

Module 08

Metadata & cataloging

Metadata work is labor intensive, detail oriented, and increasingly supported by AI tools. This module covers the real state of AI in cataloging: where it saves hours, where it introduces errors, and how to design quality-controlled workflows.

By Yulia Brusova

20 min read · Reviewed October 2026

WHAT YOU'LL BE ABLE TO DO

- 1 Apply AI to a batch metadata cleaning task in your current workflow
- 2 Evaluate at least one AI-assisted metadata tool relevant to your collection
- 3 Identify where AI-assisted subject tagging adds value and where human review is required
- 4 Draft a workflow for AI-assisted record enhancement in your system
- 5 Understand the limitations and bias risks in automated classification

Cataloging and metadata work is, in some respects, the library function most naturally suited to AI assistance. It involves pattern recognition, classification decisions, and normalization tasks that recur at scale across large and growing collections. It is also the function where errors are most consequential over time. A poor subject heading or a mislabeled record degrades discoverability in ways that are hard to detect and costly to correct at scale. To integrate AI into cataloging and metadata workflows responsibly, a digital librarian needs to know where AI speeds the work without sacrificing quality. They also need to know where professional judgment remains the only reliable quality control available.

Where AI adds genuine value in metadata work

The most reliable applications of AI in metadata work share a characteristic familiar from research support and instruction: AI is most useful when working with content the professional has already assembled, rather than generating metadata from training data alone. Suppose AI receives an item's title, abstract, table of contents, and existing partial record, and generates candidate subject headings or a descriptive summary. It is working from provided content, which carries lower hallucination risk and higher practical reliability. AI asked to produce authoritative catalog records from a title string alone is generating from training data patterns, with the attendant risk of confident errors.

Batch record cleaning and normalization is the area of most immediate practical value for most cataloging workflows. Collections accumulated over decades carry inconsistencies. Date formats vary across cataloging eras, and name authority entries appear in multiple forms. Some controlled vocabulary terms have been superseded. Encoding errors affect display and indexing, and MARC field inconsistencies cause problems in ILS export and

discovery. Given a clear normalization ruleset and a batch of records, AI can identify inconsistencies, flag records that need attention, and apply standardized corrections at a scale that would otherwise take hundreds of hours of manual review. For example, a digital librarian managing a historical collection with inconsistent birth and death date formatting across several thousand records can give Claude a CSV export, a description of the target format, and a few examples. The corrected output then goes to review. The task requires no professional classification judgment from AI, only consistent rule application, which is among the most reliable AI capabilities.

Subject heading suggestion from provided item content is a second area of consistent value. Some items have inadequate description, such as digitized photograph collections whose records contain only title and date, or dissertation repositories whose records carry no subject headings at all. For these, AI can generate candidate subject headings from provided content. For example, giving Claude a dissertation abstract and asking for LCSH-compliant subject headings produces a candidate list. The cataloger reviews it and accepts, rejects, or modifies each heading instead of constructing the list from scratch. These suggestions require professional review against the actual LCSH vocabulary and item content, but they provide a starting point that substantially reduces per-record processing time.

Description generation for digital collections addresses one of the most common digitization bottlenecks. Items have been technically processed, but minimal records without descriptive text limit their discoverability. AI can draft catalog descriptions from provided metadata, OCR text, or transcription content, which catalogers review and refine. For example, an archivist working with digitized correspondence can give AI the letter transcriptions and ask for three- to five-sentence catalog descriptions capturing sender, recipient, date, and primary subject. These drafts speed up the description workflow while leaving the substantive accuracy review in professional hands. That is the appropriate division of labor between AI efficiency and human expertise.

Batch record cleaning and normalization: a practical workflow

The workflow for AI-assisted batch record cleaning is simple enough to implement immediately. Its success depends on the quality of the instructions given to the AI, not on the model's autonomous judgment. That clarity is the professional contribution that makes AI batch processing reliable: the normalization rules come from the cataloger, not from the model.

The workflow proceeds in four stages. First, identify the specific normalization problem to address: inconsistent date formatting, name authority variants, outdated controlled vocabulary terms, or missing MARC fields of a specified type. Being precise about scope keeps AI from making changes beyond the intended correction. For example, a task specified as "normalize all 100 field entries to match the RDA form used in our authority records, no other changes" is more reliable than "improve the name authority entries." The first gives AI a specific rule to apply; the second invites judgment it is not qualified to exercise.

Second, prepare a representative sample of records in a format AI can process, typically a CSV export with relevant fields or MARC records converted to a readable format. Give AI three to five examples of the current format alongside the target format, making the transformation rule explicit rather than inferred. For example: "The 100 field currently reads 'Smith, John (1945-).' The target format is 'Smith, John, 1945-' (note the removal of parentheses). Apply this transformation to all 100 fields in the attached records and flag any record where the date information is ambiguous or missing." Explicit instructions like these produce consistent output that the cataloger can verify efficiently.

Third, process a sample before committing to the full batch. Running AI corrections on fifty records before applying them to five thousand is not overcautious. It is the quality control step that catches a misapplied rule before the error spreads across the entire collection. For example, a normalization rule that works for most name authority formats may fail on compound surnames, names with prefixes, or non-Western name structures that follow different ordering conventions. Catching these failures in a sample keeps them out of the full collection. A pilot step like this should be standard practice for any AI batch processing task, however clearly the rules were specified.

Fourth, document the transformation rules as part of the collection's maintenance record. This documentation supports quality audits: if a systematic error surfaces later, knowing which rule was applied identifies which records to check. It also supports reproducibility, since the same normalization approach can be applied to new accessions without reconstructing the workflow. A well-documented normalization workflow becomes a reusable institutional resource that gains value over time.

AI-assisted subject heading suggestion: workflow and professional oversight

AI-assisted subject heading suggestion is among the most directly labor-saving applications in cataloging, and among the most significant for quality control. The efficiency gains are

real. Giving the cataloger candidate headings, rather than requiring them to construct headings from scratch, measurably reduces per-record cataloging time, particularly for large backlogs or high-volume digitization projects. The professional responsibilities are equally real: candidate headings require review against the actual controlled vocabulary, the item's content, and the library's local practice.

The practical workflow for subject heading suggestion follows the information-in principle. Give AI as much item content as is available: title, abstract or summary, table of contents, any existing partial metadata, and a description of the collection context. For example: "Suggest LCSH subject headings for an academic dissertation with the following abstract: [paste abstract]. The library uses LC subject headings without local modifications. Suggest between three and five headings, ordered from most to least specific, and note which headings you are less confident about." This produces a candidate set from which the cataloger selects, adjusts, and authorizes headings against the actual LCSH authority file.

The most important quality control step is verification: every AI-suggested subject heading must be confirmed against the LCSH authority file before it is applied. AI suggestions will sometimes use outdated terminology or produce plausible headings that do not exist in the controlled vocabulary. They may also apply real headings that the item's content does not fully warrant. For example, AI might suggest a broad geographic or demographic heading for an item better served by a more specific term. That is not a fabrication but a judgment error, and the cataloger's expertise is needed to catch it. AI suggestions for specialized, technical, or interdisciplinary subjects also carry higher imprecision risk. The relevant controlled vocabulary may be less well represented in the model's training data than mainstream disciplinary literature.

Subject heading suggestion also needs heightened professional attention for items about communities and experiences that LCSH has historically underrepresented or misrepresented. AI trained on existing catalog records inherits their vocabulary, including terminology that contemporary reparative cataloging practice recognizes as inadequate or harmful. The section on bias and representation examines this more fully. It is the reason that "AI suggests" can never be treated as equivalent to "cataloger authorizes" in any responsible workflow.

Description generation for digital and archival collections

Description generation is the AI application with the clearest return on investment in digital collections work. The bottleneck it addresses is the gap between items that have been technically digitized and items described well enough to be discoverable, a persistent

challenge for nearly every digital collections program. Staff time for original description is limited; the number of items needing description is not. Because AI can draft descriptions from provided content, the work shifts from origination to review. That substantial efficiency gain makes backlogs tractable without sacrificing the professional oversight that description quality requires.

The most reliable approach is to give AI every relevant piece of content available for the item: digitized text if OCR is available, existing partial record fields, physical description, provenance information, and collection context. For example, for digitized historical correspondence, a prompt might read: "Write a 75-word catalog description for the following letter based on the provided transcription: [paste transcription]. The description should identify the sender, recipient, date, primary subject, and any significant details that would help a researcher determine relevance. Write in past tense, third person, plain professional language." The archivist then reviews the description for accuracy and completeness instead of drafting from scratch. Professional time goes to verification and refinement rather than initial origination.

Professional review is where AI-drafted descriptions most often need correction. AI may infer geographic settings incorrectly from sparse metadata or misidentify persons or institutions from context clues in the text. It may assign dates with more precision than the evidence warrants, or describe images in ways that reflect training data biases about what subjects and settings look like. For example, a photograph labeled only "Community gathering, 1930s" may receive an AI description that assumes a specific ethnic or regional community based on visual pattern matching. That assumption may be factually wrong, and it carries representational significance. These errors require professional correction, which is why description review must be substantive rather than cursory.

For archival finding aids, AI description generation serves a specific, well-defined function. It turns container-level entries with minimal description into fuller series and box summaries that improve discoverability without item-level cataloging for every folder. For example, consider a large archival collection with adequate series-level descriptions but no box-level descriptions. AI can draft box descriptions from the folder titles within each box, a pattern-based task it handles reliably when the folder-level data is provided as context. This does not replace the archivist's professional interpretation of the collection's significance. It extends that interpretation to a level of granularity that would otherwise take far more staff time than is available. AI-drafted finding aid descriptions can also be reviewed at a pace that matches staff availability. That makes the approach sustainable for ongoing digitization programs, not only for one-time retrospective projects.

Bias and representation in automated classification

Automated classification systems inherit the biases of their training data. In library cataloging, this is not a theoretical concern. It is a direct professional issue, with consequences for discoverability and community representation that practicing catalogers encounter in specific, identifiable ways. AI systems trained on existing catalog records will replicate, and in some cases amplify, the biases embedded in those records. Librarians deploying AI-assisted classification need to understand this mechanism well enough to design workflows that catch and correct for it.

The Library of Congress Subject Headings vocabulary has drawn sustained professional critique for decades. Scholarship and community advocacy have documented systematic inadequacy in its representation of Indigenous communities, LGBTQ+ people and experiences, communities of color, non-Western cultural and intellectual traditions, and many other historically marginalized groups. The problems go beyond obviously offensive historical terms. They include structural issues: dominant-culture framing as the unmarked default, the absence of headings that reflect community self-identification, and a granularity imbalance that gives Western European topics detailed vocabulary and non-Western cultures and communities much coarser vocabulary. For example, LCSH has historically provided far more granular terms for the political and cultural histories of Western European nations than for comparable topics in African, Asian, or Indigenous contexts. That disparity reflects the collections the headings were originally designed to serve, not the relative significance or complexity of those subjects.

AI trained on existing LCSH-governed catalog records reproduces this vocabulary. A system that suggests subject headings from existing catalog patterns will suggest the terms that appear most frequently in its training data. It will therefore suggest inadequate terms for materials about underrepresented communities, just as existing records use inadequate terms for those materials. Because AI weights high-frequency training examples, the terms most commonly applied to materials about marginalized communities are precisely the ones it will suggest most confidently. These include terms that reparative cataloging practitioners are actively working to replace. Such a system does not amplify bias by inventing new harms. It entrenches existing harm by applying it efficiently and at scale.

The professional responsibilities that follow are clear. AI subject heading suggestions require human review for all materials, not as a formality but as a substantive quality control step. For materials about historically underrepresented or marginalized communities, reviewers should be catalogers who know reparative cataloging principles and current community-driven vocabulary initiatives, such as the Cataloging Lab and relevant SACO funnel projects. The

ARL Guiding Principles for AI also explicitly call on librarians to understand and raise awareness of AI bias. In cataloging, that means applying the awareness in daily practice and contributing to institutional and professional conversations about how AI-assisted cataloging workflows are designed and governed. The efficiency gains from AI-assisted subject tagging are professionally defensible only when the quality control layer is substantive and documented. It must be applied with particular care to the collections and communities that existing cataloging systems have historically served least well.

Tools in the AI metadata space: an honest evaluation

The tools available for AI-assisted metadata work fall into two broad categories. One is integrated features within existing library systems, which catalogers reach through their established platforms. The other is general-purpose AI tools applied to metadata tasks by catalogers who have built their own workflows. Understanding what each category offers, and what responsible use requires of the cataloger, guides the practical tool decisions digital librarians face. The landscape is changing faster than professional documentation can track.

Integrated library system AI features carry the most institutional significance. They operate within systems the library already manages and are governed by existing vendor relationships and data agreements. Ex Libris has been integrating AI features into Alma and Primo across multiple release cycles, including AI-assisted metadata enrichment, subject heading suggestions, and authority control matching. For example, the AI Metadata Assistant in Alma drafts descriptive metadata and suggests subject headings based on existing record content, for a cataloger to accept, edit, or reject. Check current availability and configuration options for these features through the library's vendor account manager and the Ex Libris developer network, because releases move faster than formal documentation. OCLC, similarly, has invested in AI applications for WorldShare record quality, including duplicate detection, subject heading suggestions, and automated quality scoring. Librarians at institutions using WorldShare Management Services should review the OCLC AI feature documentation regularly, as capabilities have been expanding over the same period.

AERIE (AI Enhanced Record Improvement) is a tool designed specifically for academic library catalog record enhancement. It has been discussed in professional library technology forums and LTI assessments. Its focus on batch record improvement and subject heading suggestion addresses the same cataloging bottlenecks covered in this module. As with any specialized tool in this space, verify current availability, institutional pricing, and feature scope through direct vendor contact and current community assessments. Documentation may be outdated, a caution that applies across the entire AI metadata tool landscape.

General-purpose AI tools such as Claude and ChatGPT are a practical option for smaller-scale normalization and description tasks that do not require integration with a cataloging system. For example, a cataloger with a batch of digital object records that need descriptive summaries can develop a reliable prompt workflow in Claude without any specialized tool. The draft descriptions then go through the existing quality control process. This approach suits rule-based, content-grounded tasks, not tasks requiring authoritative MARC validation or ILS integration. To evaluate any tool in this space, apply the framework from Module 03. Understand the business model and review the privacy terms; patron search data entering external AI systems is a specific concern for catalog-integrated tools. Check for community assessments from LTI and ALA Core, and pilot with low-stakes records before applying the tool to production collections.

OCLC's December 2025 cataloging AI: classification, subject headings, and provenance

On December 8, 2025, OCLC announced a substantial expansion of AI-assisted cataloging within WorldShare Record Manager and Connexion, the two interfaces most catalogers use daily for original and copy cataloging against WorldCat. The new features suggest Dewey Decimal Classification and Library of Congress Classification numbers, along with Library of Congress Subject Headings, drawn from patterns across WorldCat's aggregate holdings data. For example, a cataloger working on a title with minimal existing copy can receive a suggested DDC number and a set of candidate LCSH terms. These come from how WorldCat's vast body of existing records has classified and described similar titles, not from a general-purpose language model's training data. Grounding in WorldCat's own holdings is the same retrieval-based design this module has favored throughout. The suggestion is anchored in records the cataloging community has already produced, not generated from an unrelated corpus. As with the subject heading workflow described earlier, the cataloger retains full accept-or-reject control over every suggestion; nothing is applied to a record automatically.

In pilot use, catalogers described the new features as a "safety net": a second set of eyes that surfaces a classification number or heading the cataloger might not have considered, particularly for titles in unfamiliar subject areas, without removing the cataloger from the decision. Pilot participants also reported saving up to twenty minutes per title. The figure is worth taking seriously as a signal of where the time savings are concentrated: the search and comparison work that previously meant manually checking similar records. It should still be treated as a vendor- and pilot-reported estimate, not an independently audited benchmark. A library piloting these features should measure its own time savings against its own baseline,

as this module has recommended for any AI-assisted cataloging workflow. Establish a calibration period and measure actual per-title time across a representative sample. Let that measurement, not the vendor's figure, inform staffing and workflow decisions.

OCLC paired this expansion with an update to its own documentation. Bibliographic Formats and Standards §3.5 now specifies how to record provenance for AI-generated metadata: how a record notes that a classification number, a subject heading, or another element originated as an AI suggestion that a cataloger reviewed and accepted. A provenance note is a small addition to a record, but it is professionally significant. A future cataloger, or a future audit, can identify which elements of a record passed through an AI-assisted step. That is exactly the documentation this module has argued for at the workflow level, now formalized in the standard itself. For a library designing its own AI cataloging workflow, BFAS §3.5 is worth reading directly. It shows how local documentation practices, such as the quality audit protocols discussed earlier, can align with a national standard instead of inventing a parallel local convention.

OCLC has also been testing AI at the collection level rather than the record level. On February 11, 2025, OCLC ran an AI-assisted WorldCat deduplication test targeting roughly five hundred thousand record pairs of print English-language books. These are pairs of records that may describe the same physical item but exist as separate WorldCat entries, for reasons ranging from cataloging-era differences to minor transcription variation. Deduplication at this scale is a different kind of task from the per-record suggestions described above. It is a pattern matching problem across an enormous record set. This module has consistently identified such tasks as well suited to AI, precisely because they require no new bibliographic content, only comparing existing records and flagging likely matches for review. Whether two records actually describe the same item, and which record to retain or merge, remains a cataloger's judgment. What AI changes is the scale at which candidate pairs can be surfaced for that judgment.

None of these developments changes the professional relationship between cataloger and AI suggestion that this module has described throughout. A classification number drawn from WorldCat patterns, a subject heading flagged as AI-suggested under BFAS §3.5, and a deduplication pair surfaced from half a million records are all, in the end, suggestions. They are inputs to a decision a cataloger makes, not decisions made on the cataloger's behalf. If the twenty-minute-per-title figure holds up under a library's own measurement, it represents time returned to catalogers for exactly this kind of judgment. That judgment applies to titles, collections, and communities that a national aggregate dataset cannot know as well as the cataloger holding the item.

Designing a quality-controlled AI cataloging workflow

A responsible AI-assisted cataloging workflow is not the existing workflow with AI added at the beginning. It is a deliberately designed process that specifies what AI will do, what human review will verify, what quality standards apply, and how the workflow will be audited and adjusted over time. That design takes more upfront thought than informal AI adoption, and it produces substantially more reliable and defensible results.

Workflow design begins with task specification: identifying precisely which cataloging tasks will be AI assisted, and which will remain entirely in professional hands. Not every cataloging task benefits equally from AI. Some require professional judgment that AI cannot reliably provide: original cataloging of complex, specialized, or culturally sensitive materials; classification decisions for items in emerging subject areas; and authority work that requires interpreting identity and community naming. For example, consider a workflow that uses AI for batch normalization of date and name authority fields while keeping original subject heading assignment entirely in professional hands. It is more defensible than one that uses AI for all MARC field generation. The first concentrates AI assistance where rule-following is sufficient and keeps professional control where judgment is required.

The review architecture is the most critical design decision. What percentage of AI output will be reviewed, at what depth, and by whom? The answers determine whether the efficiency gain from AI assistance is real or illusory. A workflow that processes one thousand records with AI and reviews ten percent is genuinely efficient only if the unreviewed ninety percent are reliably correct. Establishing that baseline reliability requires an initial period of higher review rates. Review fifty to one hundred percent of output across a range of record types to measure AI accuracy on the specific tasks the workflow handles. This calibration is not optional. It is the professional foundation that makes reduced review rates defensible and the efficiency claim honest.

Quality auditing is the ongoing component that separates sustainable AI workflows from initial experiments. After AI-processed records go live in the catalog or digital collection, review a sample at regular intervals to measure error rates, identify systematic problems, and adjust the workflow. For example, a quarterly audit may reveal that AI-suggested subject headings for a specific collection type are inaccurate more often than the overall baseline. That collection type should then move to a higher review category until the error source is understood and addressed. This auditing is consistent with ACRL subcompetency 4.1. Applying AI for task efficiency requires that the efficiency be real, which means the output must be accurate enough to serve the collection's users.

To communicate the workflow's design and quality standards to administrators and colleagues, document the workflow in a brief written protocol. Specify the tasks included, the review percentages applied, the quality thresholds that trigger workflow adjustment, and the audit schedule. This documentation supports the principle that human professionals are responsible for AI-assisted outputs. It also provides the institutional record needed to evaluate the workflow's performance and defend its use if questions arise about record quality.

PRACTITIONER NOTE

A compelling case for AI-assisted cataloging often comes from date normalization: legacy records with inconsistent date formats accumulated across cataloging eras. Consider a collection where birth and death dates in the 100 field appear in four different formats across approximately eight thousand records. It is a known problem that gets deferred for years because manual correction time is prohibitive. Run those records through Claude with a precise normalization rule and a set of examples. Review the output on a sample of two hundred records before committing to the full batch, then correct the outliers where the rule did not apply cleanly. That can shrink a project estimated at several weeks to two days of work. The experience does not reduce professional care in cataloging. If anything, seeing how precisely the rules must be specified to produce clean output sharpens attention to what professional specification requires. To find similar opportunities, audit the catalog for its most common consistency problems and ask whether each has a rule precise enough to delegate to AI. If the rule can be written clearly, the task is a candidate.

KEY TAKEAWAYS

- AI is most reliable in metadata work when applying explicit rules to provided content; batch normalization, description drafting from supplied text, and subject heading suggestions from provided abstracts all carry substantially lower risk than AI generating records from sparse inputs alone.
- Batch record cleaning (normalizing date formats, name authority variants, and controlled vocabulary) is the highest-value, lowest-risk AI application for most cataloging workflows; the rules come from the cataloger, not from the model.
- AI-suggested subject headings require professional verification against the actual LCSH authority file before application; AI inherits the vocabulary of its training data,

including outdated, superseded, and biased terminology.

- Automated classification systems replicate and can amplify the representational biases of existing catalog records; materials about historically marginalized communities require heightened human review regardless of workflow efficiency pressures.
- Integrated library system AI features (Alma, WorldShare) and specialized tools like AERIE should be evaluated through current vendor documentation and LTI assessments; general-purpose AI is appropriate for smaller-scale, rule-based normalization without ILS integration.
- OCLC's December 2025 AI features in WorldShare Record Manager and Connexion suggest DDC, LCC, and LCSH from WorldCat data with full cataloger accept/reject control; BFAS §3.5 now specifies how to record AI provenance, and a February 2025 test applied AI deduplication to roughly 500,000 WorldCat record pairs. Reported time savings of up to twenty minutes per title are vendor- and pilot-reported, not independently audited.
- A defensible AI cataloging workflow specifies task scope, establishes initial review rates calibrated to measured accuracy, and includes regular quality audits that can trigger workflow adjustment when error rates exceed acceptable thresholds.

References

APA 7TH EDITION

1. Association of College and Research Libraries. (2025, October). *AI competencies for academic library workers*. American Library Association. <https://www.ala.org/acrl/standards/ai>
2. OCLC. (2025, December). *AI-assisted cataloging in WorldShare Record Manager and Connexion*. <https://www.oclc.org/en/news/releases/2025/20251208-ai-recordmanager-connexion.html>

ACRL AI Competencies covered

Use & Application

Analysis & Evaluation

Sub-competencies: 4.1, 3.4, 4.4 · ACRL AI Competencies (2025)