Gauging Library Needs for Advanced AI-Assisted Cataloging

1. Introduction. The University of North Texas (UNT) Department of Information Science, as the project leader, is in collaboration with UNT libraries, and requests ($128,130) from the **National Leadership Grants for Libraries** for a **2-year planning** project that aims to investigate the needs and applicability of Large Language Models (LLMs) running locally to assist the subject cataloging of digital resources. The project will take formative assessment methods to address two questions: **RQ1:** What are the perceived needs, facilitators, and barriers of applying LLMs into the cataloging processes according to library professionals? **RQ2:** How can LLMs serve as auxiliary tools to librarians in the processes of cataloging? The project will build knowledge for future development and deployment of LLM-based applications for cataloging, which is aligned with Objective 3.1: create new processes and procedures needed to advance broadly defined digital inclusion, and Objective 3.2: support innovative approaches to digital collection management and improve cataloging and inventory practices.

2. Project Justification. LLMs have unparalleled capabilities in processing and generating human-like text based on vast amounts of data, which may benefit the domain of cataloging by assisting librarians in managing the ever-increasing amount of new content being created. In previous works, various artificial intelligence (AI) methods have been investigated for their potential to aid in the comprehension, summarization, and categorization of textual content in cataloging. However, the practical application is still limited. A preliminary experiment by the PI shows that LLMs may possess greater potential for practical applications in real-world cataloging scenarios. Using the Library of Congress Classification to classify books based on titles and abstracts, LLMs running locally with zero-shot or few-shot learning can generate superior results (respectively 0.48 and 0.57 in the first level with GPT3.5 for 21 classes) compared to most existing models trained with samples. LLMs that are fine-tuned can reach an accuracy of 0.85. Although LLMs are still limited in accuracy in challenging tasks, the study demonstrated that: i) LLMs retain an advantage over traditional AI methods in the absence of training. ii) The performance of LLMs with training may generate applicable results for practical use. The capabilities of LLMs could offer a pragmatic and readily deployable framework for enhancing cataloging processes.

However, little study has yet been done to understand *how LLMs might be applied to meet librarians' real-world needs*. Preliminary interviews that the PI conducted at the UNT Libraries revealed that librarians there had given it some thought: they believed LLMs might help with (i) routine cataloging of resources submitted to the institutional repository for open access, such as Electronic Theses and Dissertations (ETDs), (ii) managing the proliferation of new controlled vocabularies by adding new terms for already-cataloged resources, and (iii) updating terminology that might be problematic due to historical bias and lack of cultural understanding. They had not, however, taken any concrete steps to explore integrating LLMs into their workflows. This reveals a potential disconnect worth further study. The investigators suspect that the real-world application of LLMs in libraries depends on multiple complex factors: the steep requirement that LLM inputs must already be digitized; libraries' needs, resources, and policies; fears around catalogers' job security; ethical considerations around copyright, privacy, and bias; and other practical considerations. A deeper investigation is needed, with different types of libraries, to gain a more comprehensive understanding of their perceived needs and barriers. *How to incorporate LLMs into cataloging librarians' workflows* is another key question. What kinds of outputs are the most helpful: summaries, keywords, controlled vocabulary terms, keywords with explanations, etc.? What are their expectations about the output quality and how tolerant are they of errors? Do they like the applications to be integrated into their current system or separate? Cataloging librarians' perceptions of the applicability will be important for developing and deploying LLM-based applications. Study results will inform future research into the practical applications of advanced AI in assisting library cataloging and digital resource management.

3. Project Work Plan. The project will include two phases using formative assessment methods to identify the needs and applicability of LLMs in cataloging. **Phase 1**: Interview and survey library administrators and cataloging librarians to gauge needs using the Delphi approach. The purpose of this study is to obtain a perceived understanding of the needs and have a comprehensive scanning of environmental factors related to the adoption of LLMs. The study will engage 12-16 participants comprising a mix of administrators and cataloging librarians from various sizes and types of libraries – i.e., academic, public, and school. The interview will include structured questions covering their perceived needs and consideration of factors from multiple perspectives – e.g., technological, organizational, ethical, and policy – that could relate to the adoption of LLMs in cataloging. Interviews will involve three rounds. In the first round, the project team will interview participants to obtain detailed inputs through phone-based or online meetings. Responses will be analyzed and summarized in a report.

In the second round, which will also be phone-based or online, the report from round one will be provided to participants. Past responses may elicit new ideas, and participants will also be asked for their opinions on the responses. In the third round, a survey with a comprehensive list of factors from different perspectives will be sent to participants. Participants will indicate their level of agreement with each item and provide answers to additional structured questions.

Phase 2: Interview and survey cataloging librarians on the applicability of LLMs, using sample LLM outputs and evaluation metrics of LLMs in subject cataloging-related tasks. The sample outputs and metrics will provide cataloging librarians with tangible outcomes of LLMs, thereby enabling them to discern how applicable the results are and understand how such outcomes might optimally support their professional tasks. The project team will first use LLMs to generate outputs for digital resources. For sample outputs, prompt strategies will be explored to generate different types of results (e.g., subject headings from controlled vocabularies, keywords, and short summaries) for various resources (e.g., ETDs, scholarly publications, and government reports). Then, programs will be developed to process digital resources automatically using zero-shot, few-shot learning, and fine-tuning of LLMs. The models will be trained and evaluated with cataloged records. Evaluation metrics will be calculated for generated subject headings (e.g., precision, recall, Jaccard similarity), keywords (e.g., relevance and specificity), and summaries (e.g., ROUGE, BLEU, and consistency). The study plans to survey 50-70 cataloging librarians from diverse backgrounds to ensure that insights are inclusive. The first part of the survey will present sample outputs and ask participants how the results should be presented and how they would like to interact with the results. The second part will present models and the performance of subject cataloging-related tasks for digital resources by LLMs (e.g., the accuracy of subject headings for ETDs with zero-shot LLM), and ask participants' opinions about whether the outputs of a certain quality could be helpful and how the results may assist cataloging. Following the survey, the project team will choose 15-20 participants for follow-up online interviews with semi-structured questions to further elicit their ideas about using LLMs to assist cataloging.

Project team and resources: The PI Dr. Lingzi Hong will lead the project, focusing on the development of programs that utilize LLMs for summarizing and categorizing content, and the design of questions and protocols for surveys and interviews. She has done preliminary work on the utilization of pre-trained deep learning models for retrieval (Luo et al., 2022) and the classification of books using LLaMA, GPT3.5, and ChatGLM. Co-PI Jason Thomale is a discovery systems librarian from UNT Libraries who works closely with cataloging librarians and will assist in evaluating LLMs' outputs, recruiting participants, and developing and conducting the survey and interviews. The UNT Digital Library can provide metadata records and full text for resources including ETDs, Congressional Research Service Reports, End of Term Publications, etc.

4. Project Results. The table below shows the expected outcomes from each phase of the project. The project will lay the foundation for future work in (1) evaluating the readiness of libraries to adopt such applications, (2) implementing applications that integrate LLMs to assist librarians in cataloging, and (3) promoting the use.

Objective	Methods	Expected Outcomes
1. Assess needs and barriers of adopting LLMs	Interview and survey with librarians using the Delphi approach.	A report on the needs and consideration of factors related to the implementation of LLMs in cataloging.
2. Identify the applicability of LLMs in cataloging	Experiment with zero-shot, few-shot, and fine-turning of LLMs for cataloging tasks. Generate preliminary results of applying LLMs in cataloging.	Prototype programs for automatic processing of resources using LLMs. Research papers on comparative analysis of the performance of LLMs and existing models in cataloging tasks.
	Survey and interview with cataloging librarians using LLMs' outputs on the applicability of applying LLMs.	A report on the preferability of application features, e.g., methods of integration and interaction, by cataloging librarians.

5. Budget Summary. The proposed project costs will be $128,130 with no cost sharing. This includes $16,104 for PI and Co-PI's summer salary, $37,894 for one graduate student's salary, $7,847 for fringe benefits, and $18,530 for one student's tuition. The budget includes $8,240 for PI/Co-PI and student travel support to conferences in Year 2 and $3,750 for participant payments. The indirect cost is $35,795 based on UNT's federally negotiated off-campus rate of 48.5% MTDC.