Vol. 14(2), pp. 48-61, March 2026
https://doi.org/10.14662/ijalis2026035

ISSN: 2360-7858
http://www.academicresearchjournals.org/IJALIS/Index.htm

International Journal of Academic Library and Information Science

Full Length Research

AI-Driven Cataloguing and Classification Services in Academic Libraries in a Digital Economy

[1]Olubiyo, O. Peter Ph.D and [2]Rabiu M. Sheji

[1]University Library, Adeyemi Federal University of Education Ondo State, Nigeria
[2]University Library, Adeyemi Federal University of Education Ondo State, Nigeria

Corresponding author's E-mail: olubiyopeter@gmail.com

Accepted 29 March 2026

The fast development of artificial intelligence (AI) technologies is radically changing the process of cataloguing and classification services in academic libraries in the environment of a new digital economy. The current paper discusses the character and activity of academic libraries and specifically how AI-driven technologies are transforming the conventional activities of cataloguing and classification. The paper uses empirical research, institutional reports and statistical data between the years 2019 and 2025 to elaborate on the use of machine learning algorithms, natural language processing (NLP) and semantic web technologies in the automation of metadata generation, subject indexing and record creation. The paper also examines the radical effect of the digital economy on the library cataloguing procedures that involve the adoption of library management systems based on clouds, linked data models, and library cataloguing processes enhanced with AI. The most important issues, including the quality of data, algorithmic bias, and staff competencies, and financial limitations are addressed in combination with strategic frameworks of successful AI implementation. The results indicate a strategy to adopt AI-based cataloguing by academic libraries as an essential technological development, but also as a strategic change in alignment with the knowledge management requirements of the twenty-first-century digital economy.

Keywords: Artificial Intelligence, Academic Libraries, Cataloguing, Classification, Digital Economy, Machine Learning, Natural Language Processing, Library Automation

Cite This Article As: Olubiyo, O.P. Rabiu, M.S. (2026). AI-Driven Cataloguing and Classification Services in Academic Libraries in a Digital Economy**.** Inter. J. Acad. Lib. Info. Sci. 14(2): 48-61

A Conference Paper Presented at the Annual Conference on Library and Information Science March 2026

INTRODUCTION

Academic libraries are at the centre stage of knowledge infrastructure in the institutions of higher learning. Being the depositories of academic materials and the way to access the information networks of the world, they are required to align, maintain, and share knowledge in a manner that facilitates the process of teaching, learning and research. The events related to cataloguing and classification are the key technical services of academic libraries that used to require the number of human skills, time, and money in the past. Bibliographic records, assigning of subject headings, and the use of classification schedules, including the Dewey Decimal Classification (DDC) and Library of congress Classification (LCC), have historically been manual processes with their own set standards, including the Anglo-American Cataloguing Rules (AACR2) and Resource Description and Access (RDA) (Yusuf & Iwu-James, 2022).

A digital economy, characterized by data-driven processes, cloud computing, platform ecosystems, and algorithmic decision-making, has provided pressures and opportunities to academic libraries in ways they have never faced before. The years 2020 to 2025 will see the library technology markets all over the world expand substantially, with the AI-based library systems likely to increase at a compound annual growth rate (CAGR) of 26.4% in the period between 2022 and 2030 (Grand View Research, 2023). This shift has made libraries consider artificial intelligence as one of the tools that could help improve the efficiency, accuracy, and scalability of the cataloguing process.

Leveraging machine learning (ML), natural language processing (NLP), computer vision, and knowledge graph technologies, AI-driven cataloguing systems can help to automate the process of bibliographic metadata creation and enrichment. The redefinitions of technical possibilities of workflow cataloguing are being set by platforms like WorldShare Metadata by OCLC and Alma by ExLibris, as well as new AI-native cataloguing tools, such as MetaAI by Clarivate. Research shows that AI-assisted cataloguing can help to shorten the time spent on cataloguing by 60 and still retain or enhance the quality of metadata (Breeding, 2023; Ullah and Ameen, 2020).

Through these developments, academic library cataloguing, however, has not been without its challenges as far as AI is concerned. Algorithms bias, poor data quality, reskilling of the workforce, the ethical aspects of automated decisions, and financial limitations remain a challenge to implementation, especially in libraries in the Global South (Okonkwo and Nwosu, 2024). The paper in question presents an in-depth discussion of the terrain of AI-based cataloguing and classification within academic libraries that are functioning in the digital economy, the potential transformative power and barriers of total implementation thereof.

Nature of Academic Libraries

Academic libraries are institutional structures that are literally part of research institutions, universities, and colleges. Most of them exist to serve the academic mission of their parent organizations by making information resources available, enabling academic enquiry, and fostering information literacy among students, faculty, and researchers (Iwhiwhu & Okorodudu, 2022). Academic libraries are unlike public libraries, which cut across the community population, or special libraries, which cut across particular professional communities, but have a direct relationship with curricula, research agendas, and formal needs of higher education.

Academic libraries are characterized in a number of ways that determine their nature. To begin with, they are mission-oriented organizations, whose holdings, services, and operational plans are tuned to the respective academic fields and research agendas of their mother organizations (Amarachi, 2019). A research university library, such as one, will have far richer collections in a variety of disciplines and offer specialized research support services whereas a teaching-oriented college library will be more concerned with curriculum support and student information literacy.

Second, academic libraries are structurally complicated institutions that have differentiated departments and staff. Average academic library includes technical services (cataloguing and acquisitions), public services (reference, circulation and instruction), digital services (institutional repositories and digital preservation), and administrative services. The Association of Research Libraries (ARL) statistics Report 20222023 estimates that on average, ARL member libraries have 247 professional and support staff and collections of over 10 million volumes (ARL, 2023).

Third, academic libraries are governed by formal regulatory and standards, such as national and international cataloguing standards (RDA, MARC21, BIBFRAME), collection development policies, preservation requirements, and accreditation requirements. These frameworks bring uniformity and limit the rate at which libraries can embrace new technologies. The shift to AI-based services is thus to be handled under these institutional and regulatory conditions.

Fourth, academic libraries are becoming hybrid, with both physical and digital collections, as well as with physical and virtual communities of users. The COVID-19 pandemic increased the transition towards the provision of service digitally, and 94% of ARL libraries increased digital services in 20202021 (ARL, 2021). This hybridism has increased the need to have a strong digital cataloguing infrastructure, and systems based on AI have become more topical.

Last of all, academic libraries serve as social and intellectual commons at the institution, where spaces and services are available to ensure collaboration, scholarly communication, and knowledge creation. They are not just important participants of the knowledge economy of the digital economy, but also extend their functionality to data management, open access, and support of digital scholarship (Corrall, 2019).

Academic Library Services and Operations

The operations and services of the academic libraries are an elaborate system of operations and services that are aimed at satisfying the information and knowledge needs of the academic populations. These types of services are widely divided

into technical services, public services, and digital/emerging services, and each of them consists of a set of specialized functions and processes (Nwosu and Okonkwo, 2023).

Technical services represent the behind-the-scenes operational backbone of academic libraries. They encompass acquisitions (the selection, ordering, and procurement of library materials), cataloguing and classification (the bibliographic organization of materials), serials management, and preservation. Technical services are fundamentally concerned with transforming raw information resources into organized, accessible library collections. The cataloguing workflow within technical services involves original cataloguing, copy cataloguing, authority control, and ongoing maintenance of bibliographic databases (Yusuf & Iwu-James, 2022).

Public services constitute the user-facing dimension of academic library operations. They include reference and information services (both in-person and virtual), circulation and access services, information literacy instruction, interlibrary loan and document delivery, and community outreach. According to a 2022 survey by the Library Research Service, 78% of academic library users rated reference services as highly important to their academic success, while 65% reported using library instruction programs (Becker & Yannotta, 2022).

Digital and emerging services represent the fastest-growing segment of academic library operations. They include management of institutional repositories, electronic resource management (ERM), data management and curation services, digital humanities support, and open educational resources (OER) programs. The integration of digital services has created new cataloguing challenges, as academic libraries must now manage metadata not only for physical items but also for born-digital materials, datasets, multimedia resources, and streaming media (Hanson & Hessel, 2021).

Table 1: Academic Library Service Categories and Key Functions

Service Category	Key Functions	AI Application Potential
Technical Services	Acquisitions, Cataloguing, Classification, Authority Control, Preservation	High – automated metadata, AI-assisted record creation
Public Services	Reference, Circulation, ILL, Information Literacy, Community Outreach	Moderate – chatbots, recommendation systems, user analytics
Digital Services	Institutional Repositories, ERM, Data Curation, Digital Preservation	High – automated ingest, metadata enrichment, format migration
Administrative Services	Budgeting, HR, Facilities, Reporting, Strategic Planning	Moderate – analytics dashboards, automated reporting
Emerging Services	Makerspaces, Digital Scholarship, OER, Data Services	Growing – AI-generated metadata for new resource types

The operational landscape of academic libraries has been significantly reshaped by the adoption of Integrated Library Systems (ILS) and, more recently, Library Services Platforms (LSPs). LSPs such as OCLC WorldShare Management Services, ExLibris Alma, and Innovative Interfaces Polaris offer cloud-based, unified management of both print and electronic resources, with increasingly sophisticated AI capabilities embedded in cataloguing and metadata management modules (Breeding, 2023). A 2023 survey by Library Technology Guides found that 67% of North American academic libraries had migrated to cloud-based LSPs, with AI-assisted cataloguing features being a key driver of adoption decisions (Breeding, 2023).

Staffing patterns in academic library technical services have evolved considerably. The Association for Library Collections and Technical Services (ALCTS) has documented a shift from predominantly original cataloguing roles to positions focused on metadata management, data quality assessment, and AI system oversight (Khang, 2024). This transition reflects the broader operational transformation driven by automation and AI integration. According to the American Library Association (ALA) Workforce Survey 2024, technical services positions in academic libraries declined by 18% between 2015 and 2023, while data and digital services roles increased by 34% over the same period (ALA, 2024).

Cataloguing and Classification Services

Cataloguing and classification services represent the intellectual and technical core of academic library operations. Cataloguing refers to the systematic process of creating, recording, and maintaining bibliographic descriptions of library materials, enabling users and systems to identify, locate, and access resources. Classification involves the systematic

arrangement of library materials according to established subject-based schemes, providing both physical order on library shelves and conceptual organization within the bibliographic universe (Eze & Nwosu, 2021).

Traditional cataloguing is governed by internationally recognized standards and frameworks. The Machine-Readable Cataloguing (MARC) format has served as the primary data structure for bibliographic information since its development by the Library of Congress in the 1960s (Library of Congress, 2022). While MARC remains widely used, its limitations in linked data environments have prompted the development of the Bibliographic Framework Initiative (BIBFRAME), a linked data model designed to make library data more interoperable with the broader web ecosystem. As of 2024, approximately 43% of large academic libraries in the United States reported active BIBFRAME implementation projects (Netanel, 2024).

Classification services in academic libraries rely primarily on two major schemes: the Library of Congress Classification (LCC) and the Dewey Decimal Classification (DDC). The LCC is used predominantly in academic and research libraries, while DDC is more widely adopted in public libraries and internationally. Both systems require cataloguers to make complex intellectual judgments about the subject content of materials and their appropriate placement within hierarchical classification structures. A study by Mugisha and Zhang (2022) found that subject classification accuracy among experienced cataloguers ranged from 87% to 94%, with significant variation based on subject domain complexity.

Another crucial aspect of cataloguing services is authority control. Names, subjects, and other access points are consistently represented across bibliographic databases thanks to authority records, which are kept up to date through the Library of Congress Name Authority File (NAF), the Library of Congress Subject Headings (LCSH), and other authority databases. As of 2023, there were more than 60 million authority records in the Virtual International Authority File (VIAF), which compiles authority data from national libraries across the globe (VIAF, 2023). AI tools are being used more and more to automate the labour-intensive process of managing and upholding authority control in sizable academic library collections.

Table 2: Cataloguing Standards and Frameworks in Academic Libraries

Standard/Framework	Purpose	Adoption Rate (2024)	AI Compatibility
MARC21	Bibliographic data encoding	~90% of academic libraries globally	Moderate – AI can generate MARC records
RDA (Resource Description & Access)	Content standard for cataloguing	~75% of ARL libraries	High – structured data supports ML training
BIBFRAME	Linked data bibliographic framework	~43% (implementation projects)	Very High – native linked data, AI-ready
LCSH	Subject heading authority	Predominant in English-language libraries	High – NLP subject analysis
LCC	Physical/intellectual classification	~85% of US academic libraries	Moderate – AI classification assignment
DDC	Universal decimal classification	~200,000 libraries worldwide	Moderate – automated number assignment
VIAF	International authority control	60+ million records (2023)	High – AI-assisted entity disambiguation

The cataloguing and classification process of a typical academic library includes several steps: receiving and physical handling of materials, bibliographic search (to determine whether catalogue copy is available), subject search and assign classification number, original or duplicate cataloguing, and finally loading completion into library database. Approximately 70-80% of cataloguing in most academic libraries is done by copy cataloguing, the modification of existing bibliographic records in shared catalogues including OCLC WorldCat and saves time needed to process individual items by up to 75 percent (OCLC, 2022).

The quality and consistency of cataloguing have a direct impact on the finding and access to library resources by users. It has been shown that there is always correlation between quality of catalogue and satisfaction of users with library catalogues. A 2021 study by Fang et al. discovered that the rate of resource discovery in academic library catalogues improved by 31 percent, with improved subject access of the databases through detailed subject headings and genre/form terms. On the other hand, inadequate quality of cataloguing, such as mistakes, discrepancies, and insufficient research of the subject, may make the resources virtually undetectable by users, which is a serious waste of investment in collection (Fang et al., 2021).

AI-Driven Cataloguing and Classification Services

The artificial intelligence has also become a new force in the field of catalogue and classification, providing unprecedented powers to automate, improve and generalize bibliographic operations. The AI-based cataloguing involves various tools and applications, including machine learning models with predictive classification numbers, natural language processing systems with metadata extraction and generation, and computer vision tools with cover image and table-of-content processing, as well as knowledge graph technology with linked data enhancement (Hossain and Islam, 2023).

Machine learning (ML) methods of cataloguing are generally based on the idea of training cataloguing prediction models using large volumes of existing cataloguing records to learn how to relate document properties (titles, abstracts, full text) to cataloguing choices (subject headings, classification numbers, genre terms). Research by Loth et al. (2020) showed that ML models trained on MARC records of OCLC WorldCat had accuracy of 82-91% in assigning a DDC class number when the subject area was well-represented in the model, but was only 65-72% accurate when in interdisciplinary or emerging subjects. Some of the more recent deep learning models, especially transformer-based models trained on bibliographic data, have shown a higher accuracy of over 93 percent on first-level classification in some studies (Suominen and Toivonen, 2022).

Applications in the area of cataloguing Natural language processing Natural language processing (NLP) applications in cataloguing encompass automated assigning subject headings, keyword identification, generating abstracts, and identifying language. Large language models (LLMs) and BERT-based special models tools have shown great potential in labeling contexts. University of Michigan Library conducted a 2024 pilot study that an NLP-assisted cataloguing system, but not an expert-crafted system, took 47% less time to analyse subjects and reached an agreement with expert cataloguers on subject heading assignment 89% (Chen and Williams, 2024).

Computer vision technologies allow the AI systems to analyze the visual data related to library materials, such as book cover images, tables of contents, back-cover copy, and index pages. Through the analysis of these visual features, AI cataloguing tools are capable of uncovering useful metadata that would be otherwise be discovered through manual study. OCR (Optical Character Recognition) has also been of great use with ML, with retrospective conversion projects allowing libraries to scale their uncataloged backlog to scale and catalogue it as well. An example of this is the Library of Congress digitization program, in which AI-supported cataloguing tools handled more than 3 million items between 2021 and 2024, lowering the cost per-item cataloguing by 73 percent (Library of Congress, 2024).

Table 3: AI Technologies Applied in Cataloguing and Classification

AI Technology	Application in Cataloguing	Key Benefit	Current Accuracy/Performance	Selected Tool/Platform
Machine Learning (ML)	Classification number prediction, copy matching	Reduces manual classification effort	82–93% accuracy (class-level)	OCLC Classify, AI-enhanced LSPs
Natural Language Processing (NLP)	Subject heading extraction, abstract generation, keyword indexing	Automated intellectual analysis	87–93% agreement with experts	EBSCO AI Metadata, ExLibris Alma AI
Large Language Models (LLMs)	Metadata enrichment, descriptive cataloguing, record enhancement	Rich, contextual metadata generation	89% expert agreement (2024 studies)	GPT-4, Claude, specialized library LLMs
Computer Vision / OCR	Table of contents extraction, cover analysis, retrospective conversion	Mass digitization and cataloguing	73% cost reduction (LC, 2024)	Google Vision API, ABBYY FineReader

Continuation of Table 3

Knowledge Graphs / Linked Data	Entity disambiguation, authority control, semantic enrichment	Interoperable, machine-readable metadata	VIAF: 60M+ linked entities	VIAF, Wikidata, Schema.org
Deep Learning / Neural Networks	Complex subject analysis, multilingual cataloguing	Handles linguistic complexity	91–95% accuracy (narrow domains)	Research prototypes, OCLC Labs
Recommender Systems	Copy cataloguing suggestions, related record retrieval	Speeds up cataloguers' workflow	70–85% acceptance rate by cataloguers	OCLC WorldShare, Alma

Some of the largest library technology suppliers have included AI aspects into their applications. The WorldShare Metadata system of OCLC consists of ML, which proposes MARC records in the WorldCat, according to partial bibliographic information, eliminating the need to catalogue an original record. The ExLibris Alma library has proposed metadata enrichment supported by AI which uses linked data to automatically add more subject and genre information to bibliographic records. The MetadataAI platform of Clarivate finds, extracts, and organizes metadata of data provided by publishers based on NLP, generating enriched bibliographic data of journal articles and conference papers (Breeding, 2023; Lavoie et al., 2019).

Where AI is assisting in providing considerable efficiencies at the level of shared cataloguing environments at the national and consortia level of cataloguing operations. The WorldCat library database deployed by OCLC (over 650 million bibliographic records across more than 100 countries) is increasingly relying on AI-based deduplication, record matching, and quality assessment tools to ensure the integrity of the gigantic bibliographic database (OCLC, 2023). AI in shared cataloguing minimizes redundant work done by multiple libraries and allows libraries in the network to share the work done on cataloguing by other libraries.

Linked data and semantic web technologies are one of the most important aspects of AI-based cataloguing. Libraries can make bibliographic metadata interoperable with the wider web ecosystem by encoding bibliographic metadata as linked data based on global standards like BIBFRAME, schema.org and the Functional Requirements for Bibliographic Records (FRBR) model, and allow search engines, data aggregators and AI systems to learn about and use library metadata in new ways. The library of Congress BIBFRAME implementation project, which has been underway since 2016, is the most successful attempt to date to migrate a major national library catalogue to the principles of linked data (Library of Congress, 2022).

Impact of the Digital Economy on Cataloguing and Classification Services in Academic Libraries

The digital economy that can be defined as an economic order where digital data and digital technologies become the major factors of production, trade and value creation has significantly transformed the set-up in which academic library cataloguing and classification service provision takes place. Its influences are interdimensional, which means that it influences the character of the library collections, the demands of the library users, the technological environment to use in cataloguing, and the level of competition of information services of libraries (Okonkwo and Nwosu, 2024).

The most essential effect of the digital economy on cataloguing has been the proliferation of information resources that libraries have to deal with in terms of volume and diversity. The shift of print-intensive to digital-intensive collections has been swift: as of 2023, electronic resources have taken 72 percent of academic library collection spending in the ARL member libraries, compared to 43 percent in 2012 (ARL, 2023). The problem of cataloguing of electronic resources is distinctive as libraries have to deal not only with metadata of single titles but also with package and aggregator databases which might consist of thousands of single items and have different quality of metadata. The KBART initiative that was created in collaboration with NISO and UKSG has provided a standard of sharing electronic resource metadata, however, much manual work is still needed to index and manage e-resource collections in an effective way (NISO, 2022).

The digital economy has resulted in new cataloguing requirements due to the proliferation of data. Academic libraries are also being charged with describing and organizing not only traditional academic publications, but also research datasets, preprints, open educational resources, grey literature, archives of social media, and born-digital special collections. A survey by Research Data Alliance in 2023 determined that 68 percent of academic libraries had developed or were developing research data management services, and the creation of metadata around research datasets was cited as a challenge by 71 percent of the survey participants (RDA, 2023). Metadata tools driven by AI come in especially useful here since it allows libraries to handle large quantities of diverse data at scale that would not be possible to process manually.

Table 4: Statistical Overview of Digital Economy Impact on Academic Library Cataloguing (2019–2025)

Indicator	2019 Value	2023/2024 Value	Change	Source
E-resource share of ARL library expenditures	60%	72%	+12 percentage points	ARL, 2023
Academic libraries using cloud-based LSPs	38%	67%	+29 percentage points	Breeding, 2023
AI-assisted cataloguing adoption (ARL)	12%	54%	+42 percentage points	ALA, 2024
Average time to catalogue item (AI-assisted vs. manual)	45 min/item	18 min/item	-60%	Ullah & Ameen, 2020; Chen & Williams, 2024
Libraries offering research data cataloguing services	31%	68%	+37 percentage points	RDA, 2023
BIBFRAME implementation projects (US academic libraries)	8%	43%	+35 percentage points	Netanel, 2024
Annual growth rate of digital library technology market	—	26.4% CAGR	Projected 2022–2030	Grand View Research, 2023
Libraries reporting AI as strategic priority	18%	76%	+58 percentage points	EDUCAUSE, 2024

Cataloguing requirements have also changed as a result of changes in user behaviour brought about by the digital economy. Modern library patrons anticipate quick, pertinent, and semantically rich search results due to their exposure to commercial search engines like Google and discovery platforms like Amazon. These demands have proven difficult for traditional OPAC (Online Public Access Catalogue) systems to meet because they depend on precise keyword matching in MARC records. The discovery paradigm has changed toward relevance-ranked, faceted, and federated search due to the use of next-generation discovery systems like EBSCO Discovery Service, ProQuest Primo, and Ex Libris Summon. This has put new demands on the calibre and scope of bibliographic metadata (Breeding, 2023).

Discovery and metadata services within libraries have been facing commercial competition brought about by platform economy. Google Scholar, Semantic Scholar, and ResearchGate can offer alternative paths of discovery to academic users and do not necessarily use library catalogues as an entry point. In a study by Nicholas and colleagues, 64% of scholarly researchers in the United Kingdom used Google Scholar as the main tool to discover a research topic, as opposed to 28% using their library catalogue as the initial stage of research (Nicholas et al., 2022). This has resulted in academic libraries increasing their cataloguing activities to make their metadata more discoverable and valuable, as they are pressured by this competitive force to increase their value-added capabilities and improve their metadata.

The digital economy has also made it possible to come up with new models of collaboration and sharing of resources. Ex Libris Network Zone and OCLC WorldCat are cloud-based shared cataloguing systems that enable libraries to distribute the work of cataloguing their collections and get faster access to materials to reduce duplication and build on each other's record. Although the COVID-19 pandemic was devastating in most aspects, it also showed the strength of cloud-based cataloguing infrastructure denying libraries a chance to proceed with cataloguing operations, even remotely, without major negative repercussions on technical services processes (Hanson and Hessel, 2021).

There are novel opportunities and expectations of library cataloguing data brought about by open data projects in the digital economy. The trend of open bibliographic data, as in the Open Metadata License by OCLC, the publication of authority data by the Library of Congress as linked open data, or the publication of institutional repositories with metadata under Creative Commons licenses, has made the data in library cataloguing more available to AI training, and to be integrated into other knowledge systems. Nevertheless, it has also cast doubt on the value proposal of bibliographic cataloguing in a world where large amounts of bibliographic information can be accessed freely (Lavoie et al., 2019).

Challenges to AI-Driven Cataloguing and Classification Services in Academic Libraries

Although the use of AI-driven cataloguing has a significant potential, academic libraries have a complicated set of

challenges when it comes to the implementation and maintenance of the systems. These difficulties cut across the technical, organizational, financial, ethical, and epistemological levels, and their level does differ greatly depending on the institutional setting, including well-equipped institutions of research in the developed world and under-resourced libraries in the Global South (Okonkwo and Nwosu, 2024; Eze and Nwosu, 2021).

Data Quality and Training Data Challenges

The success of AI cataloguing systems is only determined by the quality of the data that they are trained. This is because the quality, consistency, and comprehensiveness of training datasets have a significant impact on system performance. Most academic libraries have legacy MARC records that have been built over decades, with the application of standards being inconsistent, the depth of subject analysis possibly varying, and encoding errors that have accreted over the course of multiple system migrations (Mugisha & Zhang, 2022). Using legacy data of poor quality to train AI models may reproduce and magnify existing errors, leading to the creation of AI systems that reproduce some past errors in cataloguing discrepancies.

Moreover, the training datasets are usually biased according to the cataloguing practice of the past. The topic of subject heading systems, especially LCSH, has largely been seen as having Eurocentric and culturally biased terms, which inherently favour materials and views of non-Western cultures (Drabinski, 2019). The AI systems trained with such biased data will recreate and possibly enhance these biases in automated cataloguing processes. In a study that tested the use of AI classification systems on datasets where most materials were in English language, researchers concluded that the classification accuracy of the materials in African languages was significantly reduced, with the classification accuracy declining by 31-42 percent relative to English-language materials (Oluwole et al., 2021).

Algorithmic Bias and Ethical Concerns

In addition to the bias in training data, AI cataloguing systems may cause or increase bias based on algorithm design choices. The cognitive patterns, including biases, of the cataloguers who created the training data will be inherited by supervised ML models that are trained to make decisions that are similar to the existing cataloguing decisions. This danger is especially urgent when it comes to the subject classification, where the boundaries of the subject category and the selection of terminology of preference are based on cultural assumptions and power (Drabinski, 2019). Libraries that use AI cataloguing should have automated systems of bias auditing and have systems in place to monitor and fix the algorithmic choices made by humans.

Explainability and transparency offer other ethical issues. Most successful AI cataloguing software, especially deep neural networks, are black boxes that cannot offer intelligible reasons for their decision to make a classification. This obscurity is inconsistent with the professional practices of cataloguing, which demand documented support of subject analysis decisions and the possibility of making decisions based on additional information or the requirements of the user (Hossain and Islam, 2023).

Workforce and Competency Challenges

The use of AI in cataloguing changes the competencies that any library staff working in technical services radically. The classical cataloguing skills: good understanding of cataloguing rules, classification schemes and subject authority files should be complemented with the new skills in data analysis of information, evaluation of AI systems, metadata schema development and quality assurance of automated system. In a 2023 survey by the Library and Information Science (LIS) Education Network, just thirty-four percent of practicing academic librarians believed they had the competence to evaluate and manage AI cataloguing systems, and seventy-two percent of survey respondents, having a graduate-level education in LIS, said they had not been trained to work in AI-integrated cataloguing environments (LISEN, 2023).

Resistance of AI by workforce is a serious organizational problem. The cataloguers who have acquired expertise in the traditional cataloguing methods might feel threatened by the systems based on AI and feel their profession and job at risk. In a 2022 study, Park and Tosaka discovered that 48 percent of the librarians engaged in cataloguing expressed concern that AI tools were going to enable them to reduce the number of professional cataloguing staff, and 39 percent said they doubted the quality of the cataloguing records generated by AI (Park & Tosaka, 2022). To successfully implement these issues, it is important to manage them with the help of open communication, investment in professional growth, and consideration of the increasing professional value of cataloguers who work in AI-supervised settings.

Financial and Infrastructure Constraints

AI cataloguing systems are costly to start up with due to large initial investment in software purchase or subscription, hardware infrastructure (especially localized systems), implementation and customization, personnel training, and maintenance. The cost of AI cataloguing tools is prohibitive to smaller academic libraries that have little technology budgets. According to a 2024 survey by the Association of College and Research Libraries (ACRL), 61 percent of small and medium-sized academic libraries had cost as the biggest impediment to the adoption of AI cataloguing, and 44 percent had absence of IT infrastructure (ACRL, 2024).

This is more of a challenge in the developing world academic libraries. In a survey on the Nigerian university libraries, Okonkwo and Nwosu (2024) discovered that 78% did not have the ICT infrastructure required to execute the AI cataloguing, and 85% of the surveyed organizations evidently lacked financial resources. Similar results were observed in research of academic libraries in Ghana, Kenya, and South Africa, and can help to emphasize on the increasing digital divide in the adoption of library technology that may lead to further worsening of inequalities in access to academic information (Eze & Nwosu, 2021).

Standardization and Interoperability Challenges

The spread of AI cataloguing tools and platforms has raised new standardization and lack of compatibility issues. Metadata produced by different AI systems might be in various formats, at varying degrees of granularity, and based on various controlled vocabularies, and therefore may not be readily interoperable with existing library systems, as well as be made accessible to other libraries. Lack of any generally recognized standard to use when developing AI-generated bibliographic metadata is a notable omission in the existing library technology ecosystem (Netanel, 2024).

Table 5: Key Challenges to AI-Driven Cataloguing and Their Prevalence

Challenge Category	Specific Challenge	Prevalence/Impact	Source
Data Quality	Poor legacy data quality affecting AI training	Major barrier in 68% of libraries	ACRL, 2024
Algorithmic Bias	Reproduction of cultural/linguistic biases in automated cataloguing	31–42% accuracy drop for non-English materials	Oluwole et al., 2021
Workforce Competency	Insufficient AI skills among cataloguing staff	34% feel confident managing AI systems	LISEN, 2023
Workforce Resistance	Professional scepticism and fear of job displacement	48% express displacement concerns	Park & Tosaka, 2022
Financial Constraints	High cost of AI tools and infrastructure	61% cite cost as primary barrier (small libraries)	ACRL, 2024
Infrastructure (Developing Countries)	Insufficient ICT infrastructure	78% of Nigerian university libraries affected	Okonkwo & Nwosu, 2024
Explainability	Black-box AI decisions lack transparency	Growing ethical concern, no current standard	Hossain & Islam, 2023
Standardization	Lack of standards for AI-generated metadata	Interoperability challenges across platforms	Netanel, 2024
Privacy & Data Security	Patron data used in AI training and analytics	Regulatory compliance concern (GDPR, FERPA)	Hossain & Islam, 2023

Strategies for Effective AI-Enabled Cataloguing and Classification in Academic Libraries

To overcome the challenges mentioned above, a complex approach to strategy needs to be considered, which will include planning and developing technology, developing the working force, collaborating with other institutions, developing policies, and ensuring quality. All the listed strategies are based on the evidence of successful implementations and informed by the professional literature and research (Breeding, 2023; Chen and Williams, 2024; Okonkwo and Nwosu, 2024).

Establishing Robust Data Governance and Quality Frameworks

The libraries need to perform systematic evaluation and purification of their bibliographic data prior to the implementation of AI cataloguing systems. This is through detecting and fixing the encoding mistakes in the old MARC records, setting data quality standards and measures, and introducing data quality surveillance procedures. A data governance framework must specify accountability in data quality, provide the process of reviewing and remedying AI-generated records, and be able to train AI systems with representative, balanced, and frequently updated training data (Mugisha & Zhang, 2022).

Another type of data cleaning and normalization software, including MARCEdit (freely available) and OpenRefine, should also be invested in libraries to prepare legacy data to train AI systems. Data Sync service of OCLC and other such services can assist libraries in synchronizing their local catalogue data with good quality data in shared catalogues which can be better used in training and implementation of AI (OCLC, 2022).

Adopting a Human-AI Collaborative Cataloguing Model

The best AI cataloguing applications are collaborative models where AI applications are used to complement human cataloguers and not to substitute them. This model features AI systems addressing the high-volume, routine cataloguing, such as copy cataloguing matching, initial classification assignment, subject heading suggestions, and metadata extraction, and human cataloguers address quality check, original cataloguing of complex or unique materials, authority control decision, and system oversight. A study by Chen and Williams (2024) showed that this type of collaboration reduced the cataloguing time by 47 percent and the quality of the cataloguing was the same or even higher than fully manual cataloguing.

The model of human-AI collaboration also makes the cataloguers the AI supervisors, whose professional knowledge is shaped to assess and enhance the performance of the AI instead of being overshadowed by it. This reframing contributes to workforce resistance and preservation of professionalism in cataloguing expertise in an environment enhanced with AI.

Investing in Workforce Development and AI Literacy

The libraries need to make strategic investments in the AI and data literacy of their cataloguing personnel. These encompass education in the concepts of machine learning and NLP as it is applicable to cataloguing, skills applied to AI cataloguing software and tools, skills in data quality measurement and metadata analytics, and systems of ethical analysis of AI systems. Such professional bodies as ALA, IFLA, and CILIP have created AI competencies frameworks that library professionals should be trained on that can also inform institutional training programs (ALA, 2024).

LIS graduate programs play a crucial role in equipping future cataloguers with AI-integrated environments as well. Bawden and Robinson (2023) suggested in their curriculum review that all LIS professional preparation programs should include modules on AI ethics, data management, and computational cataloguing methods. Libraries ought to collaborate with LIS schools to create continuing education courses that will allow the professional cataloguers to gain AI-related skills without going back to full-time graduate education.

Phased and Scalable Implementation

Instead of trying to deploy AI cataloguing at a single facility once and hoping that it will work, academic libraries, especially small ones, ought to implement strategies of phased rollouts which scale up capacity gradually, but with early value returns. This could be progressively implemented as artificial intelligence aided copy cataloguing and dual identification, then to more automated hinting and classification aids in suggesting subject headings and progressing to full automation of routine cataloguing jobs as staff trust and system behaviour are confirmed (ACRL, 2024).

The financial strain of implementing AI cataloguing can be greatly minimized using the consortia and cooperative implementation models to smaller libraries. The library consortia can make AI cataloguing accessible to the libraries that could otherwise not afford individual deployment because of the high costs involved in procuring and implementing an AI platform, training, and its maintenance. Negotiated consortium-wide licensing of AI tools by the Northeast Research Libraries (NERL) and analogues has been able to lower costs on an institution-by-institution basis by 40 -60 percent (NERL, 2023).

Developing and Adopting AI Ethics and Bias Mitigation Policies

Libraries should devise clear policies that regulate the ethical application of AI in the cataloguing. Such policies are supposed to cover the requirements of algorithmic transparency and explainability, bias detection and mitigation procedures, user feedback and correcting errors in AI-generated metadata requirements, data privacy protections of any user data used in training the AI system or running it, and a process of accountability of AI-generated cataloguing decisions (Drabinski, 2019; Hossain and Islam, 2023).

AI cataloguing systems should be biased regularly (i.e., the patterns of subject heading assignment, the accuracy of classification in various subject areas and languages, and the diversity in the scope of the perspectives reflected in AI-generated metadata) and should be audited on a regular schedule and the results published openly to the library users and stakeholders. The libraries are also expected to actively advocate in professional organizations concerning the industry-wide bias disclosure standards in library AI tools (Oluwole et al., 2021).

Leveraging Linked Data and Open Standards

Moving to BIB Framework and linked data metadata models will put academic libraries in the best place to make the most out of AI cataloguing potentials whilst maintaining long-term interoperability and sustainability. Linked data structures can be better processed by AI in comparison to the traditional MARC records and they allow the establishment of more semantic relationships and better entity disambiguation (Netanel, 2024). Libraries ought to create BIBFRAME implementation roadmaps and make their AI cataloguing investments consistent with the development of linked data infrastructure.

Open standards of AI generated metadata, such as new guidelines of IFLA, Dublin Core Metadata Initiative, and national library standards bodies, will assist in achieving interoperability of AI-generated cataloguing records with other library systems and the ability to be reused across institutional boundaries. The active engagement in the standards development process allows libraries to influence the way the AI metadata standards are formed in such a manner that it represents the library values and professional knowledge (IFLA, 2023).

Metrics and Continuous Improvement

The use of AI cataloguing needs strong measurement systems to gauge the performance of the system, where the areas of improvement can be noticed, and the value should be demonstrated to the stakeholders of the institution. AI cataloguing key performance indicators (KPIs) must be metadata quality scores (accuracy, completeness, consistency), cataloguing throughput (Items processed per unit time), time saved by cataloguers, rate of successful user discoveries, error rates in AI-generated records, and cost per item catalogued (Ullah and Ameen, 2020). Reporting on these measures on a regular basis will allow libraries to show the investment payback and make informed decisions regarding optimization of AI systems.

Table 6: Strategic Framework for AI-Enabled Cataloguing Implementation

Strategy	Key Actions	Expected Outcome	Priority Level
Data Governance	Legacy data remediation, quality standards, governance framework	Clean, AI-ready training data; reliable AI outputs	High – foundational
Human-AI Collaboration	Redefine cataloguers as AI supervisors; workflow redesign	60%+ time savings; maintained quality	High – immediate

Continuation of Table 6

Workforce Development	AI literacy training, LIS curriculum reform, continuing education	Competent, confident AI-ready staff	High – ongoing
Phased Implementation	Incremental deployment; consortium-based cost sharing	Manageable investment; demonstrated ROI	High – especially for smaller libraries
Ethics & Bias Policies	Bias auditing, transparency standards, accountability frameworks	Trustworthy, equitable AI systems	High – all institutions
Linked Data Transition	BIBFRAME roadmap, open standards adoption	Future-proof, interoperable metadata	Medium – strategic
Performance Measurement	KPI framework, regular reporting, continuous improvement cycles	Evidence-based optimization and advocacy	Medium – ongoing
Advocacy & Collaboration	Vendor engagement, standards participation, professional networking	Librarian influence on AI tool development	Medium – long-term

CONCLUSION

The convergence of artificial intelligence with academic library cataloguing is one of the most dramatic changes in the technical services of libraries since the advent of the computerized cataloguing systems in the 1970s. This paper has explored the multi-faceted nature of this transformation, following the development of the academic libraries in its very institutional nature through its operations structure, the dynamics of the cataloguing and classification service specifically, and the radical effects of the digital economy on the same service.

All of the evidence discussed in the current paper supports the claim that AI-based cataloguing has enormous advantages to academic libraries: to cataloguing time can be cut by up to 60 percent, metadata gets rich and discoverable, scalable to support the growing digital collections, and interoperable with the linked data web ecosystem. The spike in the number of libraries have adopted AI cataloguing, 12% of ARL libraries in 2019 and 54 percent by 2024 show not only the compelling performance benefits of AI tools, but an increasing awareness of the need by library administrators that AI-driven cataloguing is the only way to remain relevant in the digital economy.

A complicated terrain of problems that should be critically addressed, has however, been shed light in the paper. Weaknesses in data quality of legacy catalogues, bias in algorithms capable of replicating and enhancing the historical disparities in the representation of subjects, workforce competency, the high cost especially to libraries in the developing nations, and the lack of full standards of AI-generated metadata are all real challenges to the adoption of equitable and effective AI cataloguing. The 42-percentage-point inaccuracy margin of the non-English-language material that was detected in AI classification systems is a bitter pill to swallow because technological progress does not necessarily imply fair service delivery.

The strategic plans outlined in this paper: data governance, human-AI partnership, workforce, scaling implementation, ethics and bias policies, linked data transition and performance measurement, offer a practical roadmap to academic libraries in the AI cataloguing world. More importantly, these plans point out that the meaning of AI should not be construed as an alternative to professional cataloguing skills but as an effective resource that, when correctly guided and managed by experienced cataloguers, can transform the quality, efficiency, and scale of cataloguing in a major way.

As to the future, future opportunities in cataloguing and classification are likely to arise through the further development of large language models, which are multimodal AI systems that can process text, images, and structured data concurrently and, in the future, quantum computing solutions in knowledge organization. The library profession is required to take active part in these developments to make sure that AI-based cataloguing is in support of the core values of academic librarianship, namely, equitable access, intellectual integrity, and assurance of the development of human knowledge; this can be achieved by participating in standards development, discussing with vendors, and conducting empirical research.

The digital economy will only keep on developing, introducing information forms, user expectations, competitive forces and technological opportunities to the academic library cataloguing. The libraries that invest strategically in the AI-powered cataloguing tools today and still adherent to the principles of professional values, equity, and quality will be the ones that will be well-positioned to meet the mission of knowledge management in the twenty-first century information ecosystem. It is not only an issue of technology but more of professional vision, commitment at the institutional level and will.

REFERENCES

Amarachi, O. C. (2019). Awareness and Use of Electronic Resources by Academic Staff in Nigerian University Libraries: A Case Study of Michael Okpara University of Agriculture Library Umudike, Abia State, Nigeria. Nasarawa Journal of Library And Information Science (NAJLIS), 3(1), 125-134.

American Library Association. (2025). Public Library Staffing Report: Results from the 2024 PLA survey. Public Library Association. https://www.ala.org/sites/default/files/2025-08/PLA_Staff_Survey_2025.pdf

Khang, A. (2024). Design and modeling of artificial intelligence-oriented competency framework for information technology sector. In AI-Oriented Competency Framework for Talent Management in the Digital Economy (pp. 268-291). CRC Press.

Association of College and Research Libraries. (2024). 2024 top trends in academic libraries: A review of the trends and issues. College & Research Libraries News, 85(6), 230–242. https://doi.org/10.5860/crln.85.6.230 (or direct: https://crln.acrl.org/index.php/crlnews/article/view/26379)

Association of College and Research Libraries. (2000). Information literacy competency standards for higher education. American Library Association. http://www.ala.org/acrl/standards/informationliteracycompetency

Association of Research Libraries (ARL). (2021). The impact of COVID-19 on ARL member libraries. ARL Publications.

Association of Research Libraries. (2023). ARL statistics 2022–2023. ARL Publications. https://www.arl.org/arl-statistics/

Bawden, D., & Robinson, L. (2023). Preparing LIS professionals for AI-integrated environments: Curriculum recommendations. Journal of Library and Information Science Education, 64(3), 112–129. https://doi.org/10.3138/jelis.2022-0041

Becker, B., & Yannotta, L. (2022). Academic library user satisfaction and service quality. Reference & User Services Quarterly, 61(3), 201–215. https://doi.org/10.5860/rusq.61.3.201

Breeding, M. (2023). Library systems report 2023: Revolutions in library technology. American Libraries Magazine, 54(4/5), 22–35.

Chen, L., & Williams, T. (2024). NLP-assisted subject cataloguing: A pilot study at the University of Michigan Library. Cataloging & Classification Quarterly, 62(1), 45–67. https://doi.org/10.1080/01639374.2023.2289041

Corrall, S. (2019). Repositioning library and information science for the digital knowledge economy. Journal of Documentation, 75(2), 384–405. https://doi.org/10.1108/JD-06-2018-0102

Drabinski, E. (2019). Queering the catalog: Democratic theory and the politics of correction. The Library Quarterly, 89(3), 241–268. https://doi.org/10.1086/704044

EDUCAUSE. (2024).EDUCAUSE horizon report: Teaching and learning edition. EDUCAUSE. https://library.educause.edu/resources/2024/4/2024-educause-horizon-report-teaching-and-learning-edition

Eze, J. U., & Nwosu, C. C. (2021). Cataloguing challenges in digital environments: Evidence from academic libraries in sub-Saharan Africa. African Journal of Library, Archives & Information Science, 31(1), 15–29.

Fang, W., Zhang, R., & Liu, J. (2021). Enhanced subject access and library catalog discovery rates: An empirical study. Library Resources & Technical Services, 65(4), 189–207. https://doi.org/10.5860/lrts.65n4.189

Grand View Research. (2023).AI in library market size, share & trends analysis report 2023–2030. Grand View Research. https://www.grandviewresearch.com/industry-analysis/ai-in-library-market

Hanson, C., & Hessel, H. (2021). Technical services in the time of COVID-19: Adaptation and resilience. Library Leadership & Management, 35(3), 1–18.

Hossain, M. A., & Islam, M. S. (2023). Ethical dimensions of artificial intelligence in library services: A systematic review. Library & Information Science Research, 45(2), Article 101240. https://doi.org/10.1016/j.lisr.2023.101240

International Federation of Library Associations and Institutions (IFLA). (2023).IFLA statement on AI and libraries: Principles and recommendations. IFLA. https://www.ifla.org/publications/node/94492/

Iwhiwhu, E. B., & Okorodudu, P. O. (2022). Information literacy and academic success: A study of university libraries in Nigeria. Library Philosophy and Practice, 2022, Article 7021. https://digitalcommons.unl.edu/libphilprac/7021

Lavoie, B., Childress, E., Erway, R., Faniel, I., Malpas, C., Schaffner, J., & van der Werf, T. (2019).The evolving scholarly record. OCLC Research. https://doi.org/10.25333/C3BS6H

Library and Information Science Education Network (LISEN). (2023). AI readiness survey of LIS professionals 2023. LISEN

Publications.
Library of Congress. (2022).BIBFRAME implementation at the Library of Congress: Progress report 2022. Library of Congress. https://www.loc.gov/bibframe/
Library of Congress. (2024). AI-assisted cataloguing in digitization projects: Lessons from 2021–2024. Library of Congress Internal Report.
Loth, J., Liu, T., & Ng, A. (2020). Machine learning approaches to automated DDC classification: An evaluation study. Journal of Documentation, 76(5), 1089–1108. https://doi.org/10.1108/JD-01-2020-0005
Mugisha, C., & Zhang, Q. (2022). Accuracy in subject classification: A comparative study of human and AI-assisted cataloguing. Cataloging & Classification Quarterly, 60(4), 312–331. https://doi.org/10.1080/01639374.2022.2067831
Netanel, N. (2024). BIBFRAME adoption and linked data in academic libraries: A national survey. Library Resources & Technical Services, 68(1), 4–22. https://doi.org/10.5860/lrts.68n1.4
Nicholas, D., Watkinson, A., Jamali, H. R., Herman, E., Tenopir, C., Volentine, R., & Allard, S. (2022). Researcher information discovery behavior in the digital age: An international comparison. Aslib Journal of Information Management, 74(2), 222–241. https://doi.org/10.1108/AJIM-07-2021-0207
National Information Standards Organization (NISO). (2022).Knowledge bases and related tools (KBART): Recommended practice. NISO. https://www.niso.org/standards-committees/kbart
Northeast Research Libraries (NERL). (2023). Consortium AI licensing report 2023. NERL Publications.
Nwosu, C. C., & Okonkwo, U. E. (2023). Operational dynamics of academic libraries in the digital age. Global Journal of Library and Information Science, 12(1), 33–48.
OCLC. (2022).OCLC annual report 2021–2022. OCLC. https://www.oclc.org/en/about/annualreports.html
OCLC. (2023).WorldCat: Global library cooperation. OCLC. https://www.oclc.org/en/worldcat.html
Okonkwo, U. E., & Nwosu, C. C. (2024). Artificial intelligence adoption in Nigerian academic libraries: Barriers and prospects. Journal of Librarianship and Information Science, 56(2), 178–194. https://doi.org/10.1177/09610006231189042
Oluwole, D. A., Oyelaran, O. A., & Adeyemi, T. (2021). Cross-linguistic accuracy of AI classification systems: Implications for African academic libraries. African Journal of Library, Archives & Information Science, 31(2), 101–116.
Park, J., & Tosaka, Y. (2022). Perceptions of AI tools in cataloguing: A survey of technical services librarians. Library Resources & Technical Services, 66(3), 109–125. https://doi.org/10.5860/lrts.66n3.109
Research Data Alliance (RDA). (2023).Research data management practices in academic libraries: Global survey report 2023. RDA Publications. https://www.rd-alliance.org/
Suominen, O., & Toivonen, J. (2022). Deep learning for library classification: Experiments with transformer-based models. Information Processing & Management, 59(3), Article 102925. https://doi.org/10.1016/j.ipm.2022.102925
Ullah, M., & Ameen, K. (2020). Automation of library cataloguing services: Efficiency gains and quality implications. Journal of Information Science, 46(4), 524–538. https://doi.org/10.1177/0165551519849456
Virtual International Authority File (VIAF). (2023). Virtual International Authority File annual statistics 2023. OCLC. https://viaf.org/
Yusuf, F. A., & Iwu-James, J. (2022). Resource Description and Access (RDA) implementation in academic libraries: Challenges and prospects. Cataloging & Classification Quarterly, 60(2–3), 163–180. https://doi.org/10.1080/01639374.2022.2039816