

Considering AI in the Growth of RDA:

Joint RDA Board and RSC Working Group on Artificial Intelligence
report, *2024-June 2025*

RDA Linked Data Forum
ALA Annual Meeting, June 30, 2025

Charlene Chou

Wider Community Engagement Officer, RDA Steering Committee

Head, Knowledge Access, New York University Libraries

Email: Charlene.Chou@nyu.edu

R | D | A
Resource Description & Access

Agenda

- Background
- Membership
- Terms of Reference
 - Terms of Reference #1
 - Deliverable: White paper
 - Terms of Reference #2
- Completed tasks
- Next steps
- Q&A

Background: Who needs a Working Group?



R | D | A
Resource Description & Access

Membership, 2024-2025

- Ahava Cohen, RSC, Israel, Working Group Chair
- Christian Aliverti, RDA Board, Switzerland
- Renate Behrens, RSC, Germany
- Colleen Barbus, RDA Board, USA
- Judith Cannan, RDA Board, USA
- Charlene Chou, RSC, USA & *representative on the PCC TG on AI*
- James Hennelly, ALA Digital Reference, *ex officio*

Terms of Reference, #1

Tasks:

- Review the documents in the reference resources section for fuller context of the issues
- Research the following questions:
 - Are there examples of currently live production services in which AI is being used to create catalogue data? If so, are these using RDA or other standards? Which AI technologies are they using?

Terms of Reference, #1 (2)

Tasks:

- Research the following questions:
 - How might AI assist in the implementation of RDA? How might it hinder the implementation of RDA?
 - What are the ethical implications of using AI that might affect RDA?
 - Are the copyright or other intellectual property considerations with regard to the AI that might apply to the implementation of RDA Toolkit?

Terms of Reference, #1 (3)

Deliverables:

- Document answering questions researched
- Progress reports to the RSC and the Board
- Survey of attitudes regarding AI in cataloging

Deliverable: White paper

[Considering AI in the growth of RDA](#), published January 16, 2025

Objectives:

- To examine how AI can enhance metadata creation and cataloging within the RDA international metadata standard.
- To provide guidance for implementation in a way that encourages standardization in metadata while holding space for upholding our own communities' core values.
- To collaborate with other international library organizations, such as IFLA, PCC, and AI4LAM, to responsibly implement AI in cataloging worldwide.

Deliverable: White paper (2)

[Considering AI in the growth of RDA](#), published January 16, 2025

Key points:

- **Enhancing, Not Replacing Human Work:** AI is envisioned as a tool to automate repetitive cataloging tasks, such as metadata tagging and authority control, thereby allowing catalogers to focus on complex, interpretative work.
- **Human Oversight Remains Crucial:** Despite AI's capabilities, human expertise is essential to ensure accuracy, cultural sensitivity, and contextual understanding in cataloging.
- **Training AI with RDA Standards:** There's a need to develop machine-readable RDA guidelines to train AI models effectively, ensuring the creation of RDA-compliant metadata.
- **Misconception Addressed:** The paper clarifies that AI adoption is not intended to reduce staffing but to augment human roles, drawing parallels with past technological advancements in libraries that expanded rather than contracted staff responsibilities.

Deliverable: White paper (3)

[Considering AI in the growth of RDA](#), published January 16, 2025

Recommendations:

- Develop AI-Compatible RDA Guidelines: Create a minimal set of RDA descriptions suitable for AI training to facilitate consistent metadata generation.
- Invest in creating and maintaining well-structured metadata sets that adhere to established standards and then use that well-formed metadata to train LLMs.
- Encourage collaboration between cataloging experts and AI developers.

Terms of Reference, #2

Tasks (April 2, 2025):

- Investigate live production services using AI for cataloguing and whether they adhere to RDA or other standards.
- Identify specific ways AI can support metadata creation and maintenance while maintaining compliance with RDA standards.
- Draft guidelines for creating an AI agent to help users navigate the RDA Toolkit.
- Develop a minimal set of entities that a dataset must include to be RDA compliant, even if the metadata is automatically generated.

Terms of Reference, #2 (2)

Tasks (April 2, 2025):

- Establish quality criteria for RDA compliant metadata (including AI generated metadata) to ensure reliability, accuracy, and consistency and elaborate on how AI can assist in quality assurance.
- Provide recommendations for transparency in the AI-assisted metadata creation process.
- Propose methods for documenting the AI-generated nature of RDA data in metadata records.

Terms of Reference, #2 (3)

Deliverables:

- Recommendations for standards and quality requirements for RDA compliant AI-generated metadata.
- Guidelines for documenting AI-generated metadata and creating an AI Toolkit Guide.
- Progress reports to the RSC and the Board

Terms of Reference, #2 (4)

Out of scope:

- Ethical considerations
- Environmental considerations
- Copyright

Completed tasks

- Written White Paper.
- Fulfilled the first Terms of Reference.
- Spoken with Library of Congress *Exploring Computational Description* (ECD) team.
- Met with the team behind [CatMELK](#).
- Discussed the progress report and guiding principles from the [PCC Task Group on AI and Machine Learning for Cataloging & Metadata](#).

Completed tasks (2)

- Communications with [PCC Task Group on AI and Machine Learning for Cataloging & Metadata](#)
 - The AI White Paper: shared with PCC AI TG for discussion.
 - Discussed the [progress report](#) (Sept. 2024-Feb. 2025) and [Draft PCC Guiding Principles for Use of AI and Machine Learning Technologies in Cataloging and Metadata Work](#)
 - The [interim report](#) (May 2025): shared for discussion



ChatGPT 4o ▾

New chat

Search chats

Library

Codex

Sora

GPTs

DALL-E

CatalogerGPT

CATMELK



CATMELK

By RCG Gamage ආ

A Cataloging Assistant who converts book details (images title/copyright pages or simple text such as ISBN or Author/Title) to MARC21 records, customized for Sri Lankan libraries

Create MARC21
record using
book images

පුස්තිකරභ සංකල්ප
සිංහලෙන්
පැහැදිලි කරන්න

What are the
commands
in CATMELK?

Describe how to
use this application

Ask anything



R | D | A
Resource Description & Access

Meeting with CatMELK

- Selective questions
 - What kinds of data sources do you use when training this model or tool?
 - How do you handle data cleaning and filtering to ensure quality?"
 - Do you ever use domain-specific corpora, and how do you go about selecting them?
 - Do you use error logs or feedback loops to guide fine-tuning?
 - Are there any plans to extend it to other resource types besides those included in DBIB (books and monographs, serial publications, dissertations and theses)?

Meeting with CatMELK (2)

- Responses: They don't know exactly what the models know and can't fully control them—though they attempt to improve results through fine-tuning or by adding more prompts.
- Resources
 - Link to their article about CatmeLK:
<https://jula.sljol.info/articles/8083/files/67121a0716dcd.pdf>
 - Sri Lankan [MARC 21 Descriptive Bibliographic Framework \(DBIB\)](#)
 - The video they prepared for CATMELK:
<https://youtu.be/2OFpedJPkHo>

Next steps



CatalogerGPT

By Glen A Greenly

CatalogerGPT creates MarcEdit format MARC records from book contents you provide as images, text, or PDF files. It can also determine Library of Congress subject headings, LC classification and Dewey call numbers, and find errors in MARC records. See <https://glengreenly.wixsite.com/catalogergpt>

Create a MARC record from the provided content.

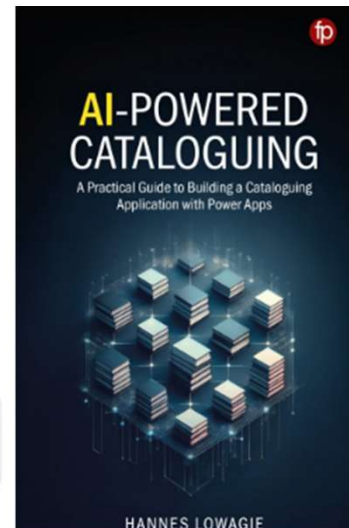
Create a MARC record for a non-existent book.

Identify and correct errors in this MARC record and add an...

Create three LC subject headings for the content.

Ask anything

+



Speak to more **tool developers**

- *CatalogerGPT*: we're still trying to get connected with Glen Greenly.
- Dr. Hannes Lowagie (*Royal Library of Belgium*): we're now trying to find a date to meet with Hannes.

R | D | A
Resource Description & Access

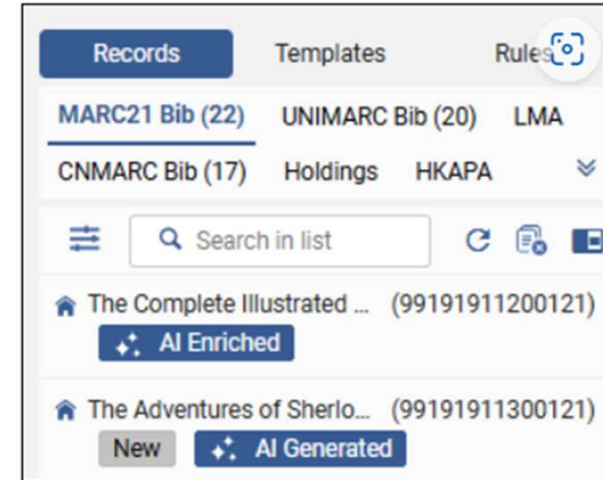
Next steps (2)

Speak to **LMS vendors**: e.g.

- Ex Libris' **AI Metadata Assistant** in the Metadata Editor: The current version is not useful, but some of the work they're doing on the next stage is interesting. A call between the WG and Ex Libris people will be scheduled in the near future.

Collaborate with the RSC Technical Working Group on **creation of prompts**

- Work on prompts which would refine what the AI engine claims to know about the Official RDA Toolkit.





Thank you!

If you have any questions, email Ahava.cohen@rdatoolkit.org or charlene@rdatoolkit.org.

*Special thanks to **Ahava Cohen** for sharing the file of her presentation from the [EURIG Meeting](#) on May 23, 2025*