Chapter II

Review of Related Literature and Studies

Related Literature

An overview of pertinent research archiving systems, academic repositories, and digital document management is provided in this section. The material is arranged thematically according to important facets of academic cooperation and thesis archiving.

Design, Development, and Testing of an Academic Repository

Handeli and Robila [1] investigated the creation of an academic repository designed to facilitate the archiving and retrieval of student projects for academic departments. The system is built on a secure database and a user-friendly interface, enabling easy access to academic resources such as Master's projects. By incorporating citation style generation, it enhances usability for both instructors and students. The repository's search engine-like features allow users to locate documents quickly using keyword queries. Initially developed for a single academic unit, the system offers a scalable and adaptable solution that could be expanded for campus-wide implementation.

Online Educational Platform Server and Method for Managing Students Database

Jeon [2] developed an online educational platform server designed to connect students and instructors, facilitating thesis writing through enhanced communication and document management. The platform includes a student database for storing documents and tracking

progress, as well as a communication module that links terminals to an external server. A document-sharing module allows for seamless file exchange, automatically saving uploads and retrieving relevant documents from external sources. This system improves student-lecturer collaboration and supports effective thesis writing by preventing plagiarism and teaching proper dissertation techniques.

Archives Management System in Colleges and Universities

The evolution of archival information management has become a significant area of study in higher education. A research study [3] examined the challenges posed by the lag in archives information management, which hampers university progress. To address these challenges, a solution was proposed using the J2EE framework and Oracle databases. The system leverages object-oriented techniques and software engineering principles to optimize archive management, offering an efficient and scalable approach to meet the needs of colleges and universities.

Management Skills and Learning Strategies Among University Graduands

Pulmones [4] conducted a study at the University of Southern Philippines Foundation (USPF) to explore the relationship between university graduands' learning strategies and their management skills. The study involved 263 respondents who completed questionnaires assessing these areas. Results indicated that graduands demonstrated strong management skills, with specific focus on motivating (mean = 3.43) and communicating (mean = 3.37), both interpreted as "always." Similarly, their use of various learning strategies yielded a grand mean of 3.16, interpreted as "agree." A significant correlation was identified between management skills and learning strategies, rejecting the null hypothesis. The findings suggested that higher management skills were associated with better learning strategies. Despite their strong management abilities, the

study recommended further improvement in learning strategies through the Learning Enhancement and Practice (L.E.A.P.) program.

Academic Integrity in Online Assessment

Holden, Norris, and Kuhlmeier [5] reviewed current research on academic integrity in higher education, focusing on assessment practices in online courses. The review addresses why students engage in academically dishonest behaviors and examines methods to reduce such behaviors. The authors hope this review will aid instructors and administrators in their decision-making process regarding online evaluations and encourage future study that will form the foundation of evidence-based practices. The paper highlights individual factors, institutional factors, medium-related factors, and assessment-specific factors that contribute to academic dishonesty.

The Practical Models of Electronic Archiving: An Evaluation Study

Alhazmi [6] conducted an evaluation study on practical models of electronic archiving, emphasizing the challenges faced by institutions using manual archiving, such as low productivity and poor services. The study highlighted the necessity of transitioning to electronic archiving systems, especially during the COVID-19 pandemic, which underscored the importance of efficient information management. Using a critical appraisal methodology, ten studies from 2017 to 2020 were analyzed to assess the benefits and challenges of electronic archiving models. The findings revealed that electronic archiving significantly improves institutional operations, promotes development and design innovation, and contributes to sectoral progress. Furthermore, the study emphasized the need for future research on how these models

influence sectoral development and encouraged scientific research focused on improving electronic archiving systems.

Archiving and Document Management at Taibah University: A Case Study

Alokluk [7] conducted a qualitative descriptive case study to examine how digital archiving and document management are organized at Taibah University. The study aimed to address the needs of researchers, professors, administrators, and students while ensuring secure data storage and access. The findings revealed that Taibah University follows a specific strategy for managing electronic documents but does not employ a unique data management system. Archiving is applied to a limited number of documents, and data security is guaranteed through the university's policies. The study highlights the importance of efficient digital archiving systems in educational institutions and their role in improving information management practices.

Effectiveness of Digital Archiving System of the Selected Offices in the Provincial Government of Laguna: A Basis for a Proposed Enhanced Archiving System

Acebron et al. [8] conducted a descriptive-correlational study to evaluate the effectiveness of digital archiving systems in selected offices of the Provincial Government of Laguna. The research aimed to bridge the gap in understanding how user awareness influences system effectiveness. Using a total population sampling approach, 55 respondents from key offices participated by completing questionnaires distributed via Google Forms. The findings revealed that respondents were fully aware of the system's duration, accessibility, and usage in terms of deployment and frequency. The system demonstrated high effectiveness in sustainability, security, and storage capability. Correlation analysis showed that increased user awareness positively influenced system effectiveness, particularly in promoting sustainability and security.

However, no significant relationship was observed between awareness of usage frequency and storage capability. The study concluded that enhancing user awareness could further improve the effectiveness of digital archiving systems and proposed an enhanced archiving system as a solution.

Standardization of Records Management System in Local Government Units (Phase 1: Assessment of the Records Management System of Local Government Units)

Romero [9] conducted a study to assess the Records Management System (RMS) practices in the Local Government Units (LGUs) of Laguna, Philippines, as part of Phase 1 of a broader initiative to standardize RMS across municipalities. The study evaluated nine RMS parameters, including Integrated Records Management Systems (IRMS), record management programs, regulatory environment, functionality and components, processes and controls, ARMA standards, filing methods and procedures, and indexing rules. Data were gathered from 160 administrative professionals across eight municipal governments. The results indicated a "Very Satisfactory" rating for the RMS parameters, demonstrating strong adherence to standard measures. However, one municipality showed significant differences in RMS practices compared to the others. The findings underscored the need for standardized RMS practices aligned with international standards to address inconsistencies across municipalities.

Science Metadata Management, Interoperability, and Data Citations of the National Institute of Polar Research, Japan

Kanao et al. [10] examined the metadata management and interoperability practices of the Polar Data Centre (PDC) at the National Institute of Polar Research (NIPR) in Japan. As part of the National Antarctic Data Centre and the Science Committee on Antarctic Research, the PDC

manages polar science data. During the International Polar Year (IPY 2007–2008), a significant volume of multidisciplinary data and metadata was compiled. The study highlighted collaborations with global data organizations, including the Global Change Master Directory, Polar Information Commons, and the World Data System under the International Council for Science. Metadata interoperability was enhanced through initiatives like the Open Archives Initiative Protocol for Metadata Harvesting and integration with systems such as the Data Integration and Analysis System Program and the Polar Data Catalogue of Canada. The metadata were modified using an automatic attributing system and DataCite through the Japan Link Center to assign Digital Object Identifiers (DOIs). The study emphasized the importance of interoperable metadata linkage and data citation systems for long-term preservation and publication of polar data.

Representation Matters: An Exploration of the Socio-Economic Impacts of ICT-Enabled Public Value in the Context of Sub-Saharan Economies

Arshad and Ali [11] investigated the socio-economic impacts of ICT-enabled public value in Sub-Saharan economies using a modified Networked Readiness Index (NRI) framework. The study examined whether ICT capabilities impact public value creation and whether public value is associated with the socio-economic effects of ICT capabilities. Public Value was measured using two proxies: perception-based measures (World Government Indicators) and objective measures (Cost of Business Startup Procedures). A six-step multi-method methodology, including Cluster Analysis, Correlation Analysis, Decision Trees Induction, Data Envelopment Analysis, Association Rules Mining, and Ordinary Least Squares regression, was applied to 26 Sub-Saharan economies. The findings revealed that economies with better-developed ICT capabilities are more efficient in converting these capabilities into public value. Key factors such

as affordability readiness, skills readiness, political and regulatory environment, and business usage were identified as critical for generating public value efficiently.

Usability Evaluation of a Research Repository and Collaboration Web Site

Zhang et al. [12] conducted an empirical usability evaluation of the Human-Animal Bond Research Initiative (HABRI) Central, an open-access research repository and collaboration platform for human-animal bond researchers. The system repurposed features of the HUBzero platform to host previously published materials, an extensive bibliography, tools, and datasets. Seven graduate students from Purdue University's College of Veterinary Medicine participated in the evaluation, performing tasks such as exploring the repository, submitting articles, adding bibliographic information, and using interaction features like user groups. Usability was assessed using the System Usability Scale (SUS), along with measures such as task success rates, navigational steps, task time, participant comments, and behavior notes. Results indicated satisfactory user experiences but identified usability issues related to metadata input (e.g., resource type and author information) and interface design (e.g., layout and consistency). The findings provide insights for improving HABRI Central and similar research repositories.

Comparing Repository Types: Challenges and Barriers for Subject-Based Repositories, Research Repositories, National Repository Systems, and Institutional Repositories in Serving Scholarly Communication

Armbruster and Romary [13] explored the challenges and barriers associated with four types of publication repositories—subject-based repositories, research repositories, national repository systems, and institutional repositories—in supporting digital scholarly communication. The study highlighted two major shifts in the role of repositories: the importance of a well-defined

and high-quality corpus for content and the need for services that provide high value to specific scholarly communities. Key challenges were identified in three dimensions: identification and deposit of content, access and use of services, and preservation of content and sustainability of service. The study also provided a comparative analysis of repository practices across major world regions, including Europe, North America, East Asia, and Australia. The findings emphasized that repositories are most effective when designed with user needs in mind and when they support the production of new knowledge through dedicated services.

The Process of Developing a Digital Repository for Online Teaching Using Design-Based Research

Richardson et al. [14] developed the Purdue Repository for Online Teaching and Learning (PoRTAL) as an Open Educational Resource (OER) to enhance online teaching skills and strategies for graduate students and faculty in higher education. Using a design-based research (DBR) approach, the study identified challenges faced by instructors new to online teaching through Van Tiem et al.'s Human Performance Technology (HPT) model. To address these needs, PoRTAL incorporated resources for online teaching, embedding research into practical activities, and resulted in the creation of an HPT-OER Model for Designing Digital Repositories. The findings emphasized the importance of integrating DBR methodologies to develop effective OER repositories that support instructional improvement and knowledge sharing.

CoRR: A Computing Research Repository

Halpern [15] describes the development of the Association for Computing Machinery's (ACM) "computing research repository" (CoRR), which integrated features from the Los Alamos e-print archive and Cornell University's Networked Computer Science Technical Reference Library.

CoRR is an open, permanent, online repository where submitted papers are not refereed, and anyone can browse and extract material for free. The study acknowledges challenges such as journal policies against online preprints, a cumbersome user interface, self-indexing submissions, lack of professional library staff management, and uncertain long-term funding. Despite these challenges, Halpern suggests that CoRR has the potential to revolutionize computer science publishing through its accessibility and open nature.

Organizing the Present, Looking to the Future: An Online Knowledge Repository to Facilitate Collaboration

Burchill et al. [16] describe the development of an online knowledge repository by the Manitoba Centre for Health Policy and Evaluation to manage complex linked databases and facilitate collaboration among researchers. The repository includes public and restricted-access pages on the World Wide Web, allowing users to access information through an indexed logical format or user-defined queries. Key components include a Concept Dictionary, which operationalizes health research concepts using administrative data; Research Definitions, which specify codes for medical procedures and diagnoses; a Meta-Index, which organizes concepts using the Medical Subject Headings (MeSH) system; and a Glossary for navigating research terms. The repository also supports educational resources, such as web-based graduate courses and materials for Manitoba's Regional Health Authorities. Confidential information is available on internal pages, ensuring secure access to sensitive data. The repository's design aims to enhance internal communication, facilitate collaboration, and improve access to epidemiological, institutional, and technical information.

Towards a Research Repository for Oxford University

Fraser [17] outlined the efforts of the Digital Archiving Group, a working group of Oxford's Information and Communication Technologies Committee (ICTC), to develop a pilot project for digital archiving of scholarly papers at Oxford University. A workshop held in June 2005 investigated the opportunities and challenges of establishing an open-access institutional repository for research. The resulting document summarized progress, recommendations, and issues discussed during the workshop. The repository aimed to address wider issues related to eprints and institutional repositories, including enhancing collaboration with organizations such as the Oxford Internet Institute, Research Technologies Service (OUCS), and Systems and Electronic Resources Service (OULS). The study emphasized the importance of creating a repository that supports open access while addressing challenges in policy, infrastructure, and long-term sustainability.

Lab2go — A Repository to Locate Educational Online Laboratories

Zutin, Auer, Maier, and Niederstätter [18] discuss the development of Lab2go, a repository designed to locate and describe online laboratories using semantic web technology. The repository leverages Web 3.0 technologies to enhance data handling through improved search mechanisms and facet-based browsing. Online laboratories, categorized into software simulations and real hardware experiments, are essential for technology-enhanced learning (TEL), particularly in engineering education. The study highlights the significance of semantic web technology in creating a general description model for online laboratories and improving their discoverability based on user-defined properties. This initiative addresses pedagogical

challenges in engineering education by supporting collaborative and interactive learning environments.

Research Data Repositories: Review of Current Features, Gap Analysis, and Recommendations for Minimum Requirements

Austin et al. [19] conducted a comprehensive review of 32 Canadian and international online data platforms to assess their features, identify gaps, and provide recommendations for minimum requirements in research data repositories. The study, led by Research Data Canada (RDC), focused on storage, data transfer, curation activities, preservation, access, and sharing features. A checklist was developed to compare platforms, revealing significant heterogeneity in features and services, non-standardized terminology, uneven compliance with standards, and a lack of certified repositories. Key recommendations included persistent dataset identification and citation, data reliability, version control, metadata standardization, privacy controls, long-term preservation, and repository certification. The study emphasized the need for investment in an integrated national digital infrastructure in Canada to support evolving researcher and end-user needs.

Functional Requirements for Research Data Repositories

Kim [20] examined the functional requirements for research data repositories, emphasizing their importance in supporting Open Science initiatives. The study analyzed features of research data, common repository platforms (e.g., Fedora Commons, iRODS, DataONE, Dataverse, Open Science Data Cloud, and Figshare), and community requests to design effective repositories. Using data curation profiles from Purdue University's Data Curation Center and requirements from the RDA Repository Interest Group and DataNest Community Platform, the study

identified 75 functional requirements grouped into 13 categories. These categories included metadata, identifiers, authentication, data access, policy support, publication, preservation, and user interface. The findings highlighted the need for repositories to ensure data accessibility, sharing, long-term preservation, and compliance with standards. The study provides a framework for organizations to design or deploy repositories that meet evolving researcher needs.

Related Studies

PromptSource: An Integrated Development Environment and Repository for Natural Language Prompts

Bach et al. [21] describe PromptSource, a system for creating, sharing, and using natural language prompts. It provides a templating language for defining data-linked prompts, an interface for iterating on prompt development, and community-driven guidelines for contributing prompts. PromptSource currently includes over 2,000 prompts for approximately 170 datasets, addressing challenges in NLP prompt engineering.

MIBiG 2.0: A Repository for Biosynthetic Gene Clusters of Known Function

Kautsar et al. [22] introduce *MIBiG 2.0*, an updated repository for biosynthetic gene clusters (BGCs) of known function. It enhances genome mining efforts for specialized metabolites by providing curated data on BGCs, supporting drug discovery and microbiome ecology studies. The repository includes 851 newly added BGCs, improved data curation, an updated schema for future compliance, and enhanced user features such as query searches and direct links to chemical structure databases.

ReCOVery: A Multimodal Repository for COVID-19 News Credibility Research

Zhou et al. [23] introduce *ReCOVery*, a repository designed to facilitate research on combating misinformation related to COVID-19. It collects and analyzes 2,029 news articles and 140,820 associated tweets to assess news credibility based on multiple factors, including textual, visual, temporal, and network-based information. The dataset enables researchers to study how COVID-19 misinformation spreads and provides benchmark statistics for evaluating news credibility prediction methods.

CNSA: A Data Repository for Archiving Omics Data

Guo et al. [24] introduce the *CNGB Sequence Archive (CNSA)*, a repository developed by China National GeneBank (CNGB) for archiving omics data. CNSA organizes sequencing data into six categories: Project, Sample, Experiment, Run, Assembly, and Variation, ensuring efficient data correlation and traceability. By adhering to life sciences data standards, CNSA aims to provide a comprehensive, curated, and accessible repository to support scientific research and the bio-industry.

RNAInter: An Enhanced RNA Interactome Repository

Lin et al. [25] present RNAInter, an updated RNA interactive repository that expands on RAID v2.0 with increased coverage and annotation. RNAInter features over 41 million RNA-associated interactions across 450,000 unique molecules, including RNA, protein, DNA, and compounds. Enhancements include eight times more interaction data, 94 additional species, refined annotations, and four embedded RNA interactome tools. This resource facilitates RNA research

by enabling a deeper understanding of RNA functions and molecular mechanisms.

JARVIS: An Integrated Repository for Data-Driven Materials Design

Choudhary et al. [26] introduce the Joint Automated Repository for Various Integrated Simulations (JARVIS), a computational infrastructure designed to accelerate materials discovery and design. JARVIS integrates density functional theory (DFT), classical force fields (FF), and machine learning (ML) techniques to facilitate data-driven materials research. The platform includes JARVIS-DFT, JARVIS-FF, JARVIS-ML, and JARVIS-tools, providing access to extensive datasets of materials, force fields, and machine learning models. By aligning with the Materials Genome Initiative (MGI), JARVIS enhances open-access databases and tools to optimize materials development and deployment.

ProteomeXchange: Advancing Big Data Approaches in Proteomics

Deutsch et al. [27] discuss the ProteomeXchange (PX) consortium, a global initiative that has standardized the submission and dissemination of mass spectrometry proteomics data since 2012. The consortium has expanded to six members, including PRIDE, PeptideAtlas, PASSEL, MassIVE, jPOST, iProX, and Panorama Public. By implementing updated submission guidelines and promoting open data policies, PX has facilitated the rapid growth of publicly available proteomics datasets. As of June 2019, over 14,100 datasets had been submitted, with a significant rise in data reuse and ‘big data’ applications, driving novel research in the field.

HGMD: A Comprehensive Database for Human Gene Mutations

Stenson et al. [28] present the Human Gene Mutation Database (HGMD®), a curated repository of germline mutations in nuclear genes associated with human inherited diseases. As of June

2020, HGMD contained over 289,000 gene lesions spanning more than 11,100 genes, extracted from 72,987 articles across 3,100 peer-reviewed journals. Designed for both research scientists and clinical diagnosticians, HGMD serves as a crucial resource for genetic analysis in clinical diagnostics and biomedical research, facilitating the understanding of genetic mutations and their implications in human health.

miRTarBase 2020: Advancing the MicroRNA–Target Interaction Database

Huang et al. [29] present miRTarBase 2020, an updated database of experimentally validated microRNA–target interactions (MTIs). With over 13,404 validated MTIs curated from 11,021 articles, miRTarBase integrates a text-mining system to enhance the recognition of MTI-related literature. The update also incorporates multiple biological databases to provide regulatory network insights, including upstream and downstream miRNA interactions. A significant expansion of high-throughput experimentally validated MTIs, particularly through CLIP-seq technology, further solidifies miRTarBase as a leading resource in microRNA research.

CellPhoneDB: A Database for Inferring Cell–Cell Communication

Efremova et al. [30] introduce CellPhoneDB, a database designed to analyze cell–cell communication by considering the subunit architecture of ligand–receptor complexes. Unlike other repositories, CellPhoneDB accurately represents heteromeric complexes and integrates a statistical framework to predict enriched cellular interactions from single-cell transcriptomics data. The updated version, CellPhoneDB v.2.0, includes additional functionalities for introducing new interacting molecules and optimizing computational efficiency, making it a valuable resource for researchers analyzing cell communication networks.

AI-Based FDA-Approved Medical Devices and Algorithms: A Database Overview

Benjamens et al. [31] analyze the current landscape of artificial intelligence (AI) and machine learning (ML)-based medical devices and algorithms approved by the U.S. Food and Drug Administration (FDA). Their study highlights the regulatory challenges and the importance of transparency in AI/ML-based medical technology approvals. They identified 64 FDA-approved AI/ML-based medical devices, of which only 45% explicitly mentioned AI/ML in official FDA announcements. The study provides an open-access, continuously updated database for tracking these technologies.

Development of a Web-Based Records Management System for the Office of Senior Citizen Affairs

Libadia *et al.*[32] explored the development of a web-based Electronic Records Management System (ERMS) tailored for the Office of Senior Citizen Affairs (OSCA) in the Philippines. The study aimed to address inefficiencies in manual record-keeping by centralizing data storage and streamlining workflows. The system incorporated web and data-capturing technologies to enhance record management, enabling faster and more accurate information retrieval. The methodology involved requirements analysis, system design using MySQL and Laravel, and evaluation based on the ISO 25010 product quality model. Findings highlighted the system's usability, reliability, and efficiency in managing member records while ensuring data security and compliance with government regulations.

Developing a Record Archiving System for Eastern Visayas State University – Burauen Campus

Villarosa [33] developed a record archiving system for Eastern Visayas State University – Burauen Campus to improve records management efficiency. The system aimed to provide a paperless solution, ensuring easy data retrieval, reliable database backup, and streamlined documentation processes. The research followed a system developmental approach, consisting of an analysis phase to assess user needs and a design and development phase for system creation. Findings indicated that the system could enhance institutional records management and overall university operations.

Capstone Archiving Management System with Citation and Reference Generator

Doromal and Soberano [34] developed the "Capstone Archiving Management System with Citation and Reference Generator" to improve archiving practices at CHMSU-FT. The system replaces traditional hardbound archiving with digital storage, reducing paper usage and optimizing physical space. Key features include efficient digital archiving, automated citation and reference generation, and enhanced accessibility for students, professors, and librarians. The study employed a descriptive-developmental methodology, with user acceptance testing conducted among 20 respondents. Results demonstrated high usefulness (mean score = 4.71), ease of use (mean score = 4.66), and system usability (mean score = 4.66), confirming its effectiveness in modernizing academic archiving.

Preserving Scholarly Endeavors: A System for Archiving Thesis and Dissertations

Mondero et al.[35]developed the Thesis and Dissertations Archiving System to enhance the storage and organization of research papers at the University of Industrial Technology.The system was designed with features such as a user-friendly interface, efficient search functionality, seamless tracking of thesis and dissertation projects, and streamlined book borrowing and returning processes. Using descriptive and developmental research methodologies, data collection was conducted through standardized instruments, while the Rapid Application Development (RAD) Model guided the software's development. The system was evaluated using the Post-Study System Usability Questionnaire (PSSUQ), which yielded a mean score of 2.28, indicating strong usability and acceptance among respondents. The findings demonstrated the system's effectiveness in enhancing accessibility, efficiency, and information retrieval capabilities, making it an indispensable tool for faculty, students, and external researchers. The study recommended institution-wide implementation to optimize archiving processes and support academic excellence.

Web-Based System for Archiving Activities of University College Staff

Taha and Abdulqader [36] proposed a web-based system to collect and archive the activities of employees and staff at an Iraqi university. The system, developed using Node.js for both frontend and backend frameworks, incorporates features such as role-based access for administrators, deans, department heads, and employees. The primary goal of the system is to replace manual archiving processes with a digital solution that ensures efficient documentation, performance evaluation, and compliance with institutional policies. By enabling global access through internet connectivity, the system facilitates quick and reliable retrieval of archived data. The study highlights the importance of such systems in improving organizational efficiency, reducing administrative workloads, and supporting decision-making processes.

Systems Development for Records Archiving and Digital Documents Repository

Gamba et al.[37] developed a records archiving and document repository system to address data access and transfer challenges at Sorsogon State College, utilizing a secure web-based platform with encryption protocols. The college, with four campuses in different municipalities, faced challenges in communication and real-time updates due to geographical distances. The developed system features a General User Interface (GUI) accessible on web-enabled computers and mobile devices, requiring only a web browser for rendering. The web-based archiving and repository system connects the satellite campuses to the college's private server through a cost-effective virtual private network. Security is enhanced through user credentials, a 1024-bit Rivest-Shamir-Adleman (RSA) private/public key exchange, and 256-bit Advanced Encryption System (AES) encryption. Contents of uploaded files are encrypted at 128-bit to prevent

unauthorized access. The developed system addresses the need for readily accessible and secure data transfer among the college's campuses.

A Web-Based Document Archiving System Using Indexing and Machine Learning for Research and Innovation Grant Allocation

Lupyani [38] developed a web-based document archiving system that integrates machine learning to enhance the research and innovation grant allocation process, ensuring efficiency, transparency, and fairness.. The study addressed the challenges of traditional manual grant selection, which can be time-consuming, resource-intensive, and prone to bias. Using historical grant data, the system employed machine learning algorithms to ensure fairness, efficiency, and transparency in funding decisions. A comparative analysis of three models—K-Nearest Neighbour (KNN), Naive Bayes, and Support Vector Machine (SVM)—revealed that SVM outperformed the others with an accuracy of 88%, precision of 86%, recall of 87%, and an F1 score of 87%. Consequently, SVM was integrated into a web-based application to categorize proposals into topic areas such as engineering, science, and technology. The system also facilitates document archiving, tracking project progress, and monitoring funded research initiatives. This innovative approach aims to streamline grant allocation processes while promoting equity and supporting research advancements.

Archiving and Document Management at Taibah University

Montaño et al. [39] developed an electronic document archiving system known as eDoc for Taibah University, aiming to improve document storage, accessibility, and security. The system integrates data management and archiving features to increase efficiency and reduce costs. eDoc offers a web-based interface for document access, ensures long-term preservation, and includes

backup and restoration capabilities. The study employed the Rapid Application Development methodology to deliver a reliable and efficient archiving solution.

Bigtable: A Distributed Storage System for Structured Data

Chang et al. [40] introduced Bigtable, a distributed storage system for managing structured data, designed to scale to petabytes across thousands of commodity servers. Bigtable has been used in several Google projects, such as web indexing, Google Earth, and Google Finance, meeting varied demands in terms of data size and latency. Despite these challenges, Bigtable provides a flexible and high-performance solution. The article describes Bigtable's simple data model, which allows clients dynamic control over data layout and format, and details the system's design and implementation.