ARTICLE

Exploring AI-assisted Classification for Collection Analysis

Jennifer Moon-Chung and Berenika M. Webster

ABSTRACT

Academic libraries face ongoing difficulty in aligning their collections with the needs of the communities they serve. Although the literature identified the value of using research, publications, and teaching data to inform collection development, maintaining crosswalks between these activities and library classification schemas made such work at scale impractical. This study developed a scalable workflow that maps institutional outputs to Library of Congress (LC) classification categories using a large language model to support an ongoing, evidence-based approval plan review. The workflow integrated six datasets across two domains of academic activity: research outputs and teaching activities. Where bibliographic metadata existed, they were used directly; otherwise, Anthropic's Claude Sonnet inferred LC classifications through data-specific prompts. A Python pipeline combining the Claude and bibliographic metadata APIs produced enriched datasets at scale, followed by automated LC range validation and selective human review. The outputs feed a three-page interactive Tableau dashboard: a summary landing page, a Research Outputs Explorer, and a Teaching Activities Explorer. The collection development team has used the dashboard alongside expenditure and usage data to review approval plans, highlighting the value of unifying previously siloed data sources. The workflow is designed to refresh with new outputs and improved models, offering a sustainable foundation for ongoing collection assessment, including gap analysis and alignment with institutional priorities.

INTRODUCTION

Academic libraries have always faced the challenge of ensuring that what they collect meets the needs of the communities they serve. For much of the 20th century, the common approach to that challenge was comprehensiveness. Research libraries operated under an expansive collecting mandate, assuming that breadth of holdings guaranteed relevance. The ideal collection anticipated need across disciplines, assembled on a "just in case" basis, before any specific need arose. Size was both strategy and metric: the larger the collection, the greater the probability that any research requirement could be met locally. This approach reflected a print world, where information was scarce, held locally, and slow to move.

The assumptions in this traditional collection model have been dismantled over the past three decades by a convergence of technological, economic, and institutional forces that have fundamentally altered the landscape in which research libraries operate. The shift began with the "serial crisis" of the late 1980s and 1990s, with the unsustainable inflationary costs of journal subscriptions. The move from print to electronic resources accelerated this transformation, with licensing largely replacing ownership. The internet, discovery platforms, preprint servers, and institutional repositories expanded discovery beyond the local collection, weakening the link

About the Authors

Jennifer Moon-Chung (jym20@pitt.edu; corresponding author) is Assessment Coordinator, University of Pittsburgh. **Berenika M. Webster** (bwebster@pitt.edu) is Associate University Librarian for Collections and Organizational Effectiveness, University of Pittsburgh.

Submitted: 30 April 2026. Accepted for Publication: 28 May 2026. Published: 21 September 2026.

https://doi.org/10.5860/ital.v45i3.17644

between library holdings and user access. The open access movement and associated funder requirements also accelerated this phenomenon.

Lorcan Dempsey's facilitated collection framework describes this new paradigm and the shift in how libraries understand their collecting responsibilities.[1] The traditional collection was built in anticipation of need, while the facilitated collection responds to need through a range of access models, including licensed content, resource sharing, shared print programs, patron-driven acquisition, and open access.

Meeting the information needs of students and researchers, through collections, is now a balance between multiple formats and access models. Continuous monitoring of the intellectual activity of the university community can provide invaluable input to maintain this balance and adjust collection investments accordingly.

PROJECT BACKGROUND

Against this backdrop, in the last two years the University of Pittsburgh Library System (ULS) has revised its general collections development policy and subject-specific collection principles.[2] The new policy emphasizes a sustainable approach to collecting and facilitating access to materials that meet current and future needs of the university community, support special and archival collection priorities, and focus on a few areas with historically strong collections.

General collections are built through a range of acquisition channels across the library system. Approval plans account for 15% of overall spending, with the remaining 85% directed toward journal subscriptions and databases. Unlike subscriptions, which provide broad, continuous coverage, approval plan purchasing is designed to be curated, with each profile reflecting a deliberate commitment to specific subject areas. Since these profiles can be calibrated to the institution's actual intellectual activity, selective budgets can be directed where they matter most.

Historically, these plans were seldom reviewed, and the reviews, when conducted, relied solely on the judgment of a relevant liaison librarian. Under our new policy, we aim to review them biannually, combining librarian expertise with robust data on the current teaching and research activities and granular information on the expenditure and usage of materials purchased through these plans. To understand how closely our approval plan selections align with our institution's intellectual output, we needed crosswalks between the two. Recent large language models (LLMs) enabled the testing of these mappings at scale and the creation of sustainable assessment workflows.

LITERATURE REVIEW

Collection Assessment

The systematic assessment of research library collections has evolved over decades from an impressionistic practice into a data-driven discipline grounded in standardized methodologies, quantitative metrics, and user-centered evaluation frameworks. Moreover, in line with the "just in case" model, these early approaches aimed to assess the depth and breadth of collections, largely ignoring relevance to user needs.

These early evaluations were conducted using checklist methods, expert judgment, and ad hoc comparisons with peer institutions. Such approaches were difficult to apply at scale. The 1984 paper by Wiemers et al. provides an overview of these early methods: application of standards, checklists, verification studies, and early usage studies including citation analysis.[3] A 1983 article

by Nancy E. Gwinn and Paul H. Mosher introduced Conspectus methodology.[4] Conspectus was the first standardized and reproducible methodology for describing and comparing library collections. It organized collections along Library of Congress (LC) classes and used a scale indicating collecting level, assessing the breadth and depth of the collection. Jennifer Beals noted that Conspectus-derived tools measure what is held, not what is needed; they assess depth, not relevance.[5]

The shift toward use-based and data-driven assessment gathered momentum through the 1990s and accelerated significantly in the digital era, relying on proxy measures such as circulation data or interlibrary loan (ILL) requests. George S. Bonn proposed a "use factor" ratio in which the collection use was normalized for its size,[6] while William Aguilar introduced the "ratio of borrowing to holdings" to normalize ILL transaction data.[7] Michael Hughes conducted a long-term collection use study using an automated, LC classification–based tool that incorporated circulation and holdings data to analyze use patterns over time.[8]

Bibliometric methods were applied to uncover the use of collections in the institutional research enterprise, transforming the assessment question from "what do patrons borrow?" to "what do researchers use to produce knowledge?" A 2007 study by Shi-Jian Gao, Wang-Zhi Yu and Berenika Webster examined the changing citation behaviors of postgraduate students at a Chinese university and the impact of these on collection decisions.[9] Gu et al. assessed the alignment between the University of Wisconsin–Madison's holdings and its scholars' output by measuring overlap between cited references in Web of Science publications and library items.[10] Their study uncovered gaps between what researchers cited and what the library held, framing the analysis as a replicable model for strategic collection decisions.

Another strand of the literature examines how well collections support institutional curricula. Bartolo et al. used information about interdisciplinary academic programs at Kent State University as a lens to develop approval plans that can meet their needs.[11] Sarah W. Sutton, Hunter Tolbert, and Khiana Harris examined course catalogs, using text-mining techniques, as an underused source of data to inform collection development decisions.[12]

The approach developed at Yale University Library, described by Julie M. Linden, Sarah Tudesco, and Daniel Dollar as "collections as a service," is a comprehensive framework of data-driven collection assessment.[13] It integrates circulation data, expenditure records, approval plan outputs, patron purchase requests, ILL data, and e-book usage into a unified evidence base for collection decision-making. Linden et al. frame collection assessment not as a periodic audit but as an ongoing research and service agenda.

This literature establishes several important points. First, the alignment of library collections with the activity of their communities is a recognized professional imperative. Second, achieving and verifying that alignment requires methods that bridge heterogeneous collections with research and teaching outputs. Third, the most rigorous approaches move beyond usage statistics to examine the actual knowledge work of the community against which collection adequacy should be measured. The present study contributes by introducing an approach that uses an LLM to assign LC classes to research publications and courses, creating a shared framework through which the intellectual activity of the institution and the subject scope of its library collecting foci can be directly compared.

Automated Subject Indexing

Automated subject indexing has been studied primarily in the context of cataloging workflows. Traditionally, subject indexing relies on controlled vocabularies within knowledge organization systems (KOS), which require significant human expertise and time to apply.[14] While these systems provide structure and consistency, their manual application is labor-intensive and can lead to delays and backlogs. Because subject metadata are also used beyond cataloging, for example, in collection assessment, these limitations affect not only access but also evaluation at scale.

As the volume and diversity of information resources continue to grow, computational approaches to subject indexing have been explored to improve efficiency and scalability. Early work focused on supervised machine learning techniques that leverage structured metadata and textual features. Weber et al. demonstrate that classification models can assign subject categories using large-scale metadata.[15] While these approaches reduce reliance on manual indexing, their effectiveness depends on the availability and quality of structured input data.

To address these limitations, hybrid approaches integrated supervised classification with rule-based extraction methods, improving coverage by identifying a broader range of relevant subject terms. Wartena and Franke-Maier show that combining classification and extraction methods improves recall while maintaining acceptable precision.[16] These approaches are often embedded within multistage workflows combining several processing and filtering steps. Asula et al., for example, implement such a pipeline for subject indexing, demonstrating gains in efficiency while requiring carefully designed system configuration.[17]

With advances in natural language processing, the field has shifted toward deep learning models capable of capturing contextual meaning in text. Transformer-based models, such as Bidirectional Encoder Representations from Transformers (BERT), have shown improved performance by leveraging contextualized language representations. However, while promising, these approaches often require additional components and task-specific configurations. Chou and Chu indicate that subject assignment is typically supported through combinations with similarity search and filtering techniques.[18] Morales-Henandez et al. further highlight the importance of multilabel classification, as many documents span multiple subject areas.[19] At the same time, Golub identifies ongoing challenges in achieving consistent semantic interpretation and robust evaluation.[20] Taken together, deep learning has improved automated subject assignment, but it typically relies on task-specific training data and carefully designed workflows.

More recently, generative artificial intelligence (GenAI), particularly LLMs, has been explored as a more flexible approach to subject assignment. Unlike earlier methods requiring task-specific training data and custom model development, LLMs can be applied through prompt-based interaction, drawing on broad pre-trained knowledge. This suggests that subject assignment may be applied more readily across heterogeneous data sources without the need to build and train new models for each task.

Several studies have explored this through applied library projects. Chow et al., for example, use ChatGPT to generate Library of Congress Subject Headings (LCSH) for theses and dissertations from titles and abstracts, showing that LLMs can assist within existing cataloging standards, although still requiring human validation.[21] Similarly, Morgan applies AI-based methods to assign United Nations Sustainable Development Goals (SDGs) to graduate theses in an institutional repository, highlighting practical considerations such as cost, usability, and varying tagging

accuracy across different approaches.[22] Martins further examines the use of LLMs in classification workflows, finding that while these models can reduce cognitive effort, their outputs remain inconsistent and require human oversight.[23] These studies highlight both the potential and the limitations of LLM-based subject assignment. While challenges remain, these limitations also highlight opportunities to extend subject representation beyond traditional cataloging. In particular, there is a need for approaches supporting a shared subject framework across heterogeneous data sources, enabling comparisons among research outputs, teaching activities, and library collections, as explored in this study. This direction suggests flexible workflows that can be refined and sustained within library contexts without extensive model development.

METHOD

This study integrates six datasets drawn from two domains of academic activity, research outputs and teaching activities, and applies LC classification to each record through a combination of existing bibliographic metadata and GenAI-assisted inference. These datasets represent two dimensions of academic activity that serve as proxies for user needs. Because library acquisitions are organized using LC classifications, mapping these activities supports alignment between user demand and collection structure. The method comprises four components: (1) data collection and preparation, (2) prompt design, (3) workflow automation, and (4) quality review.

Data Collection and Preparation

Research output data were drawn from two sources. One was an internal database that tracks faculty scholarly outputs for books, book chapters, and journal articles. Dissertations were obtained from the university's institutional repository. Teaching activity data were exported from the university course catalog and covered both graduate and undergraduate course sections. Faculty affiliation and departmental grouping data were validated against the university's human resources system to support the assignment of each record to its corresponding academic unit.

The temporal scope of each dataset was defined based on its relevance to recent academic activity. Journal articles, books, and book chapters were filtered to a five-year window spanning 2019 to 2025. Dissertations covered the three-year period from 2022 to 2024. Course data encompassed nine academic terms across three academic years, 2023 through 2025.

Preparation procedures varied by source but followed a common logic: each record set was cleaned, deduplicated, and mapped to the university's top-level academic unit structure, referred to as responsibility centers (RC). This linkage is central to the utility of the resulting data, as it enables subject librarians to examine research and teaching activity not only by LC subject area but also by the academic units generating it. In this way, the prepared datasets provide a structured basis for analyzing patterns of academic activity as indicators of user demand. Table 1 summarizes the final record counts after preparation.

Table 1. Final record counts by data type.

Record Type	Unique Records
Articles	13,547
Books	332
Book chapters	642
Dissertations	893
Undergraduate courses	5,241
Graduate courses	1,888

Prompt Design

The next step was to assign an LC classification to each prepared record, establishing a shared subject framework across heterogeneous data sources. For books and book chapters, classification codes were obtained directly by querying the OCLC Metadata API with ISBNs. All ISBNs in the dataset returned valid LC classification codes, yielding complete coverage for these two record types.

Because no comprehensive source of LC classification was available for journal articles, dissertations, and courses, a GenAI approach was adopted as a scalable alternative. After comparative testing with OpenAI ChatGPT (GPT-4 Turbo) and Anthropic Claude models (initially Claude 3.5 Haiku), the team selected Claude as the primary model family for subsequent prompt refinement.[24] A sample drawn from disciplines in which team members held subject expertise was used to evaluate classification quality. Within this sample, Claude's outputs were judged to be more contextually appropriate and consistent in assigning LC classifications. This evaluation was limited in scale and disciplinary breadth; a larger, more systematic comparison across models could strengthen the rationale for model selection.

Prompt design varied by data type but followed a common principle: combining multiple contextual signals to guide the model toward appropriate LC classification.

Articles. For journal articles, iterative prompt testing showed that supplying the digital object identifier (DOI) alongside the author's departmental affiliation produced stable classifications. The DOI provides journal-level context, while departmental affiliation helps resolve ambiguity in multidisciplinary venues. Because the objective was to estimate subject demand within academic units, departmental affiliation helped anchor classifications to the most locally relevant disciplinary domain. Additional testing showed that including topical keywords improved classification specificity. These keywords were retrieved via the OpenAlex API and incorporated into the prompt.[25] The final prompt (see Appendix A) requested an LC classification code.

Dissertations. For dissertations, a two-pass approach was adopted to improve classification precision. In early testing, single-pass classification based on title and abstract often produced overly broad or misaligned results. In the first pass (see Appendix B), the model extracted key subject terms from the title and abstract. In the second pass (see Appendix C), the extracted topics were combined with the title to generate a more precise LC classification. This separation of topic identification and classification improved consistency by providing the model with a distilled set of subject signals.

Courses. For course data, a single combined prompt was used (see Appendix D). Because university course titles are typically concise and contain recognizable disciplinary terminology (e.g., General Chemistry), the model was asked to recommend an appropriate textbook based on course title, subject area, and level (graduate or undergraduate), and to return the LC classification for that textbook.

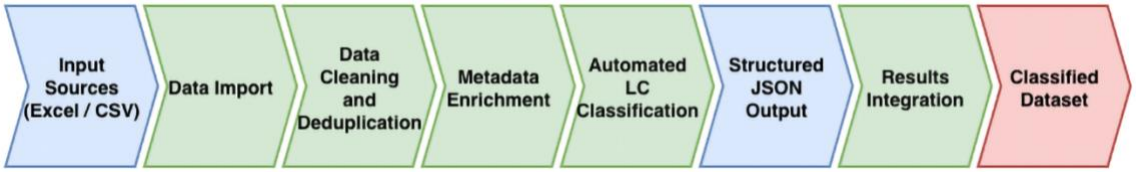
The textbook-proxy strategy was adopted because directly assigning an LC classification from course metadata is not a common task for language models and may provide limited contextual signals. However, course titles, subject areas, and level often point to recognizable curricular domains for which textbooks are commonly published and classified. Prompting the model to recommend an appropriate textbook therefore introduces a bibliographic reference point that may help connect course-level information to established LC classification practices.

Across all prompt designs, input texts were truncated to approximately 400 tokens per record to manage API costs. Early development and testing used Claude Haiku (claude-3-5-haiku-20241022), and the final production runs were completed using Claude Sonnet (claude-sonnet-4-20250514), which became available during the project.

Workflow Automation

Once prompt designs were finalized through iterative sampling, a Python script was developed to process each dataset at scale using the Claude API's message completions endpoint. The automated workflow used to generate initial LC classifications is summarized in Figure 1.

Figure 1. Automated classification workflow.



To improve consistency and operational reliability, all model calls used a temperature setting of 0, and a checkpoint-based system was implemented to reduce the risk of data loss from API timeouts or connection interruptions during large-scale processing. Results were written to a JSON Lines file after each record was processed, and the workflow checked this file on startup to identify previously completed records and resume from the point of interruption. This design allowed interrupted runs to restart without duplicating API calls or incurring unnecessary token costs.

Structured JSON output also supported integration with tabular datasets for subsequent analysis. The resulting classified dataset then moved into a separate quality assurance and deployment workflow, described in the following section, where outputs were reviewed, refined, and prepared for dashboard-based exploration and collection decision support.

Quality Review and Implementation

Following automated classification, a second-stage workflow was implemented to evaluate output quality, improve confidence in assigned classifications, and prepare the reviewed dataset for analytical use. This post-classification process combined automated checks with librarian review before results were incorporated into dashboard-ready data sources. The overall workflow is summarized in Figure 2.

Figure 2. Classification review and application workflow.

In the first stage, records were screened through automated checks. AI-generated LC classifications were compared against a reference table derived from official LC schedules.[26] The table contained 46,614 valid class ranges with associated subclass descriptions, providing broad schedule coverage at a relatively fine level of granularity. This reference was used to determine

whether returned classifications fell within recognized ranges and subclasses. For instance, a returned code of QH450.5 would be validated against the QH426-470 range for *Genetics*.

In the second stage, targeted human review was conducted through spot checking. Approximately 30 records per discipline were sampled and evaluated by librarians, who reviewed the assigned classification alongside the source metadata and model rationale. Particular attention was given to subclass specificity, interdisciplinary cases, and conceptual fit. These reviews were intended as informed quality checks rather than expert judgments in each disciplinary area.

Following review, the validated dataset was restructured into formats suitable for aggregation and visualization in Tableau Server.

DISCUSSION

Outputs and Applications

The project resulted in a three-page interactive Tableau dashboard environment consisting of (1) a summary landing page, (2) a Research Outputs Explorer, and (3) a Teaching Activities Explorer. These interfaces allowed users to move from high-level institutional summaries to more focused subject and academic unit–level exploration.

The summary page shown in Figure 3 provides a high-level overview of recent scholarly and instructional activity, including counts by output type, major LC classes, and RCs. This page serves as an entry point for identifying broad trends before moving into more detailed exploratory views.

The Research Outputs Explorer shown in Figure 4 enables liaison librarians to focus on the academic units they primarily serve. By filtering to a selected RC, users can examine how scholarly outputs are distributed across LC subject areas and publication formats. For example, when filtered to the School of Computing and Information, the Research Outputs Explorer shows a concentration of activity in QA-Science Mathematics, with journal articles representing the largest share of outputs. Users can then move from this broad LC class to a more specific range, such as QA75.5-76.95, which covers electronic computers and computer science. From there, users can review individual outputs, including title, author, year, department, and the source of the LC classification. This structure allowed users to move from disciplinary patterns to record-level context.

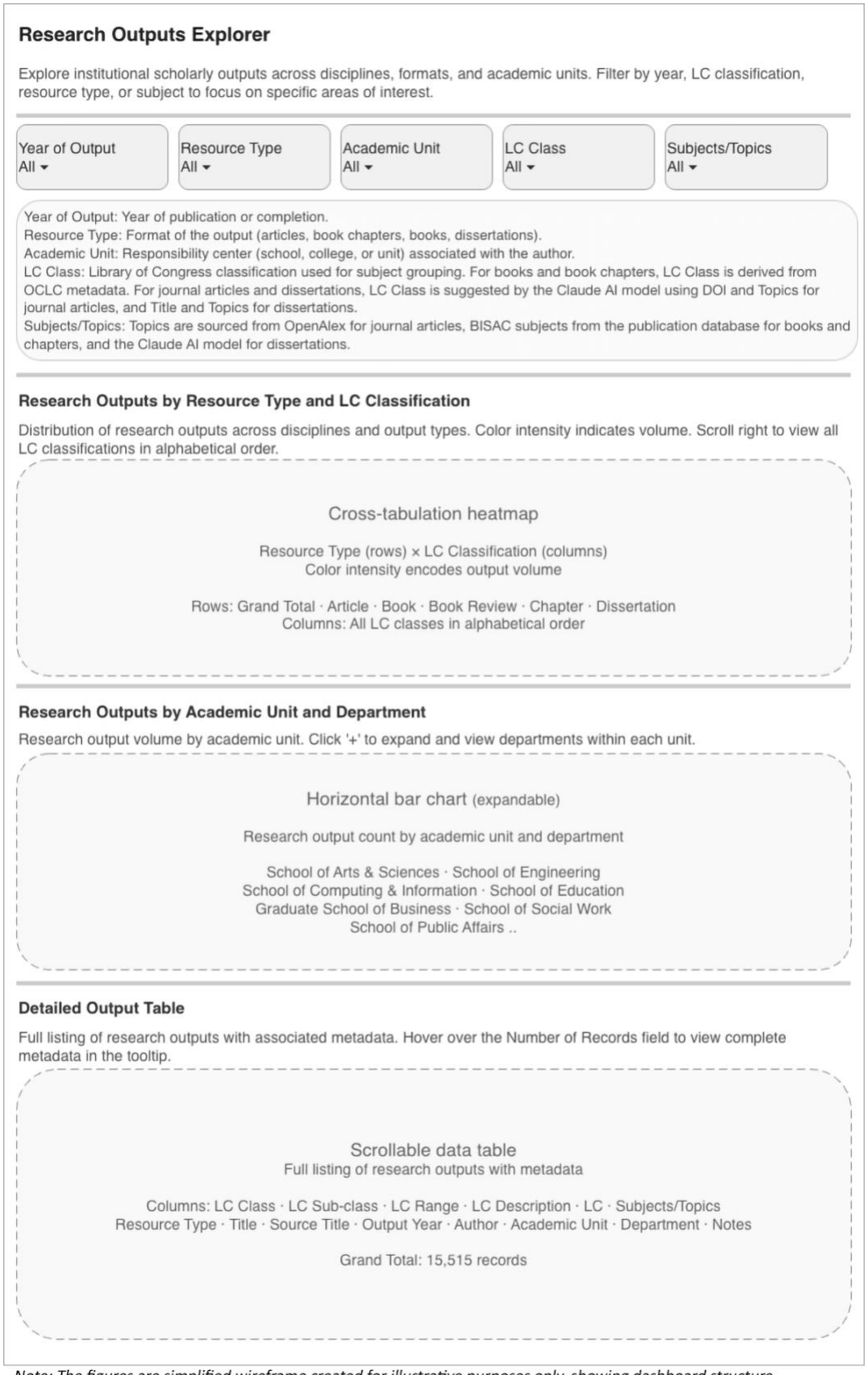
Conversely, collection decision makers could begin with a selected LC class and identify which RCs were generating activity in that area. For example, when filtered to HA, HB, HC, HD, HG, and HJ, classes broadly associated with economics, finance, commerce, and business, the dashboard showed strong concentrations from the Graduate School of Business, followed by the School of Arts and Sciences. Users can then drill down to review contributing departments and programs. This pathway helped users understand where campus demand was concentrated at both school and department levels.

The Teaching Activities Explorer shown in Figure 5 applies the same interaction logic to course data. Users can filter by academic unit, course level, or LC subject areas to examine curricular activity across departments. This view complemented research output data by providing evidence of institutional demand relevant to collection planning.

Figure 3. Summary page presenting top subject areas.

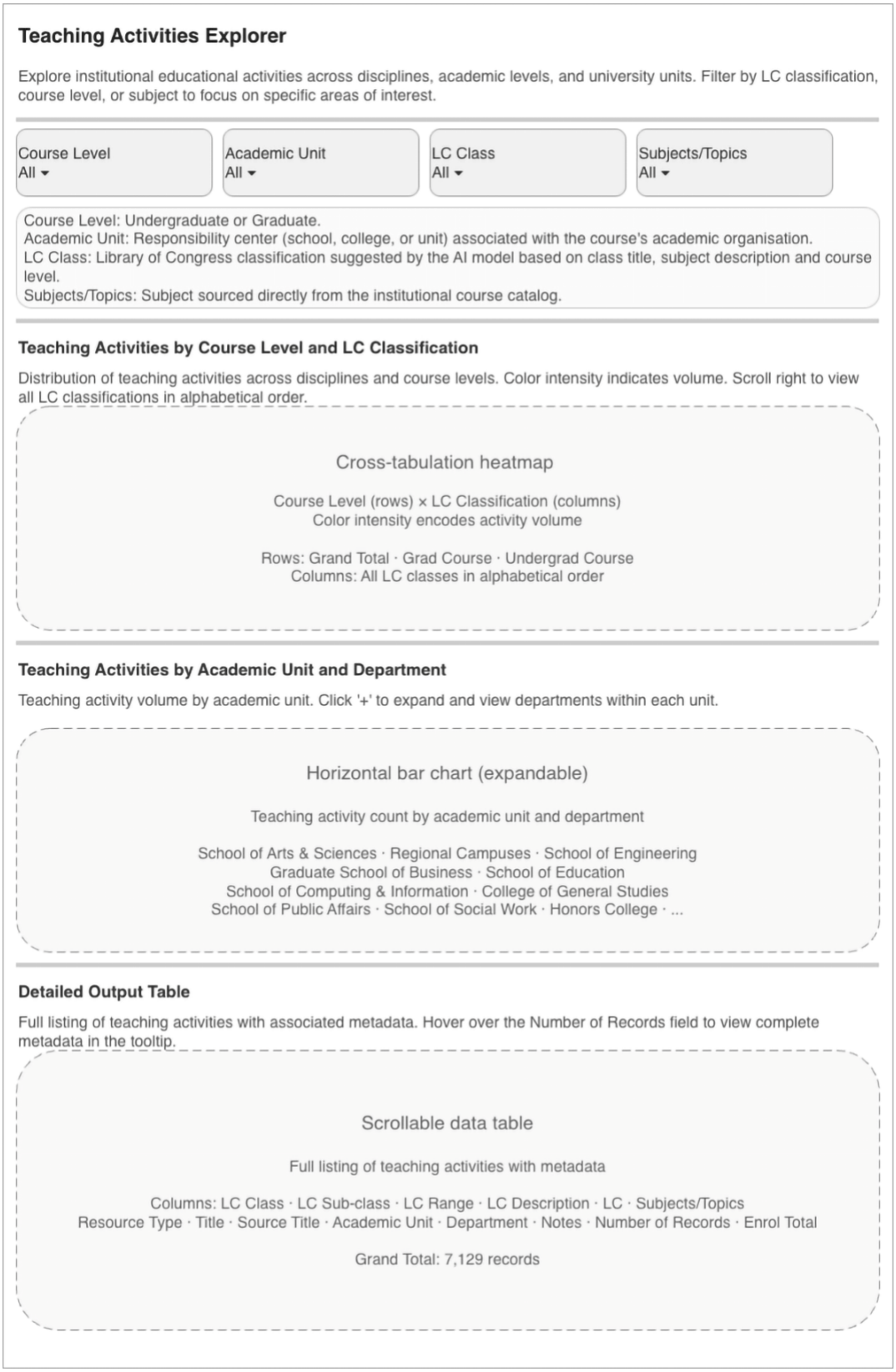
Note: The figures are simplified wireframe created for illustrative purposes only, showing dashboard structure, navigation, and example use cases. The values and labels shown do not represent actual institutional data or results.

Figure 4. Research Outputs Explorer.



Note: The figures are simplified wireframe created for illustrative purposes only, showing dashboard structure, navigation, and example use cases. The values and labels shown do not represent actual institutional data or results.

Figure 5. Teaching Activities Explorer.



Note: The figures are simplified wireframe created for illustrative purposes only, showing dashboard structure, navigation, and example use cases. The values and labels shown do not represent actual institutional data or results.

Strengths and Limitations

A primary strength of this project lies in demonstrating the possibility of using GenAI to build a scalable shared subject framework across heterogeneous institutional datasets without the need for custom model development or specialized technical infrastructure; the framework provides actionable information for collection decision-making, particularly when evaluating new subscriptions, renewal priorities, or potential cancellations. This may be especially valuable in periods of budget pressure, when decisions increasingly require evidence of institutional relevance.

A second strength is the practical nature of the workflow. The project combined automated processing, structured validation checks, human review, and a multipage interactive dashboard to create outputs that were transparent and operationally useful. Because the workflow relied on commonly available tabular data sources and repeatable scripts, it was designed to be updated on a recurring basis as new data become available. This suggests a sustainable model for ongoing decision support.

Several limitations should be acknowledged. First, the results are only as complete as the data sources used in the project. In the present work, publication data were sourced primarily from a publication database, which largely reflects tenure-stream faculty activity on the main campus. As a result, some forms of scholarly output and contributions from other campus populations may be underrepresented. Future iterations could strengthen institutional coverage by incorporating additional sources, such as institutionally maintained self-reporting systems.

Second, while classifications were refined through iterative prompt testing, spot checking, and practitioner feedback during the review process, the project did not include a formal large-scale accuracy evaluation using metrics such as precision, recall, or F-scores. The results should therefore be interpreted as operationally useful approximations rather than benchmarked classification outputs. More systematic validation across disciplines would strengthen future implementations.

Third, some limitations stem from the use of LC classifications as the organizing framework. While practical for comparison with LC-based approval plans and acquisitions workflows, LC schedules do not always reflect the most current terminology or emerging fields. In addition, many scholarly works and activities span multiple subject areas, whereas classification systems generally rely on primary placements or simplified categories.

Fourth, course data presented distinct challenges, and prompt development revealed several complexities unique to this data type. One challenge was that course titles alone were sometimes too generic to indicate subject focus (e.g., Capstone or Seminar). Another was that academic departments did not always reliably signal course discipline. Not all courses offered within a department were directly tied to that department's primary field, as some served broader institutional or general education purposes. A further challenge was that course descriptions varied in quality. Some were detailed and informative, while others were brief, outdated, inconsistent, or missing altogether. In response, multiple prompt strategies were tested. The combination of course title, subject area, and academic level produced the most balanced results by emphasizing subject context while avoiding overreliance on administrative department labels or inconsistent descriptions. For example, a course administratively associated with Economics but described within Gender, Sexuality, and Women was assigned an LC classification in the HQ range relating to women and feminism, suggesting that the model was able to prioritize

instructional content over departmental placement. While this approach produced practical results, limitations remain. Continued refinement and the incorporation of richer sources such as updated catalog descriptions, syllabi, course reserves, or reading lists could further improve future implementations.

NEXT STEPS

The data generated using the methodology described in this paper were already being used, in conjunction with usage and expenditures information, by our collection development teams to initiate review of existing approval plans and, in one case, to establish a new one. Feedback was positive, particularly regarding access to previously siloed data.

The dataset produced by this methodology could also serve as a foundation for more comprehensive collection evaluation analyses, including the extent to which current holdings support the full range of institutional academic activity. When integrated with circulation, ILL, and acquisition data, the LC-classified output data can provide a richer, evidence-based picture of collection strengths and weaknesses.

The methodology is designed to evolve alongside both the institutional outputs and the rapid developments of AI capabilities. Periodic updates will incorporate the most recent research outputs and teaching activities, while improvements in LLMs are expected to provide increasingly precise and reliable classifications over time.

APPENDIX A. JOURNAL ARTICLE CLASSIFICATION PROMPT

For each record, the prompt was populated with the article's DOI, the author's departmental affiliation, and topical keywords retrieved from the OpenAlex API. The model was instructed to return a verified LC classification, up to three Library of Congress Subject Headings, a confidence score, and a one-sentence justification in strict JSON format.

```
Your task is to categorize a journal article using ONLY officially
registered Library of Congress Classification (LCC) codes and Library of
Congress Subject Headings (LCSH)terms. Do not include any terms that are not
registered in LCSH, even if they may seem relevant. Only use terms that are
officially standardized and registered in the LCSH.

If the most relevant term is not in LCSH, do not include it.

Respond strictly in this JSON format with NO additional text:

{
  "lc_classification": "[verified LCC code]",
  "lcsh": ["[verified LCSH term 1]", "[verified LCSH term 2]"],
  "confidence_score": "[0.0-1.0]",
  "explanation": "[One-sentence justification for classification]"
}
```

APPENDIX B. DISSERTATION CLASSIFICATION PROMPT (PASS 1, TOPIC EXTRACTION)

The model received the dissertation title and abstract and was asked to extract the most relevant subject headings or topical areas, returned as a comma-separated list in a JSON object. This output served as structured input for the second classification pass.

```
Extract the most relevant subject headings or topical areas. Output should
be in one line, separated by commas, and preceded by the label:
{
  "Subject Headings/Topical Areas": "subject1, subject2, subject3",
  "confidence_score": float
}
Do not include any explanation, formatting, or text outside the JSON object.
```

APPENDIX C. DISSERTATION CLASSIFICATION PROMPT (PASS 2, LC CLASSIFICATION)

The model received the dissertation title along with the topics extracted in Pass 1 and was asked to assign the most appropriate LC classification, a confidence score, and a brief rationale explaining the fit between the assigned class and the work's disciplinary domain.

```
Your task is to categorize a dissertation into the most appropriate Library
of Congress(LC) Classification based on its title and topics. Focus on
assigning the classification that best represents the disciplinary and
academic domain of the work, not just keyword matches.

Respond strictly in the following JSON format and do not include any
explanations, extra text, or markdown formatting. Respond with only the JSON
object:

{
  "LCC": "HG3755.3.Z34",
  "confidence_score": "0.95",
  "brief_rationale": "HG3755 is for Finance--Credit. Debt. Loans--Consumer
credit.Personal loans--General works."
}


Guidelines:

- "LCC" must be the most precise and commonly used LC Classification number
appropriate for the dissertation's disciplinary field.

- Avoid classifications from unrelated LC subclasses unless clearly
warranted by the domain.

- "confidence_score" is a numerical value between 0 and 1.

- "brief_rationale" must summarize why this LCC fits the academic field and
content.

- Do not include any additional explanation, notes, or markdown formatting.
```

APPENDIX D. COURSE CLASSIFICATION PROMPT

The model received the course title, subject area, and intended audience (graduate or undergraduate) and was asked to recommend the most appropriate textbook, provide a brief explanation, and return the LC classification for that textbook.

```
You are a library collections expert. Based on the course title, subject,
and audience information below, recommend the most appropriate textbook.
Return your response in JSON format with the following fields:

- "suggested_textbook": Title and author(s) of the textbook appropriate for
graduate or undergraduate level depending on the audience.

- "explanation": A concise reason why this textbook is appropriate for the
course.

- "lc_classification": The most relevant Library of Congress Classification
code or code range for this course.


Respond strictly in the following format:

{
  "suggested_textbook": "Title by Author",
  "explanation": "Brief explanation of why this textbook is appropriate.",
  "lc_classification": "LC Classification code or range"
}
```

ENDNOTES

[1] Lorcan Dempsey, "The Facilitated Collection," LorcanDempsey.Net, January 31, 2016, https://www.lorcandempsey.net/towards-the-facilitated-collection/.

[2] University of Pittsburgh Library System, "General Collections Development Policy," ULS Collections, accessed April 27, 2026, https://library.pitt.edu/collections-overview.

[3] Eugene Wiemers et al., "Collection Evaluation: A Practical Guide to the Literature," *Library Acquisitions: Practice & Theory* 8, no. 1 (1984): 65–76, https://doi.org/10.1016/0364-6408(84)90055-3.

[4] Nancy E. Gwinn and Paul H. Mosher, "Coordinating Collection Development: The RLG Conspectus," *College & Research Libraries* 44, no. 2 (1983): 128–40, https://doi.org/10.5860/crl_44_02_128.

[5] Jennifer Beals, "Assessing Library Collections," *Electronic Journal of Academic and Special Librarianship* 7, no. 3 (2006), https://southernlibrarianship.icaap.org/content/v07n03/beals_j01.htm.

[6] George S. Bonn, "Evaluation of the Collection," *Library Trends* 22 (January 1974): 265–304.

[7] William Aguilar, "The Application of Relative Use and Interlibrary Demand in Collection Development," *Collection Management* 8, no. 1 (1986): 15–24, https://doi.org/10.1300/J105v08n01_02.

[8] Michael Hughes, "A Long-Term Study of Collection Use Based on Detailed Library of Congress Classification, a Statistical Tool for Collection Management Decisions," *Collection Management* 41, no. 3 (2016): 152–67, https://doi.org/10.1080/01462679.2016.1169964.

[9] Shi-Jian Gao et al., "A Longitudinal Investigation into the Changing Citing Behavior of Geomatics Postgraduate Students at Wuhan University, China, 1988–2004: Implications for Collection Development," *Library Collections, Acquisitions, & Technical Services* 31, no. 1 (2007): 42–57, https://doi.org/10.1080/14649055.2007.10766145.

[10] Weiye Gu et al., "Assessing the Alignment of University Library Collections with Scholarly Research Outputs: UW-Madison Case Study," *The Journal of Academic Librarianship* 51, no. 6 (2025): 103153, https://doi.org/10.1016/j.acalib.2025.103153.

[11] Laura M. Bartolo et al., "Collection Development and Interdisciplinary Endeavors: Collaborative Efforts for Educational and Work Environments," presented at the ACRL Tenth National Conference, Denver, Colorado, March 15–18, 2001.

[12] Sarah Sutton et al., "Data-Driven Collection Development: Text Mining College Course Catalogs," *Kansas Library Association College and University Libraries Section Proceedings* 14, no. 1 (2024), https://doi.org/10.4148/2160-942X.1093.

[13] Julie Linden et al., "Collections as a Service: A Research Library's Perspective," *College & Research Libraries* 79, no. 1 (2018): 86, https://doi.org/10.5860/crl.79.1.86.

[14] Fulvio Mazzocchi, "Knowledge Organization System (KOS)," in *Encyclopedia of Knowledge Organization* (Toronto: International Society for Knowledge Organization, November 11, 2024), https://www.isko.org/cyclo/kos#2.0.

[15] Tobias Weber et al., "Using Supervised Learning to Classify Metadata of Research Data by Field of Study," *Quantitative Science Studies* 1, no. 2 (2020): 1–26, https://doi.org/10.1162/qss_a_00049.

[16] Christian Wartena and Michael Franke-Maier, "A Hybrid Approach to Assignment of Library of Congress Subject Headings," *Archives of Data Science, Series A* 4, no. 1 (2018): 1–13, https://doi.org/10.5445/KSP/1000085951/22.

[17] Marit Asula et al., "Kratt: Developing an Automatic Subject Indexing Tool for the National Library of Estonia," *Cataloging & Classification Quarterly* 59, no. 8 (2021): 775–93, https://doi.org/10.1080/01639374.2021.1998283.

[18] Charlene Chou and Tony Chu, "An Analysis of BERT (NLP) for Assisted Subject Indexing for Project Gutenberg," *Cataloging & Classification Quarterly* 60, no. 8 (2022): 807–35, https://doi.org/10.1080/01639374.2022.2138666.

[19] Roberto Carlos Morales-Hernández et al., "A Comparison of Multi-Label Text Classification Models in Research Articles Labeled with Sustainable Development Goals," *IEEE Access* 10 (2022): 123534–48, https://doi.org/10.1109/ACCESS.2022.3223094.

[20] Koraljka Golub, "Automated Subject Indexing: An Overview," *Cataloging & Classification Quarterly* 59, no. 8 (2021): 702–19, https://doi.org/10.1080/01639374.2021.2012311.

[21] Eric H. C. Chow et al., "An Experiment with the Use of ChatGPT for LCSH Subject Assignment on Electronic Theses and Dissertations," *Cataloging & Classification Quarterly* 62, no. 5 (2024): 574–88, https://doi.org/10.1080/01639374.2024.2394516.

[22] Kyle Morgan, "Using AI to Auto-Tag Graduate Theses," *Information Technology and Libraries* 44, no. 4 (2025): 1–10, https://doi.org/10.5860/ital.v44i4.17381.

[23] Sugabsen Martins, "Artificial Intelligence-Assisted Classification of Library Resources: The Case of Claude AI," *Library Philosophy and Practice (e-Journal)* (2024): 8159, https://digitalcommons.unl.edu/libphilprac/8159.

[24] Anthropic, "Models Overview," Claude Platform Docs, accessed April 29, 2026, https://platform.claude.com/docs/en/about-claude/models/overview.

[25] Jason Priem et al., "OpenAlex: A Fully-Open Index of Scholarly Works, Authors, Venues, Institutions, and Concepts," preprint, arXiv, June 17, 2022, https://doi.org/10.48550/arXiv.2205.01833.

[26] Acquisitions and Bibliographic Access Directorate of the Library of Congress, "Library of Congress Classification Outline," Cataloging and Acquisitions, accessed April 27, 2026, https://www.loc.gov/catdir/cpso/lcco/.