A Study on the Difference between AI-Generated and Human-Expert-Generated Default Cataloging Frameworks

Ruwan Gamage[1, 2], Priyanwada Wanigasooriya[3]

ABSTRACT

This research distinguishes Artificial Intelligence (AI) and human expert-generated default cataloging frameworks used in metadata entry into library catalogs. The National Library of Sri Lanka has initiated to build a s eries of human-generated frameworks. The experts have so far selected MARC21 tags to represent books & monographs, serials, theses and dissertations in catalogs. In the current study, authors utilized ChatGPT to pick MARC21 tags for the above three formats and compared them with human-selected ones. The *Jaccard similarity index* revealed significant similarity in tag selection across resource types, with coefficients of 0.33, 0.38, and 0.36 for books & monographs, serials, and theses & dissertations respectively. While AI showed competence in identifying standard fields, it displayed occasional geographical bias. Human experts considered situational needs and displayed a profound understanding of compliance rules. However, AI's suggestions for additional fields highlighted its grasp of cataloging principles. In conclusion, ChatGPT presented a robust foundation for default cataloging frameworks. The study posits that AI, in collaboration with human expertise, can enhance cataloging practices. Moreover, AI's suggestions to include additional fields should be considered for future improvements.

Keywords: Cataloging frameworks, Library cataloging, MARC21, AI, ChatGPT

[1] National Institute of Library and Information Sciences, University of Colombo, Sri Lanka

[2] Faculty of Arts, Maldives National University, Maldives

[3] Department of Library and Information Sciences, University of Kelaniya, Sri Lanka

Email : mailruga@gmail.com

Introduction

This research examines the correspondence between AI-crafted and human expert-designed default cataloging frameworks in library cataloging, a fundamental aspect of librarianship (Bowman, 2002). Default cataloging frameworks are specified sets of MARC21 tags or metadata entry fields ensuring consistency in cataloging practices throughout library systems (Biswas, 2020).

The National Library of Sri Lanka (NLSL) spearheaded the creation of these frameworks for diverse library material formats in November 2019, aiming to establish a standard guideline for metadata entry. The NLSL's Bibliographic Control Division (BCD) assembled a committee of subject experts who, by December 2022, had developed frameworks for books & monographs, serials, and theses & dissertations. Only the monograph framework (Gamage et al., 2020) is so far available to the public. This special committee is named the Descriptive Bibliographic Framework (DBIB) Committee. It includes the relevant staff from the NLSL and outside experts[1]. The authors of this paper are also members of this expert panel.

Delphi method was employed for framework creation, incorporating a validation verification step involving practitioners outside the committee. The goal was to design straightforward frameworks applicable to all library types and adhering to widely used cataloging standards and Resource Description and Access (RDA) rules as far as possible.

With the advent of ChatGPT in November 2022, AI is being explored in numerous practical contexts, including AI-expert collaboration in clinical decision-making. As the construction of DBIB frameworks for Ola Leaf Manuscripts, audio-visual materials, etc., is still ongoing, the findings of this study will be useful for future DBIB Committee endeavors.

Objectives

The primary objective of this study was to analyze and compare the similarities between AI-generated and human-generated frameworks in terms of appropriate MARC21 tag selection, specifically tailored for different resource formats.

[1] See Gamage et al. (2020) for details.

Methods

In this study, the authors utilized both ChatGPT Model 3.5 and 4 (version 2023.05.12) to propose default cataloging frameworks. Both models yielded encouraging results, yet GPT-4 delivered a more structured and comprehensive output. As a result, the conversation was continued on the GPT-4 platform, andto produce tabulated frameworks for books and monographs (B), serial publications (S), and theses and dissertations (T). Following this, ChatGPT was tasked with suggesting enhancements for RDA compliance in these frameworks. A comparison was made between the AI-constructed and human-constructed frameworks to identify the degree of similarity, considering only the main MARC21 fields and disregarding subfields to facilitate a simplified analysis.

In this study, researchers utilized the deviation from average number of tags to see overall deviation of the two methods. The Jaccard index, also known as the Jaccard Similarity Coefficient (JSC), was used to quantify the similarity between AI and human-created frameworks. The index, calculated as the size of the intersection of the two sample sets divided by the size of their union ($J(A, B) = |A \cap B| / |A \cup B|$), provides a value between 0 and 1, with a value closer to 1 denoting a higher similarity.

Although these techniques are useful, they have limitations; the deviation only considers the tag count and the JSC takes into account only the tag numbers, neither considering the relevance of tags. As a result, a qualitative analysis of the importance of selected and non-selected tags was performed.

Results

Table 1 shows the total number of tags selected in each case, with averages for comparison.

Table 1: Total number of selected MARC 21 tags for different formats by humans and AI, with the average number of tags selected and the deviation from the average.

Total number of tags	AI	DBIB	Average	Deviation
Books	17	31	24.0	7
Serial	19	22	20.5	1.5
Theses and dissertations	15	20	17.5	2.5

Table 2 shows the similarity of AI selected and human selected MARC21 tags for the three DBIB frameworks.

Table 2.Jaccard Similarity Coefficient of AI selected and human selected MARC21 tags for different types of materials

Type of Material	Jaccard Similarity Coefficient
Books	0.33
Serials	0.38
Theses and Dissertations	0.36

Discussion and Conclusion

Table 1 offers a comparative look at the total tags used across different frameworks for each resource type. A higher deviation is evident in the books & monographs framework, while lower deviations, 1.5 and 2.5, are found in the serials and theses & dissertations frameworks respectively. These findings emphasize the variability in tag selection across different frameworks, illuminating the need for diverse approaches when designing default frameworks.

JSC calculations yielded interesting results, as summarized in Table 2. All three formats demonstrated a similarity of over a third, pointing to a significant concurrence between AI and human-generated frameworks in selecting tags for different resource categories. The AI framework's alignment with the human-selected ones underscores its reliability and efficacy in this regard. Furthermore, AI suggested several additional fields for each category, including 336-338 (content, media, and carrier types), 700-710 (relator terms), and 264 (imprint and copyright notice) tags, in response to the request for RDA compliance.

While the AI successfully identified the commonly used fields in library catalogs, it occasionally exhibited geographical bias, such as the selection of tag 050. The human experts, on the other hand, considered more contextual requirements. For instance, they opted for the Main Entry Uniform Title – 130 over the Added Entry Uniform Title – 830. The AI successfully proposed RDA fields suggesting its comprehension over the newcataloging standard.

In conclusion, ChatGPT effectively lays a robust foundation for the construction of default cataloging frameworks, proposing tags in a reasonable manner. Its additional suggestions also provide valuable insight into future framework revisions.

Acknowledgment

The DBIB framework committee includes specialized staff from the NLSL and is advised by the Senior Administration of the NLSL. It also comprises researchers and practitioners engaged in cataloging and automating catalogues. The authors recognize their contribution to building the default cataloging frameworks recommended by the NLSL.

References

Biswas, R. (2020). Use of Customized MARC Framework for Recording Books in Koha: With Special References to Public Libraries of West Bengal. Library Philosophy and Practice, 1–15.

Bowman, J. H. (2002). Essential Cataloguing: The Basics. https://www.facetpublishing.co.uk/page/detail/essential-cataloguing-by-jh-bowman/?k=9781856044561

Gamage, R., Wanigasooriya, P., Hettiarachchi, N., Balasooriya, H., Gangabadaarachchi, V., Wijesinghe, A., De Silva, C. H., Subasinghe, K., Shiham, M. S. M., &Seneviratne, C. (Eds.). (2020). MARC21 Descriptive Cataloging Framework recommended by National Library of Sri Lanka. National Library and Documentation Centre of Sri Lanka.Retrieved July 2, 2023, from http://www.natlib.lk/download/Marc21.php

Jaccard Similarity. (n.d.). Retrieved July 2, 2023, from https://www.learndatasci.com/glossary/jaccard-similarity/