Most of the questionnaire and interview respondents in this study were students from different faculties at the following three universities NM-AIST, MU and UDSM. A total of 1217 (73%) out of 1667 respondents were aware of EIRs. Machine learning techniques have been utilized to improve the usability and accessibility of the digital library system by developing PEM EIR. The performance of the developed digital library system was then assessed after ML had been turned into blacklight tool. Then the digital library systems PEM, NM-AIST, MU and UDSM were thoroughly examined in this chapter and finally the developed PEM system was highly ranked following user testing and validation.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 Introduction

This chapter presents the work's conclusion and recommendations, as well as future work, ethical issues and a summary of the chapter's content. In the developed DL discovery system, some components have been highlighted that are concerned with providing environmental quality, delivery quality, outcome quality, and standard with DL procedures. This work contributes to the conceptual design and development of the discovery tool in the university libraries concerning the transition from normal DL resources to modern DL resources. It also offers in-depth user forecasts and instruction on the management of library tools for references and information processing. Nonetheless, most libraries are now starting to introduce one-stop shopping programs for their customers, where at the same time several different types of resources are being queried. Additionally, this work offers widely library user awareness, adoption, and usage of DL resources in which the use of DL discovery systems is distinct from the use of other library services such as the online catalog and the repositories.

5.2 Conclusion

The current study established that more than half of the respondents, 1217 (73%) were aware of EIRs and had knowledge of what they constitute but have restricted access due to numerous obstacles. The main challenges of online DL are information search and retrieval attributes related to label relevance and feature correlation segments. To improve availability and usefulness of online DL, this work adopted machine learning techniques to enhance the access and utilization of library resources. The four library websites (Developed RDT, NM-AIST, MU and UDSM) have been successfully tested for the compliance with the WCAG 2.1 standards. It has been established that only developed RDT website out of 4 meet the accessibility standards of WCAG 2.1 AA with the compliance score of 90%.Concerning the satisfaction of a provision of access to EIRs by developed RDT, 94% (33 numbers of respondents out of 35) strongly agreed that robust EIR infrastructure was the most demanding characteristic that was incorporated in developed RDT. The users preferred to search e-resources with simple and user-friendly interfaces. After system development and validation, the proposed system showed a huge significance of having a platform that is suitable for both

users with in-depth ICT skills and non ICT skilled for online information searching. Furthermore, the system offers full-text downloads as well as an upgrading and updating of its different features to accommodate a large number of users at once. This study concludes that there should be key steps to increase the confidence, awareness, and use of e-library resources by users. One of the actions that should be done is to set up proactive end-user training sessions in the departments and libraries. Library users have recommended the use of developed RDT due to its simple, easy and user-friendly tools for online information searching. The main achievement of this work, the online search tool for library users has been developed to allow one-point access to different repositories and databases. The application of RDT is responsible to the user needs, as stated in related articles, and as shown in this work. Study results show that developed RDTs will be further advanced to provide simplification and adaptability to meet the needs of various user groups. Within the longer-term steps, researchers could conduct research to match various designs, which could result in an additional suite of information resources retrieval for anticipated RDTs disabled users.

5.3 Recommendations

To improve the use and awareness of e-library resources the following suggestions have been put forth after the analysis of e-resources conducted among research scholars of NM-AIST, UDSM, MU and globally. All these stakeholders should adapt to the following highlited recommendations for enhancing their institutions library digital systems. There should be several department-wise training and orientation programs for students and faculty members to optimize the utilization of e-resources. Similarly, from time to time, e-library resources demonstrations and trials should be organized for the new resources from publishers for users to create awareness of recent developments within the e-library resources technology. Additionally, a high-speed internet network should be availed within campuses and hostels with security measures agreeable to the users to limit the time users spend searching for information. Regarding the provision and use of e-resources, library managers should introduce services such as content alert, bulletin board, library newsletters, pamphlets, training tutorials/modules on use of the subscribed e-resources/databases. Librarians have to be compelled to improve the infrastructure of libraries EIRs which may increase the literacy skills of the users. Library managers should also develop web-based library services for librarians to survey and obtain feedback regarding the awareness and usefulness of the available e-resources. Libraries may prefer a discovery service based on its functionality and

content coverage as well as resource management system from vendors based on sets of distinct requirements. Users require a discovery system excelled at providing precise methods for interacting with librarians in a variety of ways. The genre of discovery services will continue to be enhanced to add new functionality and capabilities in response to request from libraries and users to improve their competitive position. Some of technical improvements and enhancement should be in the areas of application programming interfaces, expanding API ecosystem, enhancing social features, rich media materials and collections, research data sets, discovery and access related to special collections material, analytics and altimetrics.

5.3.1 Future Work

Based on the information processing, the two major components considered in RTD systems development are ease of use and access to digital library resources. The use of Machine Learning techniques in stored resources in libraries with an increasingly large share and access of digital information resources was an approach addressed to achieve the main objective in this study. The developed machine learning proposed ensemble model (PEM) was accomplished by tuning different models for labels relevance (weights) and their attributes (features) correlation and it was two stage process (binary label classification). Error correction was performed by applying algorithm model built from historical dataset to a new dataset that improved the classification efficiency. At this instant the future work should be on fine-tuning different combination of machine learning algorithms to attain better label categorization and significantly improve the performance.

5.3.2 Ethical Considerations

All materials provided by officials to aid in various tasks for this study, including data gathered and information acquired in various ways, have been used only for the purpose of this research activity.

5.4 Chapter Summary

The results of the current study have shown that more than half of the respondents, 1217 (73%) knew about EIRs and understood what they comprised, but face accessibility issues due to social and technological barriers. The created RDT allows users to make their choices about the library's search utilities. In the longer-term phases, researchers might do study to match different designs, which could result in an additional suite of information resources

retrieval for anticipated RDTs disabled users. In order to increase the usage and knowledge of e-library resources the following recommendation has been made in response to an exploration of e-resources conducted among research scholars of NM-AIST, UDSM, and MU as well as globally. All of these stakeholders should adapt the developed RDT in this research to improve the digital library systems at their respective organizations. The next phase of this research should focus on fine-tuning various machine learning algorithm combinations in order to get better label classification and substantially enhance performance.

REFERENCES

[1] O. Bonnie and L. Tijani, "Availability , Awareness and use of Electronic Information Resources amongst Library Registered Students in Federal University Lafia and Nasarawa State University Keffi," vol. 7, no. October, pp. 97–107, 2020.

[2] M. Mani, S. S. Hameed, and A. Thirumagal, "Impact of ICT knowledge, Library Infrastructure facilities on students' usage of E-Resources - An empirical study," *Libr. Philos. Pract.*, vol. 2019, no. January 2021, 2019.

[3] M. N. Masrek and J. E. Gaskin, "Assessing users satisfaction with web digital library: the case of Universiti Teknologi MARA," *Int. J. Inf. Learn. Technol.*, vol. 33, no. 1, pp. 36–56, 2016.

[4] M. A. Ugc-net and I. P. Commission, "E-learning, e-searching and e-resource management in the libraries," *IP Indian J. Libr. Sci. Inf. Technol.*, vol. 3, no. 1, pp. 16–18, 2020.

[5] A. Temboge and U. A. Ga'anda, "Use and User Satisfaction With Library Information Resources and Services in Federal University of Kashere Library," *Researchgate.Net*, no. October, 2022.

[6] I. S. Isibika and G. F. Kavishe, "Utilisation of subscribed electronic resources by library users in Mzumbe university library, Tanzania," *Glob. Knowledge, Mem. Commun.*, 2018.

[7] A. S. Samzugi, "User preference on use of print and Electronic Resources in Selected Universities in Tanzania: A surve," *Libr. Philos. Pract.*, vol. 2019, 2019.

[8] E. F. Wema, "The role of Tanzania library services board in enhancing information literacy skills among the public in Tanzania," *Inf. Dev.*, 2022.

[9] N. Tom-george and N. Onyema, "Digital Library Services for Sustainable University Education," no. July, 2022.

[10] R. Felder, L. Silverman, B. Joyce, B. Showers, B. Ghirardini, and University of South Africa, "Planning learning and the Teaching and Learning Programme model," *Learning and Skills Improvement Service (LSIS)*, 2010. .

[11] D. Dickenson, "How Academic Libraries Help Faculty Teach and Students Learn: The 2005 Colorado Academic Library Impact Study," *A Closer Look*, 2006.

[12] N. A. Aderibigbe and B. A. Ajiboye, "User education program as determinant of electronic information resources usage in Nimbe Adedipe University Library, Nigeria," *Electron. Libr.*, 2013.

[13] A. Toledo, "Open access and OER in latin america: A survey of the policy landscape in chile, colombia and uruguay," in *Adoption and Impact of OER in the Global South*, 2017.

[14] M. Enis, "Accessibility Upgrade," *Libr. J.*, 2013.

[15] P. B. Babu and M. Krishnamurthy, “Library automation to resource discovery: A review of emerging challenges,” *Electron. Libr.*, 2013.

[16] S. Marikyan, D. & Papagiannidis, “Unified Theory of Acceptance and Use of Technology: A review. In S. Papagiannidis (Ed),” 2021.

[17] H. Abuhassna *et al.*, “Trends on Using the Technology Acceptance Model (TAM) for Online Learning: A Bibliometric and Content Analysis,” *Int. J. Inf. Educ. Technol.*, vol. 13, no. 1, pp. 131–142, 2023.

[18] S. N. Mohanty, “Emotion and Information Processing: A Practical approach,” *Emot. Inf. Process. A Pract. approach*, no. January 2021, pp. 1–221, 2020.

[19] P. T. C. de O. Salvador, M. D. S. Bezerril, C. M. S. Mariz, M. I. D. Fernandes, J. C. A. Martins, and V. E. P. Santos, “Virtual learning object and environment: a concept analysis,” *Revista brasileira de enfermagem*. 2017.

[20] M. Sitnicki, “DEVELOPMENT OF A MODEL OF DIGITAL RESEARCH UNIVERSITIES,” *Balt. J. Econ. Stud.*, 2018.

[21] M. Mutula Stephen and S. Mudogo Mutula, “The Electronic Library Status of digital heritage preservation management in Eastern Africa Status of digital heritage preservation management in Eastern Africa,” *Electron. Libr. Collect. Build. Iss Aslib Proc.*, 2014.

[22] A. Jose, “Evaluation of Digital Libraries: A Case Study,” in *International Conference of Semantic Web & Digital Libraries*, 2007.

[23] D. Dukić and J. Strišković, “Croatian university students’ use and perception of electronic resources,” *Libr. Inf. Sci. Res.*, vol. 37, no. 3, pp. 244–253, 2015.

[24] M. Koivunen and K. Saranto, “Nursing professionals’ experiences of the facilitators and barriers to the use of telehealth applications: a systematic review of qualitative studies,” *Scandinavian Journal of Caring Sciences*. 2018.

[25] E. Matheus and R. Abankwah, “Learners’ perceptions on online library resources at Namibian College of Open Learning,” *Inf. Learn. Sci.*, 2018.

[26] A. G. Larson and M. Owusu-Acheaw, “Information needs of distance learners: A case of Winneba study center, university of education, Winneba, Ghana,” *Turkish Online J. Distance Educ.*, 2016.

[27] D. Jotwani, “Library resources and services in Indian Institutes of Technology,” *Ann. Libr. Inf. Stud.*, 2013.

[28] J. A. Adams and S. C. Bonk, “Electronic Information Technologies and Resources: Use by University Faculty and Faculty Preferences for Related Library Services,” *Coll. Res. Libr.*, 2014.

[29] J. Y. Hwang, J. Kim, B. Lee, and J. H. Kim, “Usage patterns and perception toward e-books: Experiences from academic libraries in South Korea,” *Electron. Libr.*, 2014.

[30] A. Rubel, “Libraries, Electronic Resources, and Privacy: The Case for Positive Intellectual Freedom,” *Libr. Q.*, 2014.

[31] D. Agaba, I. Kigogo-Bukenya, and J. Nyumba, “Utilization of electronic information resources by academic staff at Makerere University,” *Univ. Dar es Salaam Libr. J.*, 2011.

[32] M. Eusuf, M. S. Ahmed, and M. S. R. Eusuf, “a Conceptual Framework for Digital Library and That Apply To the Engineering Education At Higher Learning Institution,” *Khulna Univ. Stud.*, no. November, pp. 332–338, 2022.

[33] D. Ojakaa *et al.*, “CSOs HSS support proposal,” *World Heal. Organ.*, 2014.

[34] E. Mączyńska, “The economy of excess versus doctrine of quality,” *Kwart. Nauk o Przedsiębiorstwie*, 2017.

[35] A. D. Alharthi, M. Spichkova, and M. Hamilton, “Sustainability requirements for eLearning systems: a systematic literature review and analysis,” *Requir. Eng.*, pp. 1–21, 2018.

[36] A. Balog, “Testing a multidimensional and hierarchical quality assessment model for digital libraries,” *Stud. Informatics Control*, 2011.

[37] J. Jeng, “What Is Usability in the Context of the Digital Library and How Can It Be Measured,” *Inf. Technol. Libr.*, 2013.

[38] M. N. Masrek, A. Jamaludin, and S. A. Mukhtar, “Evaluating academic library portal effectiveness: A Malaysian case study,” *Libr. Rev.*, 2010.

[39] M. H. Vinagre, L. G. Pinto, and P. Ochôa, “Revisiting digital libraries quality: A multiple-item scale approach,” *Perform. Meas. Metrics*, 2011.

[40] M. Kyrillidou and S. Giersch, “Developing the DigiQUAL protocol for digital library evaluation,” 2005.

[41] F. D. Davis, “Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology,” *MIS Q.*, 1989.

[42] R. Hussein, N. S. A. Karim, and M. Hasan Selamat, “The impact of technological factors on information systems success in the electronic-government context,” *Bus. Process Manag. J.*, 2007.

[43] W. Hong, J. Y. L. Thong, W. M. Wong, and K. Y. Tam, “Determinants of user acceptance of digital libraries: An empirical examination of individual differences and system characteristics,” *J. Manag. Inf. Syst.*, 2001.

[44] G. T. Lee, N. Dahlan, T. Ramayah, N. Karia, M. H. Abu, and H. Asaari, “Impact of Interface Characteristics on Digital Libraries Usage,” *Malaysian Online J. Instr. Technol.*, 2005.

[45] J. Y. L. Thong, W. Hong, and K. Y. Tam, “Understanding user acceptance of digital libraries: What are the roles of interface characteristics, organizational context, and

individual differences?," *Int. J. Hum. Comput. Stud.*, 2002.

[46] T. Ramayah and J. W. C. Lee, "System characteristics, satisfaction and e-learning usage: A structural equation model (SEM)," *Turkish Online J. Educ. Technol.*, 2012.

[47] J. C. Roca, C. M. Chiu, and F. J. Martínez, "Understanding e-learning continuance intention: An extension of the Technology Acceptance Model," *Int. J. Hum. Comput. Stud.*, 2006.

[48] E. T. Lwoga, "Measuring the success of library 2.0 technologies in the African context: The suitability of the DeLone and McLean's model," *Campus-Wide Inf. Syst.*, 2013.

[49] M. Kuźma and A. Mościcka, "Accessibility evaluation of topographic maps in the National Library of Poland," *Abstr. ICA*, 2019.

[50] H. Jeong, "An investigation of user perceptions and behavioral intentions towards the e-library," *Libr. Collect. Acquis. Tech. Serv.*, 2011.

[51] F. Xu and J. T. Du, "Factors influencing users' satisfaction and loyalty to digital libraries in Chinese universities," *Comput. Human Behav.*, vol. 83, pp. 64–72, 2018.

[52] R. Agarwal, V. Sambamurthy, and R. M. Stair, "COGNITIVE ABSORPTION AND THE ADOPTION OF NEW INFORMATION TECHNOLOGIES.," *Acad. Manag. Proc.*, 2011.

[53] N. Parsazadeh, N. M. M. Zainuddin, R. Ali, and A. Hematian, "a Review on the Success Factors of E-Learning," *Second Int. Conf. e-Technologies Networks Dev.*, pp. 42–49, 2013.

[54] G. A. Torres-Vargas, "The Design of a Digital Library for Mexican Universities," *Procedia - Soc. Behav. Sci.*, vol. 73, pp. 751–756, 2013.

[55] M. Tripathi and V. K. J. Jeevan, "Quality assurance in distance learning libraries," *Quality Assurance in Education*. 2009.

[56] P. Joore *et al.*, "Online library : digital copies," *BMJ Open*, 2016.

[57] D. L. Ferullo, "Major copyright issues in academic libraries: Legal implications of a digital environment," in *The Changing Landscape for Electronic Resources: Content, Access, Delivery, and Legal Issues*, 2013.

[58] G. Mahesh and R. Mittal, "Digital content creation and copyright issues," *Electron. Libr.*, 2009.

[59] G. Cleveland, "Digital libraries: definitions, issues and challenges," *Occas. Pap. IFLA UDT Core Progr.*, 2008.

[60] K. Wilson, C. Neylon, L. Montgomery, and C. K. K. Huang, "Access to academic libraries: An indicator of openness?," *Inf. Res.*, 2019.

[61] H. Piwowar *et al.*, "The state of OA: A large-scale analysis of the prevalence and

impact of Open Access articles," *PeerJ*, 2018.

[62] O. Bolarinwa and S. C. A. Utulu, "Open access: Perceptions and reactions of academic librarians in Nigerian private universities," *African J. Libr. Arch. Inf. Sci.*, 2011.

[63] E. T. Lwoga and F. Questier, "Open access behaviours and perceptions of health sciences faculty and roles of information professionals," *Health Info. Libr. J.*, 2015.

[64] eghana M. Sanjeeva and S. C. Powdwal, "Open Access Initiatives: Reframing the role of Librarians," *Libr. Her.*, 2017.

[65] B. K. Roy, S. C. Biswas, and P. Mukhopadhyay, "Designing web-scale discovery systems using the VuFind open source software," *Libr. Hi Tech News*, 2018.

[66] S. Popowich, "Agile Open-Source Discovery: Blacklight with EBSCO Discovery Service," in *Exploring Discovery: The Front Door to your Library's Licensed and Digitized Content*, 2016.

[67] E. Sadler, "Project Blacklight: A next generation library catalog at a first generation university," *Libr. Hi Tech*, 2009.

[68] D. N. Sudha and Y. M. Latha, "Semi Supervised Multi Text Classifications for Telugu Documents Labeled Unlabeled Text Learning unit Telugu News Articles Unlabeled data Count Vectorizer , TF-IDF Dimensionality Reduction Machine Learning Classifiers (Hyperparameter tuning)," vol. 12, no. 12, pp. 644–648, 2021.

[69] K. Shah, H. Patel, D. Sanghvi, and M. Shah, "A Comparative Analysis of Logistic Regression, Random Forest and KNN Models for the Text Classification," *Augment. Hum. Res.*, vol. 5, no. 1, 2020.

[70] Z. Liu, X. Lv, K. Liu, and S. Shi, "Study on SVM compared with the other text classification methods," *2nd Int. Work. Educ. Technol. Comput. Sci. ETCS 2010*, vol. 1, pp. 219–222, 2010.

[71] L. Bates and I. Gurevych, "Like a Good Nearest Neighbor: Practical Content Moderation with Sentence Transformers," 2022.

[72] B. Bhattarai, O. C. Granmo, and L. Jiao, "ConvTextTM: An Explainable Convolutional Tsetlin Machine Framework for Text Classification," *2022 Lang. Resour. Eval. Conf. Lr. 2022*, no. June, pp. 3761–3770, 2022.

[73] C. Sun and B. Luo, "Analysis of English Writing Text Features Based on Random Forest and Logistic Regression Classification Algorithm," *Mob. Inf. Syst.*, vol. 2022, 2022.

[74] Z. Wang, H. Xu, P. Zhou, and G. Xiao, "An Improved Multilabel k-Nearest Neighbor Algorithm Based on Value and Weight," *Computation*, vol. 11, no. 2, 2023.

[75] M. H. Nguyen, "A label-oriented approach for text classification," *Int. J. Innov. Comput. Inf. Control*, vol. 16, no. 5, pp. 1593–1609, 2020.

[76] K. Kowsari, D. E. Brown, M. Heidarysafa, K. Jafari Meimandi, M. S. Gerber, and L.

E. Barnes, “HDLTex: Hierarchical Deep Learning for Text Classification,” *Proc. - 16th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2017*, vol. 2017-Decem, pp. 364–371, 2017.

[77] H. Wiegand, “Kish, L.: Survey Sampling. John Wiley & Sons, Inc., New York, London 1965, IX + 643 S., 31 Abb., 56 Tab., Preis 83 s.,” *Biom. Z.*, 1968.

[78] M. Breeding, “Beyond the ILS: A New Generation of Library Services Platforms,” in *Robots in Academic Libraries: Advancements in Library Automation*, 2013.

[79] S. Clayton and J. M. Widener, “Beyond Embedded Librarianship: Co-Teaching with Faculty to Integrate Digital Scholarship in Undergraduate Research,” in *Undergraduate research and the academic librarian : case studies and best practices*, 2017.

[80] NM-AIST, “No Title,” 2023. [Online]. Available: https://libraryservices.nm-aist.ac.tz/.

[81] UDSM, “No Title,” 2023.

[82] MU, “No Title,” 2023.

APPENDICES

Appendix 1: Survey Questionnaire for Gathering Data from Study Case Institutions' Library Users

1. What is your institution?

Nelson Mandela African Institution of Science and Technology

University of Dar es Salaam

Mzumbe University

2. What is your gender?

Male

Female

Others

3. What is your education level?

PhD

Master's

Bachelor

Diploma

Certificate

Others

4. What is your title?

Teaching staff

Non-teaching staff

Student

Institutional library staff

5. What is your discipline area?

Science and technology

Engineering and Technology

Science and Engineering

Natural and Applied Sciences

Humanities

Social Sciences

Agricultural Sciences and Fisheries Technology

Health and Allied Sciences

Business and Development Studies

Education

Law

Journalism and Mass Communication

Kiswahili Studies

Resource Assessment

Marine Sciences

Others

6. How would you rate your basic ICT knowledge and skills in searching the electronic resources?

Very Good

Good

Average

Poor

7. Have you received announcements about electronic library resources?

Yes

No

8. Do you have a user electronic library account on your university library portal for you to access and use electronic library resources?

Yes

No

9. If not from question number 8 above why have you not set up a library account?

Lack of time

Poor internet access

No need for it

I have other access to adequate information resources

Lack of awareness

I have a library account

10. How do you get to access electronic library resources?

Physically through university library getting help from a librarian

Remotely

Both

None

11. Does Internet connectivity prevent you from accessing electronic library resources?

Yes

No

Sometime

12. How important is it for you to have access to electronic library resources?

Not important

Slightly important

Moderately important

Important

Very important

13. Are you aware of electronic library resources available at your Institution?

Yes

No

14. Have you been exposed to any kind of online electronic library resource before?

Yes

No

15. Which electronic resources have you been using for searching for relevant academic materials? (Tick all that applies)

E-Newspapers

E-thesis and Dissertations (Institutional Repository)

E-Reference sources

Scholarly database (i.e. TEEAL, Research4life)

Publishers Journal Archive

Publishers Journal Collections

E-magazines

E-books Collections (Library System/ILS, OPAC)

Bibliographic Databases

Full-Text Abstracting and Indexing

Google Scholar

Others

16. What mode/type of e-learning were you or are you using to access electronic resources? (Tick all that applies)

Computer-based learning

Mobile-based learning

Video-based/CD-ROM based learning

Webinars (a.k.a. Web seminars/video conferencing)

Virtual classroom (e.g. online portal)

Others

17. Which one of the following information sources have you used over the past 3 months? (Tick all that applies)

    World Wide Web/www

    Electronic library resources

    Institution library print materials

    Institutional repository

    Colleagues in other Universities

    Personal library collection

    None

    Others

18. Your studies/work/research work had been affected by the difficulty of accessing online library references

    Yes

    No

19. What problems are you facing when using the online searching platform for accessing electronic library materials and references? (Tick all that applies)

    A large mass of irrelevant information

    Limited to the number of downloads

    Information overload

    Lack of or poor internet connectivity

    Poor ICT Infrastructure

    License restrictions to access the digital library collection

    Lack of standard interfaces on e-resources portals

    Preference of print resources over electronic resources

    Academic staff discouraging the use of e-resources

    User authentication

    Download delay

    Lack of in-depth ICT skills and information searching skills

    The high cost of affordable online access

    Others

Appendix 2: User Acceptance Testing for Resource Discovery Tool (RDT)

This survey questionnaire tool collects information from online library users to verify their expectations of the developed resource discovery tool.

* Required

1. Do you think the system conforms to the online library guidelines? (On information sharing for teaching, learning and research) *

Check all that apply.

- ☐ Yes
- ☐ No
- ☐ I don't know

2. Does the system contain the required features/units that make easier for user to retrieve online contents? *

Mark only one oval.

- ⬭ Yes
- ⬭ No
- ⬭ I don't know

3. Based on your experience with online library in your institution, do you think this online tool will enhance the access of online materials at required level? *

Mark only one oval.

- ⬭ Yes
- ⬭ No
- ⬭ I don't know

4. Does the system compatible with limited recourse setting areas (such as poor internet, electricity, and digital literacy) *

Mark only one oval.

⬭ Yes

⬭ No

⬭ I don't know

5. How would you rate your overall satisfaction with this system as a tool for online information? *

Mark only one oval.

⬭ Satisfied

⬭ Somehow satisfied neutral

⬭ Somehow dissatisfied dissatisfied

6. Give a reason for your answer on satisfaction/ dissatisfaction with this online information tool *

_____

7. Will you recommend this online information tool to others for teaching, learning and research activities? *

Mark only one oval.

⬭ Yes

⬭ No

8. Give the reason for your answer on recommending/ not recommending others to use this online information tool *

_____

_____

_____

9. What features did you like the most in this online information tool and reasons for your answer? *

_____

10. What features did you like the least in this online information tool and what reasons for your answer? *

_____

11. What topics/ features should be added in this online information tool? Reasons for your choice/answer *

_____

12. What topics/ features should be removed from this online information tool? Reasons for your choice/answer *

_____

13. Does the system enhance the availability of online information based on demand? *

Mark only one oval.

⬭ Yes

⬭ No

⬭ I don't know

14. Specify your answer how it does *

_____

15. Do online library users need support to use this tool? *

Mark only one oval.

⬭ Yes

⬭ No

⬭ I don't know

16. Specify your answer on how online library users will need/ not need support *

_____

Appendix 3: Comparison of exploration tools for Web-scale services

S/n	Parameters/Criteria:	Tools for Resource Discovery		
		Blacklight	eXtensible Catalog	VuFind
1.	*Indexing and Organization*			
1.1	FRBRised compilation and metadata/FRBRized view	√	√	√
1.2	RDA cataloging assistance	√	*	*
1.3	Multiple author indexing	√	√	√
1.4	Refinement of outcomes of search	√	√	√
1.5	Non-MARC data support and authority support	√	√	√
1.6	Flexible support of formats other than MARC with metadata	*	*	√
1.7	As an additional function, Federated Searching	√	√	√
1.8	Navigation Faceted	√	√	√
1.9	Filtering/Search Filtering for Record Retrieval	√	√	√
1.10	Structured navigation (support for DDC/LC classification schemes listed)	√	√	√
1.11	Authority regulation system	√	*	√
1.12	The ILS (single ILS) objects' real-time status	√	√	√
1.13	Many ILS objects' real-time status (in the case of a union catalog, many ILS)	√	*	√
1.14	A scan according to the description of facets	√	√	√
1.15	Indexing of social tags generated by users	√	√	√
1.16	User Interface for the Recommender Device	√	√	√

1.17	Incorporation into effects of geo-location	√	√	√
1.18	Help for the recovery of results from more distributors of third parties	√	√	√
1.19	Integrated assistance for producing data on demand	√	√	√
1.20	Availability of the connection feature with extractor of full text	√	*	√
1.21	Availability for domain-specific metadata schema management with multiple support crosswork mapping	√	*	√
2.	*Progressive quest*			
2.1	Fuzzy lookup	√	√	√
2.2	Searching range	√	√	√
2.3	Narrow filtering for search/search	√	√	√
2.4	Searching for Full-text	√	*	√
2.5	Scoped scanning (scanning increasing/decreasing)	√	√	√
2.6	Searching smartly	√	√	√
2.7	Browsing Faceted	√	√	√
2.8	Sounded Question	√	*	√
2.9	Many structures for ranking	√	√	√
2.10	Check by word	√	√	√
2.11	Searching/Fielded Search Field Level	√	√	√
2.12	Call Number Indexing	√	√	√
2.13	Internationalization/Multilinguality	√	√	√
2.14	Check for the 'Bento Box' (categorization or records by source)	*	*	√
2.15	Significance based quest	√	√	√
2.16	Show fields for the outcome	√	√	√
2.16.	The Scale	√	*	√
2.16.	Set of documents	√	*	*
2.17	Search Term Suggestions	√	√	√

2.18	Within scan, look for	√	√	√
3.	*Rating and Ordering*			
3.1	The sorting of several results by common keys (size/geography/license)	√	√	√
3.2	Pertinence of rating	√	√	√
4.	*Improved installations*			
4.1	Enriched content (Wikipedia link, external books-data mashup)	√	√	√
4.2	Completed Auto Recommendation	√	*	√
4.3	Checker for Spell	√	*	√
4.4	The New Directives	√	√	√
4.5	User's contribution	√	√	√
4.6	Integration of Social Networking Resources	√	√	√
4.7	Links that are persistent	√	*	√
4.8	Check for Stem	√	*	√
4.9	Mobile Interface With	√	√	√
4.10	Reporting on Statistics	√	*	√
4.11	Statistics for Use	√	*	√
4.12	OpenURL-Complete Text/OpenSearch-Complete Text Quest	√	*	√
4.13	Support with Google Analytics	√	*	√
4.14	Pictures/Multimedia Objects Support	√	√	√
4.15	Support to varieties of data/metadata format	√	*	√
4.16	Real-time collaboration with apps for resource management (e.g. ILS retain claim placement)	√	√	√
5.A.	*Improved submission*			
5.1	Catalog enhancement (e.g. Book-Screen/TOC/Review)	√	√	√
5.2	Visualized search resource	√	*	√
5.3	Related Objects displayed/e.g. Snipets	√	√	√

5.4	Visual organization (e.g. tag cloud)	√	*	√
5.5	Book-cover links for photos, feedback and other material from suppliers	√	√	√
5.6	Save resources on the ordered lists	√	√	√
5.7	Authors' Biographies	√	*	√
5.8	Citation type format/Integration with reference/citation management software from APA/MLA or Chicago	√	√	√
5.9	Map View/GPS-enabled Geographical Search and Chart View/Search Results Show Display	√	*	√
5.10	Updates of new versions of the key variable	√	*	√
5.11	Evaluation diagrams and maps	√	√	√
5.12	Feedback and plugin/complement modules	√	√	√
5.13	Inclusion from different sources	√	√	√
5.B.	*Boosted Records*			
5.1	Query or object URLs continuous persistent	√	√	√
5.2	Object labeling or documents	√	√	√
5.3	Recommended quest and vehicle reviews	√	√	√
5.4	Retrieve the forwarded documents	√	*	√
5.5	Lists/folders	√	√	√
5.C.	*Pursued protocols/standards*			
5.1	Joint Minimum Requirements (OAI-PMH)	√	√	√
5.2	(OA-DOI/OAI-ORE/RDF/ORCID) Progressive	√	*	√
5.3	For the LOD connection	√	√	√
5.4	RDF XML Support	√	*	√
5.5	Connector incorporation of the	√	*	√

	industrial exploration method (e.g. Summon serials solutions)			
5.6	Local with authentication	√	*	√
6.	*Personal data settings*			
6.1	Discovery of the past	√	√	√
6.2	Search Save	√	√	√
6.3	Item List Favorite	√	√	√
6.4	List of items/subjects concerned	√	√	√
6.5	Save, quote, booking and export performance	√	√	√
7.A.	*Social Networking Instruments Support/Integration*			
7.1	Social networks incorporation	√	√	√
7.2	Integrate social networking sites	√	√	√
7.3	Exporting Bibtex, Xml, RIS & ENDNOTE	√	√	√
7.4	Wiki-like e-resources integration	√	√	√
7.5	Integration of the library's locations and live documents	√	*	√
7.6	Help for Piwik Open Source Platform Analytics	√	*	√
7.7	Social networking bookmark for applications	√	√	√
7.8	Export of quotes	√	√	√
7.B.	*System warnings*			
7.1	RSS warnings-feeds	√	√	√
7.2	Atom’s replies	√	√	√
7.3	Chat facility assistance	√	√	√

Note: * Shows ‘‘in the pipeline”

Appendix 4: Training Set Codes


```
{
  "request_id": "6a8d11c063939d6555538e6c625ab884",
  "num_samples": 1,
  "features": {
    "feat_WOS5736.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.citation": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.homepage": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.license": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.post_processed": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.supervised_keys": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.builder_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.config_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.download_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.post_processing_size": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.dataset_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.size_in_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.description": {
      "dtype": "string",
      "_type": "Value"
    },
```

```
    "feat_WOS11967.citation": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.homepage": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.license": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.post_processed": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.supervised_keys": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.builder_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.config_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.download_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.post_processing_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.dataset_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.size_in_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.citation": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.homepage": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.license": {
      "dtype": "string",
      "_type": "Value"
```

```
    },
    "feat_WOS46985.post_processed": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.supervised_keys": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.builder_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.config_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.download_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.post_processing_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.dataset_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.size_in_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.version.version_str": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.version.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.version.major": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.version.minor": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.version.patch": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.version.version_str": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.version.description": {
      "dtype": "string",
```

```
      "_type": "Value"
    },
    "feat_WOS11967.version.major": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.version.minor": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.version.patch": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.version.version_str": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.version.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.version.major": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.version.minor": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.version.patch": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.features.input_data.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.input_data.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "text": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_1.dtype": {
```

```
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_1.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_1._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_2.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_2.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "target": {
      "names": [
        "Value"
      ],
      "_type": "ClassLabel"
    },
    "feat_WOS5736.splits.train.name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.splits.train.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.splits.train.num_examples": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.splits.train.dataset_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.checksum": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.input_data.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.input_data.id": {
      "dtype": "string",
```

```
    "_type": "Value"
  },
  "feat_WOS11967.features.input_data._type": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label.dtype": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label.id": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label._type": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label_level_1.dtype": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label_level_1.id": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label_level_1._type": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label_level_2.dtype": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label_level_2.id": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.features.label_level_2._type": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.splits.train.name": {
    "dtype": "string",
    "_type": "Value"
  },
  "feat_WOS11967.splits.train.num_bytes": {
    "dtype": "int64",
    "_type": "Value"
  },
  "feat_WOS11967.splits.train.num_examples": {
    "dtype": "int64",
    "_type": "Value"
  },
  "feat_WOS11967.splits.train.dataset_name": {
    "dtype": "string",
    "_type": "Value"
  },
```

```
    "feat_WOS11967.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.checksum": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.input_data.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.input_data.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.input_data._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_1.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_1.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_1._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_2.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_2.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_2._type": {
      "dtype": "string",
      "_type": "Value"
```

```
    },
    "feat_WOS46985.splits.train.name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.splits.train.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.splits.train.num_examples": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.splits.train.dataset_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.checksum": {
      "dtype": "string",
      "_type": "Value"
    },
    "split": {
      "dtype": "string",
      "_type": "Value"
    }
  }
}
```

Appendix 5: Test Set Codes


```
{
  "request_id": "fe5618742a4a7512cab01e2d2f697fc3",
  "num_samples": 46986,
  "features": {
    "feat_Y1": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_Y2": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_Y": {
      "dtype": "int64",
      "_type": "Value"
    },
    "target": {
      "names": [
        "CS ",
        "Civil ",
        "ECE ",
        "MAE ",
        "Medical ",
        "Psychology  ",
        "biochemistry "
      ],
      "_type": "ClassLabel"
    },
    "feat_area": {
      "dtype": "string",
      "_type": "Value"
    },
    "text": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_Abstract": {
      "dtype": "string",
      "_type": "Value"
    },
    "split": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.citation": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.homepage": {
      "dtype": "string",
      "_type": "Value"
    },
```

```
    "feat_WOS5736.license": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.post_processed": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.supervised_keys": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.builder_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.config_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.download_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.post_processing_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.dataset_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.size_in_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.citation": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.homepage": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.license": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.post_processed": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.supervised_keys": {
      "dtype": "string",
      "_type": "Value"
```

```
    },
    "feat_WOS11967.builder_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.config_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.download_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.post_processing_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.dataset_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.size_in_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.citation": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.homepage": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.license": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.post_processed": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.supervised_keys": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.builder_name":
{
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.config_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.download_size": {
```

```
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.post_processing_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.dataset_size": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.size_in_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.version.version_str": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.version.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.version.major": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.version.minor": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.version.patch": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.version.version_str": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.version.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.version.major": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.version.minor": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.version.patch": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.version.version_str": {
      "dtype": "string",
      "_type": "Value"
    },
```

```
    "feat_WOS46985.version.description": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.version.major": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.version.minor": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.version.patch": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.features.input_data.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.input_data.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_1.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_1.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_1._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_2.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.features.label_level_2.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.splits.test.name": {
      "dtype": "string",
      "_type": "Value"
```

```
    },
    "feat_WOS5736.splits.test.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.splits.test.num_examples": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.splits.test.dataset_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS5736.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS5736.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.checksum": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.input_data.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.input_data.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.input_data._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.label.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.label.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.label._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.label_level_1.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.label_level_1.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.label_level_1._type": {
      "dtype": "string",
```

```
      "_type": "Value"
    },
    "feat_WOS11967.features.label_level_2.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.label_level_2.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.features.label_level_2._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.splits.test.name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.splits.test.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.splits.test.num_examples": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.splits.test.dataset_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS11967.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS11967.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.checksum": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.input_data.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.input_data.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.input_data._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label.id": {
```

```
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_1.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_1.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_1._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_2.dtype": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_2.id": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.features.label_level_2._type": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.splits.test.name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.splits.test.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.splits.test.num_examples": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.splits.test.dataset_name": {
      "dtype": "string",
      "_type": "Value"
    },
    "feat_WOS46985.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.num_bytes": {
      "dtype": "int64",
      "_type": "Value"
    },
    "feat_WOS46985.download_checksums.https://data.mendeley.com/public-
files/datasets/9rw3vkcfy4/files/c9ea673d-5542-44c0-ab7b-
f1311f7d61df/file_downloaded.checksum": {
      "dtype": "string",
      "_type": "Value"
    }
```

```
    }
}
```

Appendix 6: Computational Output


```
{
    "output": [
        {
            "Y1": "0",
            "Y2": "12",
            "Y": "12",
            "Domain": "CS ",
            "area": " Symbolic computation  ",
            "keywords": " (2+1)-dimensional non-linear optical waves; erbium-doped
optical fibre; symbolic computation; soliton solution; soliton interaction
",
            "Abstract": "(2 + 1)-dimensional non-linear optical waves through the
coherently excited resonant medium doped with the erbium atoms can be described by
a (2 + 1)-dimensional non-linear Schrodinger equation coupled with the self-
induced transparency equations. For such a system, via the Hirota method and
symbolic computation, linear forms, one-, two-and N-soliton solutions are
obtained. Asymptotic analysis is conducted and suggests that the interaction
between the two solitons is elastic. Bright solitons are obtained for the fields E
and P, while the dark ones for the field N, with E as the electric field, P as the
polarization in the resonant medium induced by the electric field, and N as the
population inversion profile of the dopant atoms. Head-on interaction between the
bidirectional two solitons and overtaking interaction between the unidirectional
two solitons are seen. Influence of the averaged natural frequency. on the
solitons are studied: (1). can affect the velocities of all the solitons; (2)
Amplitudes of the solitons for the fields P and N increase with. decreasing, and
decrease with. increasing; (3) With. decreasing, for the fields P and N, one-peak
one soliton turns into the two-peak one, as well as interaction type changes from
the interaction between two one-peak ones to that between a one-peak one and a
two-peak one; (4) For the field E, influence of. on the solitons cannot be found.
The results of this paper might be of potential applications in the design of
optical communication systems which can produce the bright and dark solitons
simultaneously."
        },
        {
            "Y1": "5",
            "Y2": "2",
            "Y": "74",
            "Domain": "Medical ",
            "area": " Alzheimer's Disease  ",
            "keywords": " Aging; Tau; Amyloid; PET; Alzheimer's disease;
Polypathology        ",
            "Abstract": "(beta-amyloid (A beta) and tau pathology become
increasingly prevalent with age, however, the spatial relationship between the two
pathologies remains unknown. We examined local (same region) and non-local
(different region) associations between these 2 aggregated proteins in 46 normal
older adults using [F-18]AV-1451 (for tau) and [C-11]PiB (for A beta) positron
emission tomography (PET) and 1.5 T magnetic resonance imaging (MRI) images. While
local voxelwise analyses showed associations between PiB and AV-1451 tracer
largely in the temporal lobes, k-means clustering revealed that some of these
associations were driven by regions with low tracer retention. We followed this up
with a whole-brain region-by-region (local and non-local) partial correlational
analysis. We calculated each participant's mean AV-1451 and PiB uptake values
within 87 regions of interest (ROI). Pairwise ROI analysis demonstrated many
positive PiB AV-1451 associations. Importantly, strong positive partial
```

```
correlations (controlling for age, sex, and global gray matter fraction, p <.01)
were identified between PiB in multiple regions of association cortex and AV-1451
in temporal cortical ROIs. There were also less frequent and weaker positive
associations of regional PiB with frontoparietal AV-1451 uptake. Particularly in
temporal lobe ROIs, AV-1451 uptake was strongly predicted by NB across multiple
ROI locations. These data indicate that A beta and tau pathology show significant
local and non-local regional associations among cognitively normal elderly, with
increased PiB uptake throughout the cortex correlating with increased temporal
lobe AV-1451 uptake. The spatial relationship between A beta and tau accumulation
does not appear to be specific to A beta location, suggesting a regional
vulnerability of temporal brain regions to tau accumulation regardless of where AP
accumulates."
        },
        {
            "Y1": "4",
            "Y2": "7",
            "Y": "68",
            "Domain": "Civil ",
            "area": " Green Building  ",
            "keywords": "building automation systems; building energy efficiency;
daytime lighting; lighting control systems; EN 15232 standard       ",
            "Abstract": "Energy used for lighting is one of the major components
of total energy consumption in buildings. Nowadays, buildings have a great
potential to reduce their energy consumption, but to achieve this purpose
additional efforts are indispensable. In this study, the need for energy savings
evaluation before the implementation of lighting control algorithms for a
specified building is highlighted. Therefore, experimental tests have been carried
out in a university building with laboratories and other rooms, equipped with KNX
building automation system. A dimmable control strategy has been investigated,
dependent on daylight illuminance. Moreover, a relationship between external and
internal daylight illuminance levels has been evaluated as well. Based on the
experimental results, the authors proposed a method for the rough estimation of
electrical energy savings. Since, according to the EN 15232 standard, Building
Automation and Control Systems (BACS) play an important role in buildings' energy
efficiency improvements, the BACS efficiency factors from this standard have been
used to verify the experimental results presented in the paper. The potential to
reduce energy consumption from lighting in non-residential buildings by 28% for
offices and 24% for educational buildings has been confirmed, but its dependence
on specific building parameters has been discussed as well."
        },
        {
            "Y1": "1",
            "Y2": "10",
            "Y": "26",
            "Domain": "ECE ",
            "area": " Electric motor  ",
            "keywords": " NdFeB magnets; Electric motor; Electric vehicle; Hybrid
electric vehicle; Recycling; Rare earth elements        ",
            "Abstract": "(Hybrid) electric vehicles are assumed to play a major
role in future mobility concepts. Although sales numbers are increasing, little
emphasis has been laid on the recycling of some key components such as power
electronics or electric motors. Permanent magnet synchronous motors contain
considerable amounts of rare earth elements that cannot be recovered in
conventional recycling routes. Although their recycling could have large economic,
environmental, and strategic advantages, no industrial recycling for permanent
magnets is available in western countries at the moment. Regarding the essential
steps, dismantling of electric vehicles as well as the extraction of magnets from
the rotors, little has been published before. This paper therefore presents and
discusses different recycling approaches for the recycling of NdFeB magnets from
```

```
(hybrid) electric vehicles. Many results stem from the German research project \"Recycling of components and strategic metals of electric drive motors.\"."
        },
        {
            "Y1": "5",
            "Y2": "43",
            "Y": "115",
            "Domain": "Medical ",
            "area": " Parkinson's Disease  ",
            "keywords": " Parkinson's disease; dyskinesia; adenosine A(2A) receptor; striatum; glutamic acid decarboxylase        ",
            "Abstract": "(L)-3,4-Dihydroxyphenylalanine ((L)-DOPA) remains the primary pharmacological agent for the symptomatic treatment of Parkinson's disease (PD). However, the development of (L)-DOPA-induced dyskinesia (LID) limits the long-term use of (L)-DOPA for PD patients. Some data have reported that adenosine A(2A) receptor (A(2A)R) antagonists prevented LID in animal model of PD. However, the mechanism in which adenosine A(2A)R blockade alleviates the symptoms of LID has not been fully clarified. Here, we determined to knock out (KO) the gene of A(2A)R and explored the possible underlying mechanisms implicated in development of LID in a mouse model of PD. A(2A)R gene KO mice were unilaterally injected into the striatum with 6-hydroxydopamine (6-OHDA) in order to damage dopamine neurons on one side of the brain. 6-OHDA-lesioned mice were then injected once daily for 21 days with (L)-DOPA. Abnormal involuntary movements (AIMs) were evaluated on days 3, 8, 13, and 18 after (L)-DOPA administration, and real-time polymerase chain reaction and immunohistochemistry for glutamic acid decarboxylase (GAD) 65 and GAD67 were performed. We found that A(2A)R gene KO was effective in reducing AIM scores and accompanied with decrease of striatal GAD67, rather than GAD65. These results demonstrated that the possible mechanism involved in alleviation of AIM symptoms by A(2A)R gene KO might be through reducing the expression of striatal GAD67."
        },
        {
            "Y1": "0",
            "Y2": "0",
            "Y": "0",
            "Domain": "CS, ECE ",
            "area": " Computer vision  ",
            "keywords": " Tea category identification; computer vision; color histogram; wavelet packet entropy; winner-takes-all; radial basis function; artificial neural network; pattern recognition; support vector machine        ",
            "Abstract": "(Objective) In order to increase classification accuracy of tea-category identification (TCI) system, this paper proposed a novel approach. (Method) The proposed methods first extracted 64 color histogram to obtain color information, and 16 wavelet packet entropy to obtain the texture information. With the aim of reducing the 80 features, principal component analysis was harnessed. The reduced features were used as input to generalized eigenvalue proximal support vector machine (GEPSVM). Winner-takes-all (WTA) was used to handle the multiclass problem. Two kernels were tested, linear kernel and Radial basis function (RBF) kernel. Ten repetitions of 10-fold stratified cross validation technique were used to estimate the out-of-sample errors. We named our method as GEPSVM + RBF + WTA and GEPSVM + WTA. (Result) The results showed that PCA reduced the 80 features to merely five with explaining 99.90% of total variance. The recall rate of GEPSVM + RBF + WTA achieved the highest overall recall rate of 97.9%. (Conclusion) This was higher than the result of GEPSVM + WTA and other five state-of-the-art algorithms: back propagation neural network, RBF support vector machine, genetic neural-network, linear discriminant analysis, and fitness-scaling chaotic artificial bee colony artificial neural network."
        },
        {
```

```
            "Y1": "6",
            "Y2": "0",
            "Y": "125",
            "Domain": "biochemistry ",
            "area": " Molecular biology  ",
            "keywords": " Brain-machine interfaces; Electrical stimulation of the
brain; Transcranial magnetic stimulation; Trends; Psychiatry        ",
            "Abstract": "(R)ecent advances in deep brain stimulators and brain -
machine interfaces have greatly expanded the possibilities of neuroprosthetics and
neuromodulation. Together with advances in neuroengineering, nanotechnology,
molecular biology and material sciences, it is now possible to address fundamental
questions in neuroscience in new, more powerful ways. It is now possible to apply
these new technologies in ways that range from augmenting and restoring function
to neuromodulation modalities that treat neuropsychiatric disorders. Recent
developments in neuromodulation methods offer significant advantages and potential
clinical benefits for a variety of disorders. Here we describe the current state
of the art in neuromodulation methods, and some advances in brain -machine
interfaces, describing the advantages and limitations of the clinical applications
of each method. The future applications of these new methods and how they will
shape the future of psychiatry and medicine, along with safety and ethical
implications, are also discussed. (c) 2017 Elsevier Inc. All rights reserved."
        },
        {
            "Y1": "1",
            "Y2": "10",
            "Y": "27",
            "Domain": "ECE ",
            "area": " Satellite radio  ",
            "keywords": " Software-defined radio; AIS receiver; VHF data exchange
services; Satellite radio applications; Flexible RF front end        ",
            "Abstract": "(T)his paper presents the concept of a software-defined
radio with a flexible RF front end. The design and architecture of this system, as
well as possible application examples will be explained. One specific scenario is
the operation in maritime frequency bands. A well-known service is the Automatic
Identification System (AIS), which has been captured by the DLR mission AISat, and
will be chosen as a maritime application example. The results of an embedded
solution for AIS on the SDR platform are presented in this paper. Since there is
an increasing request for more performance on maritime radio bands, services like
AIS will be enhanced by the International Association of Marine Aids to Navigation
and Lighthouse Authorities (IALA). The new VHF Data Exchange Service (VDES) shall
implement a dedicated satellite link. This paper describes that the SDR with a
flexible RF front end can be used as a technology demonstration platform for this
upcoming data exchange service."
        },
        {
            "Y1": "3",
            "Y2": "4",
            "Y": "56",
            "Domain": "MAE ",
            "area": " Fluid mechanics  ",
            "keywords": " plumes; thermals; turbulence modelling; turbulence
simulation        ",
            "Abstract": "A two-dimensional Boussinesq model of thermal convection
is investigated by numerical simulations. The turbulent heat transport is mainly
due to thermal plumes. It is shown that their thin, wrinkled interfaces control
the statistics of large temperature excursions. Away from the interfaces, the
temperature fluctuations are strongly suppressed and the field in those regions is
well mixed. Their statistical signature is in the low-order moments of the
temperature increments, whose scaling exponents depend linearly on the order."
```

```
        },
        {
            "Y1": "2",
            "Y2": "13",
            "Y": "46",
            "Domain": "Psychology  ",
            "area": " Prenatal development  ",
            "keywords": "D cell; Gastrin; G cell; Human fetus;
Immunohistochemistry; Radioimmunoassay; Ultrastructure    ",
            "Abstract": "The gestational time of appearance of gastrin and
somatostatin in the human fetal stomach, duodenum and pancreas was examined.
Immunoreactive gastrin (IRG) is detected in antral, duodenal and pancreatic
extracts of a 7.0-cm (crown-heel length) fetus. More IRG is extracted from the
duodenum than the antrum. Duodenal IRG concentration from fetuses of 16.0--26.0 cm
are higher than younger fetal and adult concentrations. Antral IRG concentrations
are one tenth of the adult contents. Very small IRG concentrations are present in
the human fetal pancreas. Gastrin immunohistochemical staining is positive first
in duodenal (6.5-cm fetus) and later in antral (12.5-cm fetus) mucosa; pancreatic
tissue is negative for gastrin immunohistochemistry. Type IV cells are encountered
in antral and duodenal mucosa of 4.0-cm fetuses; other endocrine cells appear with
fetal growth. Not until much later in gestation (21.0 cm) do typical G cells
appear. These results suggest that early in fetal life gastrin is produced by the
type IV cell. Somatostatin immunohistochemical staining is positive in stomach,
duodenum and pancreas in 6.5-cm fetuses. Immature D cells are found in antral and
duodenal mucosa of 5.0-cm fetuses and mature D cells in 11.0-cm fetuses."
        },
        {
            "Y1": "5",
            "Y2": "51",
            "Y": "123",
            "Domain": "Medical ",
            "area": " Sprains and Strains  ",
            "keywords": "emergency department, epidemiology, recreation, sports,
traumatic brain injury        ",
            "Abstract": "Sports- and recreation-related traumatic brain injuries
(SRR-TBIs) are a growing public health problem affecting persons of all ages in
the United States. To describe the trends of SRR-TBIs treated in US emergency
departments (EDs) from 2001 to 2012 and to identify which sports and recreational
activities and demographic groups are at higher risk for these injuries. Data on
initial ED visits for an SRR-TBI from the National Electronic Injury Surveillance
System-All Injury Program (NEISS-AIP) for 2001-2012 were analyzed. NEISS-AIP data
are drawn from a nationally representative sample of hospital-based EDs. Cases of
TBI were identified from approximately 500 000 annual initial visits for all
causes and types of injuries treated in EDs captured by NEISS-AIP. Numbers and
rates by age group, sex, and year were estimated. Aggregated numbers and
percentages by discharge disposition were produced. "
        },
        {
            "Y1": "6",
            "Y2": "5",
            "Y": "130",
            "Domain": "biochemistry ",
            "area": " Enzymology  ",
            "keywords": " enzymology; factor IXa; hemophilia; structure;
substrate; thrombosis        ",
            "Abstract": ". Background: A key feature of factorIXa is its
allosteric transformation from an enzymatically latent form into a potent
procoagulant. Although several small molecules have been found to be capable of
partially affecting FIXa function (i.e. ethylene glycol, Ca2+, and low molecular
```

weight heparin [LMWH]), the resulting modest changes in peptidolytic activity have made the study of their mechanisms of action challenging. As these effects provide hints about potential regulatory forces that may be operational in the full expression of FIXa coagulant activity, their description remains of great interest. Studies of crystal structures have yielded insights into the structural changes induced by these effectors, but there remains a paucity of information to correlate any given structural change with specific consequences for FIXa function.Objectives: To correlate structural changes induced by these modulators with defined consequences for FIXa substrate discrimination and function.Methods: A peptidomics-based mass spectrometry (MS) approach was used to examine the patterns of hydrolysis of four combinatorial chemistry-derived pentapeptide libraries by FIXa under various conditions in a soluble, active enzyme system.Results: Ethylene glycol specifically altered the S3 subsite of FIXa to render it more tolerant to side chains at the P3 substrate position, whereas Ca2+ enhanced tolerance at the S2 subsite. In contrast, LMWH altered both the S2 and S1' subsites.Conclusions: These results demonstrate the role of plasticity in regulating FIXa function with respect to discrimination of extended substrate sequences, as well as providing crucial insights into active site modulations that may be capitalized on by various physiologic cofactors of FIXa and in future drug design."
        },
        {
            "Y1": "6",
            "Y2": "8",
            "Y": "133",
            "Domain": "biochemistry ",
            "area": " Southern blotting ",
            "keywords": "coagulation factor VIII gene, hemophilia A, mutation, bioinformatics, inhibitors, China   ",
            "Abstract": "Hemophilia A (HA) is an X-linked bleeding disorder caused by heterogeneous mutations in the factor VIII gene (F8). Our aim is to identify the causative mutations in a large HA cohort from China. We studied 216 unrelated HA families. Molecular analyses of F8 were performed using a combination of molecular techniques, including polymerase chain reaction, direct sequencing, and multiplex ligation-dependent probe amplification. The deleterious consequences of the unreported missense mutations were evaluated using various bioinformatics approaches. Causative mutations in F8 were identified in 209 families, intron 22 inversion (Inv22) was identified in 89 severe families, and intron 1 inversion (Inv1) was positive in 5 severe families; 95 mutations were detected among 115 noninversion families, of which 42 were novel, including 29 null variations and 13 missense mutations for which causality was demonstrated via bioinformatics. Among the 53 previously reported mutations, more nonsense (5 of 9) and missense (10 of 26) mutation sites were found to occur at Arginine (Arg) sites and multiple small deletions/insertions (5 of 10) located within the poly-A runs of the B domain. The majority of these sequence variants frequently recurred in the database. The odds ratios for the likelihood of developing inhibitors significantly increased in the presence of nonsense mutation. Our F8 defect spectrum was heterogeneous. Small deletions/insertions in the poly-A runs of the B domain and nonsense and missense mutations at Arg sites were identified as mutation hot spots. Nonsense mutation increased the risk of developing inhibitors."
        },
        {
            "Y1": "6",
            "Y2": "8",
            "Y": "133",
            "Domain": "biochemistry ",
            "area": " Southern blotting ",
            "keywords": " age-related macular degeneration; genetics; retina; telomere   ",

```
            "Abstract": ". Purpose: To evaluate the association between telomere
length and age-related macular degeneration (AMD). Methods: Circulating leucocyte
telomere length and the proportion of telomeres <5kb were analysed in blood DNA
samples taken from 121 patients with exudative AMD (83%), large drusen (14%) or
central geographic atrophy (3%). Controls consisted of 77 age-matched subjects
without AMD. The AMD status was assessed by a masked analysis of fundus
photographs or angiographs. Telomere length was measured by Southern blotting.
Results: Mean (SD) telomere length was 7.76kb (0.68) in AMD patients and 7.83
(0.69) in controls (p=0.485). The corresponding proportions of telomeres <5kb were
10.60 (2.76) and 10.05 (2.64) (p=0.197). In this material, there was no
correlation between telomere length and age, gender or smoking status. There were
no differences between the major AMD risk single-nucleotide polymorphisms (SNPs)
of the CFH, HTRA1 or C3 genes, expect for somewhat longer telomeres in controls
with the C3 risk SNP. There were no differences in telomere length between
patients with drusen or exudative AMD. Conclusions: Telomere length is not
associated with exudative AMD or high-risk drusen."
        },
        {
            "Y1": "1",
            "Y2": "10",
            "Y": "27",
            "Domain": "ECE, CS ",
            "area": " Satellite radio  ",
            "keywords": " statistical tomography; high-latitude ionosphere;
kilometer-scale irregularities        ",
            "Abstract": "[1] First results are reported on statistical tomography
of kilometer-scale irregularities in the F layer high-latitude ionosphere from
amplitude data of satellite radio probing. Basic formulae for statistical
tomography of three-dimensionally (3-D) anisotropic small-scale irregularities are
presented. It is shown that 3-D anisotropy disguises spatial distribution of
irregularities but is not an insuperable difficulty for tomographic
reconstruction. An example is shown of imaging the spatial distribution of the
variance of electron density fluctuations over the Kola peninsula in February
1996. Iterative procedure of tomographic inversion was used in the reconstruction.
Further steps of applying statistical tomographic approach are outlined."
        },
        {
            "Y1": "5",
            "Y2": "38",
            "Y": "110",
            "Domain": "Medical ",
            "area": " Cancer  ",
            "keywords": " Fluorocholine; Positron emission tomography;
Hepatocellular carcinoma; Phosphatidylcholine; Fatty acids        ",
            "Abstract": "[F-18]fluorocholine PET/CT can detect hepatocellular
carcinoma (HCC) based on imaging the initial steps of phosphatidylcholine
synthesis. To relate the diagnostic performance of [F-18]fluorocholine positron
emission tomography (PET)/x-ray computed tomography (CT) to the phospholipid
composition of liver tumors, radiopathologic correspondence was performed in
patients with early-stage liver cancer who had undergone [F-18]fluorocholine
PET/CT before tumor resection. Tumor and adjacent liver were profiled by liquid
chromatography mass spectrometry, quantifying phosphatidylcholine species by mass-
to-charge ratio. For clinical-radiopathologic correlation, HCC profiles were
reduced to two orthogonal principal component factors (PCF1 and PCF2) accounting
for 80 % of total profile variation. Tissues from 31 HCC patients and 4
intrahepatic cholangiocarcinoma (ICC) patients were analyzed, revealing
significantly higher levels of phosphocholine, CDP-choline, and highly saturated
phosphatidylcholine species in HCC tumors relative to adjacent liver and ICC
tumors. Significant loading values for PCF1 corresponded to phosphatidylcholines
```

containing poly-unsaturated fatty acids while PCF2 corresponded only to highly saturated phosphatidylcholines. Only PCF2 correlated significantly with HCC tumor-to-liver [F-18]fluorocholine uptake ratio (rho = 0.59, p < 0.0005). Sensitivity for all tumors based on an abnormal [F-18]fluorocholine uptake ratio was 93 % while sensitivity for HCC based on increased tumor [F-18]fluorocholine uptake was 84 %, with lower levels of highly saturated phosphatidylcholines in tumors showing low [F-18]fluorocholine uptake. Most HCC tumors contain high levels of saturated phosphatidylcholines, supporting their dependence on de novo fatty acid metabolism for phospholipid membrane synthesis. While [F-18]fluorocholine PET/CT can serve to identify these lipogenic tumors, its imperfect diagnostic sensitivity implies metabolic heterogeneity across HCC and a weaker lipogenic phenotype in some tumors."
        },
        {
            "Y1": "6",
            "Y2": "5",
            "Y": "130",
            "Domain": "biochemistry ",
            "area": " Enzymology ",
            "keywords": " [FeFe]-hydrogenase; kinetics; enzymology; electron bifurcation; time-resolved infrared spectroscopy   ",
            "Abstract": "[FeFd]-hydrogenases catalyze the reversible production of hydrogen gas from protons and electrons, but the mechanism of catalysis is still the subject of debate. In this report, we describe a pre-steady-state photoreduction methodology for correlating putative intermediates to the reactivity of an [FeFe]-hydrogenase from Thermotoga maritima. In this method, MV+ is rapidly formed photochemically by pulsed laser excitation and the intermolecular ET and active site dynamics are followed by nanosecond time-resolved visible and infrared spectroscopy, respectively. The results kinetically validate the H-ox H-red and H-sred intermediates, strongly supporting a vast body of literature on their involvement in proton reduction."
        },
        {
            "Y1": "6",
            "Y2": "7",
            "Y": "132",
            "Domain": "biochemistry ",
            "area": " Northern blotting ",
            "keywords": "berberine; non-alcoholic fatty liver disease; oxidative stress; sirtuin 1; uncoupling protein 2   ",
            "Abstract": "Objective; To investigate the involvement of sirtuin 1 (SIRT1)-uncoupling protein 2 (UCP2) pathway in the development of non-alcoholic fatty liver disease and whether berberine exerts its effects by regulating this pathway."
        },
        {
            "Y1": "5",
            "Y2": "50",
            "Y": "122",
            "Domain": "Medical ",
            "area": " Sports Injuries ",
            "keywords": " Elastic therapeutic tape; Material properties; Tension test   ",
            "Abstract": "[Purpose] Elastic therapeutic tape has been widely used for rehabilitation and treatment of sports injuries. Tapes with different elastic properties serve different treatment purposes with inappropriate tension reducing tape effectiveness. Many tapes are available in the market, but studies on tape properties are limited. The aim of this study was to examine the material properties of elastic therapeutic tape. [Subjects and Methods] Brands of elastic

therapeutic tape included KinesioTex (R), ATex, Mueller, 3M, and ThaiTape. The Material Testing System Insight (R) 1 Electromechanical Testing Systems was used to apply a tensile force on elastic therapeutic tape. Ten specimens of each brand were tested. Stress, load, and Young's modulus at 25%, 50%, 75%, 100%, and maximum point were collected. One-way analysis of variance with post hoc testing was used to analyze tape parameters. [Results] Maximum elongation and Young's modulus at all percentages were significantly different between brands. There were no differences in maximum load and maximum stress. [Conclusion] Mechanical properties are different for commercial elastic therapeutic tapes. Physiotherapists and other clinicians should be aware of mechanical tape properties to correctly apply kinesio tape."
        },
        {
            "Y1": "5",
            "Y2": "50",
            "Y": "122",
            "Domain": "Medical ",
            "area": " Sports Injuries ",
            "keywords": " Sports injury; Athletes; Postural stability   ",
            "Abstract": "[Purpose] The aim of this study was to analyze stabilometry in athletes during an indoor season in order to determine whether injured athletes show different stabilometric values before injury than non-injured athletes in two different training periods (volume and pre-competition periods). [Subjects] The subjects were 51 athletes from Unicaja athletic club who trained regularly. [Methods] At the end of the preseason and volume periods, athletes were subjected to bipodal and monopodal stabilometry. In addition, all injuries happening in the periods after performing stabilometry (volume and pre-competition periods) were tracked. [Results] Variance analysis of bipodal stabilometric measurements taken at the end of the preseason period showed that athletes with higher values for the center-of-pressure spread variables suffered injuries during the volume period. The right-leg monopodal stabilometric measurements taken at the end of the volume period showed that athletes with higher values in the center-of-pressure position variables suffered injuries during the pre-competition period. [Conclusion] Athletes showing the worst values for center-of-pressure spread variables are more prone to sports injuries in the subsequent training period. In monopodal measurements, athletes with poorer mediolateral stability were more prone to injuries in the subsequent training period."
        },
        {
            "Y1": "5",
            "Y2": "48",
            "Y": "120",
            "Domain": "Medical ",
            "area": " Senior Health ",
            "keywords": " Ankle range of motion; Balance; TETRAX   ",
            "Abstract": "[Purpose] This study examined the influence of range of motion of the ankle joints on elderly people's balance ability. [Methods] We conducted a four-week experiment using 60 out of the 89 elderly people who used B senior health facility, as subjects. TETRAX stability, synchronization were conducted to measure balance ability. We examined correlations between TETRAX balance and ankle joint ROM. Student's t-test and Pearson correlation coefficients were used for the statistical analyses. [Results] There were significant correlations between measurement items of TETRAX stability and synchronization and movement of the ankle joint. [Conclusion] It was found that as had plate was stable in measurement using ankle joint ROM and TETRAX after closing eyes, balance was maintained with Dorsi Flexion."
        },
        {

```
            "Y1": "0",
            "Y2": "5",
            "Y": "5",
            "Domain": "CS, ECE ",
            "area": " Computer graphics  ",
            "keywords": "Stereoscopic images; Virtual reality; Method comparison;
Visualization techniques; Architectural space; Daylight       ",
            "Abstract": "Visualization and presentation techniques are
experiencing a rapid development and application in the architectural profession
and beyond. As such, Stereoscopic Images (SI) and Virtual Reality (VR), both
advanced visualization techniques, have found their way in daylighting research.
This paper explores the correspondence between these two techniques using both
quantitative and qualitative methods. The study focused on two small rooms with
the same dimensions, with alternately white and black surfaces, and three
different windows sizes. Seven architectural qualities were studied: Pleasantness,
Calmness, Interest, Excitement, Complexity, Spaciousness, and Amount of View. The
attributes were evaluated by 20 participants using a Likert-type scale. The
collected data was analyzed using Wilcoxon signed-rank tests and the Bland-Altman
method. The results revealed that the attributes Pleasantness, Calmness, Interest,
and Complexity in rooms with both white and black surfaces, and Excitement and
Spaciousness in a room with white surfaces are evaluated similarly in both VR and
SI. The attributes Excitement and Spaciousness in a room with black surfaces and
the Amount of View in rooms with both white and black surfaces were not evaluated
similarly with both methods. In addition, qualitative results indicated that the
visualization technique affects how a space is perceived. Indeed, the Virtual
Reality's nature as an immersive environment provides the feeling of ‘presence’
which does not apply to SI. The results of this paper can help researchers working
with daylighting in buildings in selecting the appropriate visualization technique
to reduce experimental and logistical constraints caused by varying daylight
conditions."
        },
        {
            "Y1": "1",
            "Y2": "4",
            "Y": "21",
            "Domain": "ECE ",
            "area": " Digital control  ",
            "keywords": " active power filter; vector resonant controller; digital
design        ",
            "Abstract": "\\There are many factors that affect the performance of
active power filter, which is a key factor for command current tracking control.
Aiming at the deficiency of traditional PR controller, this paper puts forward an
improved PR controller-vector resonant controller. The controller has the
advantages of stronger anti-interference ability and greater gain at the resonance
point, and it can improve the accuracy of harmonic compensation in the current
compensation control. In this paper, the digital controller in the digital control
system has been given a detailed design. Finally, the experimental results show
that the control strategy can further reduce the harmonic content and improve the
compensation accuracy."
        },
        {
            "Y1": "6",
            "Y2": "2",
            "Y": "127",
            "Domain": "biochemistry ",
            "area": " Human Metabolism  ",
            "keywords": " ALUMINUM; GALLIUM; INJECTION; RETENTION; EXCRETION
",
```

```
            "Abstract": "1 Al-26 and Ga-67 were given as citrates to a healthy
male volunteer by intravenous injection. The retention of both tracers was studied
by body radioactivity measurement. Levels in blood and excreta were determined by
gamma-ray spectrometry and/or accelerator mass spectrometry. 2 More than half of
the Al-26 had left the blood after 15 min and the decline continued, leaving <1%
in blood after 2 d; the losses occurred both to renal excretion and through uptake
by other compartments. Estimated excretion up to 13 d was 83% (urine) and 1.8%
(faeces). Whole-body retention of 15% at 13 d declined to similar to 4% at 1178 d,
when the daily reduction corresponded to a biological half-life of 7 y, suggesting
that sustained intake of dietary aluminium may lead to a progressively increasing
internal deposit. 3 The metabolism of Ga-67 differed markedly from that of Al-26
in all aspects studied."
        },
        {
            "Y1": "6",
            "Y2": "6",
            "Y": "131",
            "Domain": "biochemistry ",
            "area": " Polymerase chain reaction  ",
            "keywords": " 1; 2-dicloropropane; dichloromethane; gpt delta rat; in
vivo mutagenicity; liver        ",
            "Abstract": "1,2-Dichloropropane (1,2-DCP) and dichloromethane (DCM)
are possible causative agents associated with the development of
cholangiocarcinoma in employees working in printing plant in Osaka, Japan.
However, few reports have demonstrated an association between these agents and
cholangiocarcinoma in rodent carcinogenicity studies. Moreover, the combined
effects of these compounds have not been fully elucidated. In the present study,
we evaluated the in vivo mutagenicity of 1,2-DCP and DCM, alone or combined, in
the livers of gpt delta rats. Six-week-old male F344 gpt delta rats were treated
with 1,2-DCP, DCM or 1,2-DCP+DCM by oral administration for 4weeks at the dose
(200mgkg(-1) body weight 1,2-DCP and 500mgkg(-1) body weight DCM) used in the
carcinogenesis study performed by the National Toxicology Program. In vivo
mutagenicity was analyzed by gpt mutation/Spi(-) assays in the livers of rats. In
addition, gene and protein expression of CYP2E1 and GSTT1, the major enzymes
responsible for the genotoxic effects of 1,2-DCP and DCM, were analyzed by
quantitative polymerase chain reaction and western blotting. Gpt and Spi(-)
mutation frequencies were not increased by 1,2-DCP and/or DCM in any group.
Additionally, there were no significant changes in the gene and protein expression
of CYP2E1 and GSTT1 in any group. These results indicated that 1,2-DCP, DCM and
1,2-DCP+DCM had no significant impact on mutagenicity in the livers of gpt delta
rats under our experimental conditions. Copyright (c) 2016 John Wiley & Sons,
Ltd."
        },
        {
            "Y1": "5",
            "Y2": "36",
            "Y": "108",
            "Domain": "Medical ",
            "area": " Multiple Sclerosis  ",
            "keywords": " Immunomodulation; Intestinal absorption; Blood-brain
barrier diffusion; Cannabinoid CB2 receptor; 1   8-Naphthyridine-3-carboxamide
derivative; Neurodegeneration       ",
            "Abstract": "1,2-Dihydro-4-hydroxy-2-oxo-1,8-naphthyridine-3-
carboxamide derivative VL15 has been recently developed as a selective cannabinoid
CB2 receptor compound. Given the high selectivity of this compound at the
cannabinoid CB2 receptor and the well-known protective function of this receptor
in neurological disorders with autoimmune component like multiple sclerosis, we
assessed the immunomodulatory properties of VL15. We assessed on activated
peripheral blood mononuclear cells), proliferation and viability, cell cycle
```

progression and measured activation markers and the expression of phosphorylated proteins. We found that VL15 reduces PBMC proliferation slightly affecting cell vitality, blocks the cell cycle progression and down-regulates the levels of T cell activation markers as well as the expression of phosphorylated proteins, NF-kB, IKK alpha beta, IKB alpha, ERK and Akt. VL15 was also used in drug permeability assays on Caco-2 cell line to evaluate its oral bioavailability and on MDCKII-hMDR1 cell lines to estimate its propensity to cross the blood-brain barrier by passive diffusion, in order to potentially maintain its efficiency on the infiltrating auto-reactive lymphocytes in the central nervous system. In these models, VL15 showed high intestinal absorption and good blood-brain barrier penetration. Our findings suggest that VL15, by controlling the immune response, might find potential application as orally administered drug in pathologies like multiple sclerosis. (C) 2017 Elsevier B.V. and Societe Francaise de Biochimie et Biologie Moleculaire (SFBBM). All rights reserved."
        },

RESEARCH OUTPUT

Review Article 1

1. A. Kato, M. Kisangiri, and S. Kaijage, "A review development of digital library resources at university level," *Education Research International,* 2021.

Research Article 1

2. A. Kato, E. W. Steinhauer, M. Kisangiri, S. Kaijage, "Determining the Extent of Use of Electronic Information Resources by the Online Consumers," *Qualitative and Quantitative Methods in Libraries,* 2021.

Research Article 2

3. A. Kato, E. L. Temba, E. W. Steinhauer, M. Kisangiri, S. Kaijage, "Development of a User-Friendly Digital Library Resource Discovery Tool to Mitigate Learners' Technical Hitches," *Springer, Cham,* 2024.