



Metadata Enrichment with AI

Bringing the Power of AI to Alma CZ



Book Bibliographic Records in the Community Zone

QUALITY INDICATORS

Bibliographic ranking

- Breadth
- Depth



COMMUNITY ZONE



14M Book Bib records

76.1 Average score of book Bib records

80% Book collections in high quality
>100 activations

BOOK BIB RECORDS SOURCES

Providers | Alternative sources



Bringing AI to the Community Zone



Technology



Opportunity



Necessity

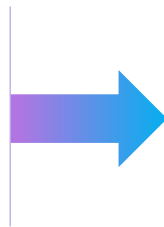
Resource Management

Provider Metadata

Generating Trustworthy Fields From the Source



Original Book



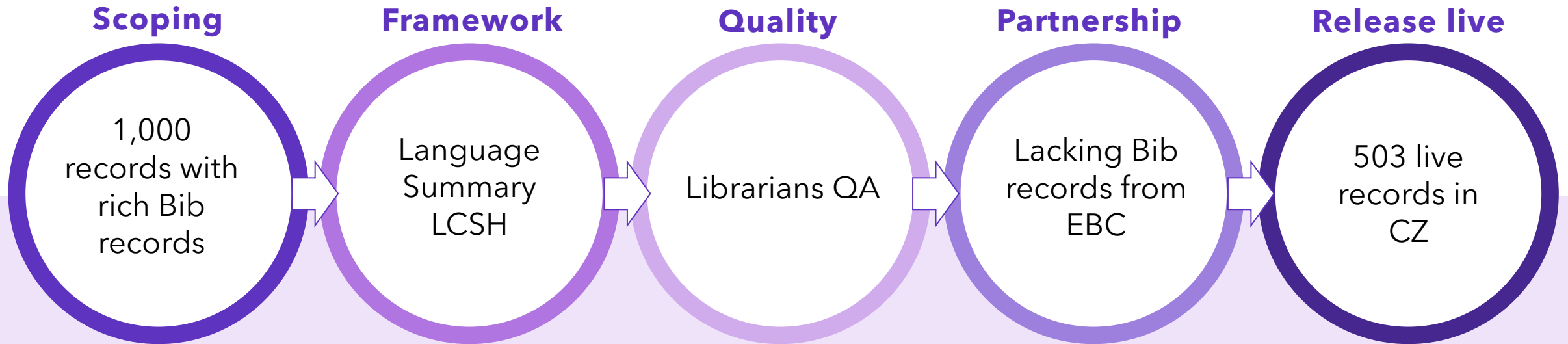
AI Metadata Generator

- Language
- Summary
- Subject Headings (LoC)
- Table of Contents



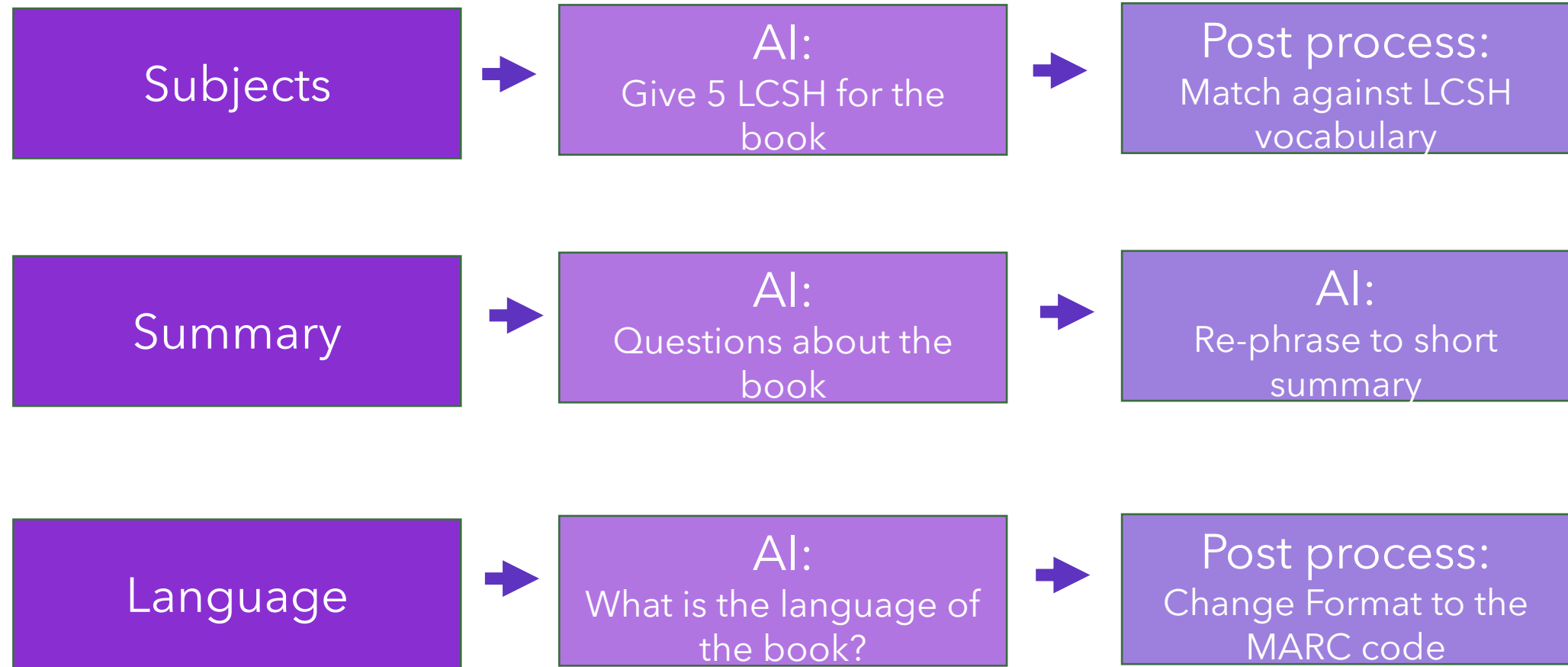
Enriched MARC Record

AI Enrichment Process



Partnership | Quality | Scale

How Does it Work?



Subjects Post Processing - Examples

AI: Artists - Correspondence

Post process: \$a Artists \$v Correspondence

(Splits to main and sub)

AI: Professional development

Post process: Career development

(Uses internal references (450->150))

AI: Business Strategy

Post process: Business planning. Similarity score: 0.929

(Finds the most similar Subject)



Taking Copyrights into Account



Legal department

- Usage
- Storage
- Output



Partner providers

- Open Access for testing
- ProQuest full text



Processing the full text

- Not copy-paste
- Actions not to risk copyright

The Challenges of Bringing AI to Providers workflow

How Do We Keep You Informed?



AI Indication

- Bib record indications
- Searchable



Transparency

- Webinar
- Knowledge Article
- Column in T&F journal




Continue to cooperate


- CZMG and CWG
- Designated email

[Institution](#) [Network](#) [Community](#)


Select All Sort by: Rank Secondary Sort by: Rank

- 1  **Tort Law in Spain.**
Book By Nasarre-Aznar, Sergio. (Alphen aan den Rijn : Wolters Kluwer Law International, 2022)
Subject: Liability (law). Torts (international Law).
Edition: 2nd ed.
Creation Date: 06/08/2023 12:06
Modification Date: 07/15/2024 09:13


● Electronic (5)

ISBN: 94-035-4694-8 and others
Record number: (CKB)6190000000011458
- 2  **Intellectual property law in Ireland**
Book (Bloomsbury Professional)
Subject: Patent Laws And Legislation. Intellectual Property.
Creation Date: 06/08/2023 12:06
Modification Date: 07/15/2024 09:13

● Electronic (3)

ISBN: 1-78043-541-X
Record number: (CKB)6180000000008082
- 3  **Conceptual Integration Theory in Idiom Modifications.**
Book By Delibegović Džanić, Nihada. (Valencia : Universidad de València Servicio de Publicaciones, 2020)
Subject: Idioms. Phraseology. Linguistics.
Edition: 1st ed.
Creation Date: 06/08/2023 12:06
Modification Date: 07/15/2024 09:13

● Electronic (4)

ISBN: 84-9134-556-6
Record number: (CKB)6180000000007611
- 4  **The Politeness/Impoliteness Divide : English-Based Theories and Speech Acts Practice in Moroccan Arabic.**
Book By Anssari-Naim, Saida. (Valencia : Universidad de València Servicio de Publicaciones, 2016)
Subject: Linguistics. Pragmatics. Speech Acts (linguistics).
Edition: 1st ed.
Creation Date: 06/08/2023 12:06
Modification Date: 07/15/2024 09:13

● Electronic (2)

ISBN: 84-370-9890-4 and others
Record number: (CKB)6180000000007295

035

__ |a (Ex-AI)996190000000011458

520	__ a Tort Law in Spain Third Edition by Sergio Nasarre-Aznar provides a comprehensive overview contractual and unlawful torts, and the different types of liabilities. Detailed explanations are provided for risky activities. It provides an analysis of strict liability including road and traffic accidents, product liability, and professional liability.
650	_0 a Liability (law). 7 Generated by AI.
650	_0 a Torts (international Law). 7 Generated by AI.

Bibliographic Rank

77



AI Bibliographic Records Enrichment



Within the Alma Community Zone, bibliographic records undergo enrichment from various sources, predominantly content providers' metadata. However, since certain book bibliographic records lack comprehensive MARC feeds, Ex Libris is exploring alternative methods for enrichment. Artificial intelligence presents an opportunity to enrich a greater number of records at scale by leveraging the full text of the book.

How does it work?

The AI-driven enrichment process utilizes the entire text of a book to create detailed metadata about the publication by extracting relevant information from the full text to populate MARC fields in the descriptive Bibliographic record.

As of February 2024, Metadata generated by AI is added to the Alma Community Zone records, focusing on enriching metadata for three specific MARC fields: Language (041 and 008), Summary (520), and LC subject headings (650). The primary focus has been on enriching Ebook Central books, which were lacking these specific metadata elements.

We are working to add more fields, including Classification Dewey and LC (082, 050), Publishing information (260/264), additional authors and contributors (100/700), table of content (505) and more.

https://knowledge.exlibrisgroup.com/Content/Knowledge_Articles/Alma/Knowledge_Articles/AI_Bibliographic_Records_Enrichment



Quality improvement



Enriching more fields

Next Steps of AI Bib Enrichment



Expanding sources



Automation



//

Do not follow where the path may lead. Go
instead where there is no path and leave a trail.

Harold R. McAlindon

//



Your Voice Matters!

Register now for our
Voice of Customer survey
launching **October 22nd**



Don't miss your chance to win a prize at IGeLU 2024 by registering for our survey!



[Sign up now](#)



Thank You

tamar.ganor@clarivate.com

AI.enriched@clarivate.com



© 2024 Clarivate

Clarivate and its logo, as well as all other trademarks used herein are trademarks of their respective owners and used under license.