Level 1: Foundations | All Librarians

Module 03

Picking the right tool

There are dozens of AI tools and new ones appear weekly. This module gives you a framework for evaluating them, a practical comparison of the major tools you'll encounter, and clear guidance on when not to use AI at all.

By Yulia Brusova

20 min read · Reviewed June 2026

WHAT YOU'LL BE ABLE TO DO

1. Compare the major general purpose AI tools on at least four practical dimensions relevant to library work
2. Identify what data privacy considerations apply when choosing an AI tool for professional library use
3. Make an informed decision about when AI is and is not appropriate for a given library task
4. Apply a structured evaluation framework to an AI tool you have not previously used
5. Match a specific library task to the most appropriate AI tool based on task type and privacy requirements

> *A common question librarians ask is which AI tool to use, and the honest answer is that the question itself requires reframing. Tool selection is not a matter of identifying the most popular option or the one with the most impressive marketing. It is a professional decision that involves evaluating capability, privacy terms, institutional context, and task fit, and making that evaluation deliberately rather than by default. In order to make sound tool choices, one must understand what each major tool does well, what it does not, and what happens to the information one puts into it.*

Tool selection as professional judgment

When a patron asks a librarian which database to use for a research question, the librarian does not answer by naming the largest or most widely advertised option. She evaluates the patron's topic, the type of sources needed, and the scope of the assignment, and recommends accordingly. For example, a nursing student looking for clinical evidence requires different resources than a history student writing a senior thesis, and a skilled librarian distinguishes between them. The same professional reasoning applies to AI tool selection.

There is no single best AI tool. There are tools that perform better for certain task types, tools that carry different privacy implications, tools that integrate with different institutional environments, and tools that differ in cost. Additionally, the landscape changes rapidly; a tool that was the strongest option in a given category one year may have been surpassed by the next. In order to make sound decisions across this shifting landscape, librarians need an

evaluation framework, not a fixed list of recommendations. Such a framework is what this module provides.

Two dimensions deserve particular weight in a library context: what the tool does with inputs (privacy), and whether it generates or retrieves (reliability). Both affect how the tool can appropriately be used in professional library work, and both are often poorly understood by librarians who adopt tools based on marketing or peer recommendation alone. Indeed, the professional judgment involved in AI tool selection is not different in kind from the judgment librarians have always applied to database selection, simply applied to a newer and faster moving category of resource.

Dakan and Feller name this capacity **platform awareness**: a working understanding of which AI systems exist, what each does well, and where each is unreliable (Dakan & Feller, AI Fluency: Framework & Foundations, Anthropic Academy, 2025). The concept resists the comparison-chart approach to tool selection. No static list of AI tool ratings stays accurate for more than a few months, and no single evaluation session produces the kind of judgment that holds across different task types. What actually builds platform awareness is repeated use - noticing, over time, that a particular system handles long documents reliably but loses track of nuance in multi-step reasoning, or that another performs well on structured tasks but drifts when the prompt leaves room for interpretation. Such observations accumulate into something more durable than any benchmark score. Furthermore, Dakan and Feller make a point that holds consistently in library practice: the most effective AI collaborators are domain experts first and AI users second. A librarian who can distinguish a credible source from a plausible-sounding fabrication brings professional judgment to AI output evaluation that no amount of tool familiarity can substitute for. Platform awareness grows on top of existing expertise. It does not replace it, and it is not a prerequisite for beginning.

The major general purpose tools: a practical comparison

The following tools represent the primary options available to academic librarians as of 2026. Each has genuine strengths and meaningful limitations, and the landscape has shifted considerably over the past year; tools that were experimental in 2024 are now production grade, and new entrants have become impossible to ignore.

ChatGPT (OpenAI) is the most widely recognized AI tool globally and has the largest user community, which means the most community resources, tutorials, and library specific examples available through professional channels. For example, the ACRL and ALA discussion boards have accumulated substantial practitioner experience with ChatGPT that is not yet available for newer tools. The free tier uses GPT-4o and is genuinely capable for most

drafting and editing tasks, though it imposes rate limits that become noticeable with daily use. The paid tier (Plus, approximately $20/month) removes those limits and adds access to newer GPT-5 series models as well as OpenAI's o series reasoning models, specifically o1 and o3, which are designed for complex multistep problems rather than drafting. For library work, the standard GPT-4o or GPT-5 models handle most tasks well; the reasoning models are valuable for tasks involving complex decision trees or multistep evaluation, such as assessing a vendor contract against a checklist of institutional requirements. For data privacy, inputs on the free tier may be used to improve OpenAI's models unless the user explicitly opts out in account settings, a step many users have not taken.

Claude (Anthropic) is particularly strong for tasks involving long documents and nuanced, sustained writing, and the Claude 4 model family, comprising Opus 4, Sonnet 4, and Haiku 4, has meaningfully expanded both capability and speed. Claude's context window is among the largest available, at 200,000 tokens, which means a librarian can paste an entire accreditation self study document and ask Claude to identify gaps relative to a specific standard without hitting length limits that would affect other tools. Claude's Projects feature allows users to attach persistent documents such as a collection development policy, an institutional style guide, and a list of approved databases that remain available across all conversations within the project. Such functionality is particularly useful for librarians who want consistent, institution specific output across recurring tasks. Anthropic is generally regarded as more conservative in its approach to data use, and the Pro plan ($20/month) provides full access with clearer privacy terms than the free tier.

Gemini (Google) is most valuable for librarians and institutions whose primary workflow runs through Google Workspace. Gemini integrates directly into Google Docs, Gmail, Drive, and Slides, allowing it to work with documents in the places where they already live rather than requiring copy paste workflows. The current Gemini 3.1 Pro model supports a context window of one million tokens and handles text, images, audio, and video in a single conversation, which is particularly relevant for digital collections librarians working with mixed media. Gemini also has access to Google Search as a grounding mechanism, which reduces (but does not eliminate) hallucination on factual questions. For institutions that have signed Google Workspace enterprise agreements, the privacy terms may differ from consumer Gemini; this requires verification with institutional IT rather than assumption.

Microsoft Copilot is the most important tool for librarians at institutions running Microsoft 365, and it is substantially underrepresented in library AI discourse relative to its actual relevance. Copilot integrates into Word, Outlook, Teams, Excel, and OneNote, and can draft emails, summarize long documents, and generate meeting notes directly within these applications. For example, a librarian who receives a lengthy vendor proposal as a Word

document can ask Copilot to summarize the key terms and flag anything requiring legal review without leaving the application. Institutions with Microsoft 365 enterprise agreements may already have access to Copilot as part of their existing contract, worth verifying with IT before purchasing a separate AI subscription.

Perplexity is designed to combine AI generation with real time web search, and every response includes citations to the pages from which it drew information. For example, a librarian wanting a quick overview of a new federal policy affecting library funding can ask Perplexity and receive a summary with source links to verify. Such citations reduce but do not eliminate hallucination risk, since the model may still misrepresent sources it cites. Perplexity is most useful for quick factual orientation where seeing sources alongside the answer matters more than depth of analysis.

Grok (xAI) is worth awareness even if it is not yet widely discussed in library professional literature. Grok is built into the X platform and is also available as a standalone product. Its distinguishing feature is real time access to X platform content, which makes it useful for tracking current professional conversations and emerging library community responses to policy changes. For library practice, Grok is most relevant as a complement for monitoring professional discourse rather than as a primary drafting or research tool; its privacy terms and institutional data governance implications require the same evaluation as any other commercial AI tool before professional use.

AI tools built for library work

Beyond the general purpose tools, a substantial category of AI products is now built directly into the systems libraries already manage, and over the past year this category has moved from announcement to general availability across most major platforms. For library practice, the shift matters because these tools retrieve from controlled, curated sources rather than generating from a language model's training data, which carries a fundamentally lower hallucination risk for bibliographic and discovery work than general purpose AI.

The clearest example is Ex Libris's Primo Research Assistant, now generally available to all Primo customers. It performs retrieval augmented generation, commonly abbreviated RAG, over the Central Discovery Index, a shared metadata index of roughly five billion records, and returns the five sources most relevant to a patron's query along with citations back to the underlying records. The University of Florida is among the institutions running it in production. Such a design means the tool is constrained to recommend sources that actually exist in the index, which is precisely the structural guardrail against fabricated citations that Module 01 identified as missing from general purpose AI. Clarivate offers a parallel feature for

institutions on its Summon discovery platform, the Summon Research Assistant, built on the same retrieval principle.

Cataloging and metadata work has seen comparable movement. Ex Libris has added an AI Metadata Assistant to Alma, and on December 8, 2025, OCLC announced AI cataloging features added to WorldShare Record Manager and Connexion, which suggest Dewey and Library of Congress classification numbers and Library of Congress Subject Headings drawn from WorldCat's aggregate data. For example, a cataloger working through a backlog of items with minimal existing metadata can use these suggestions in order to narrow the search space quickly, but verifying each suggested classification or heading against the item in hand remains the cataloger's responsibility; the suggestion is a starting point, not a substitute for judgment about what the item actually is. In digital collections work, Specto has become a generally available tool for AI assisted description, organization, and discovery of digitized materials, extending the same retrieval grounded approach to special collections and archives.

The database layer has moved as well. ProQuest Research Assistant has been integrated into ProQuest Central since February 2025, and database level AI assistants are now embedded directly in JSTOR, in EBSCO as AI Insights, in Ebook Central, and in Scopus. For a researcher working within any of these platforms, the AI assistant operates over that platform's own indexed content, which means its summaries and suggested sources are traceable back to records the library already provides access to, rather than to an unknown training corpus.

None of this displaces general purpose AI from library work; it complements it. ChatGPT, Claude, Gemini, and Copilot remain the tools librarians reach for day to day, for drafting an email, brainstorming a workshop outline, or rewording a policy paragraph, because those are generative tasks where the lack of grounding in a specific institutional dataset is not a liability. The vendor tools described above earn their place for discovery and metadata work precisely because they are grounded, and at a scale, the full Central Discovery Index, the full WorldCat database, that no individual librarian could review manually. What neither category of tool can do is decide, for a specific patron or a specific item in hand, whether the retrieved source actually answers the research question or whether the suggested subject heading actually describes what the item is about. That determination is a professional judgment, made by a librarian who knows the patron's need or has the item in hand, and it remains irreplaceable regardless of how good the retrieval becomes.

Free vs. paid tiers: what the difference means for library practice

The free tier of most major AI tools is genuinely useful and represents a reasonable starting point for a librarian exploring AI for the first time. For example, the free version of ChatGPT using GPT-4o can draft a LibGuide introduction, summarize a document, or generate a set of workshop objectives at a quality level that is immediately useful for most practitioners. Such capability at no cost is substantial, and librarians who have not yet tried these tools have access to more than is commonly assumed.

However, several meaningful differences separate free from paid tiers. Rate limits are the most immediate: free tier users encounter usage caps that interrupt workflows during periods of heavy use, particularly near semester peak periods or during preparation for major instruction programs. Paid tiers remove or substantially raise these limits. Model access is the second difference: free tiers may restrict access to the most capable model version, or provide access on a limited rotating basis. Furthermore, paid institutional plans, specifically the Team and Enterprise tiers for ChatGPT, Claude, and Gemini, typically include contractual data privacy commitments that free tiers do not, which is a significant consideration for professional library use involving any patron context.

In order to advocate effectively for institutional AI access, librarians should be prepared to make the case in terms administrators understand: time saved on drafting, reduced revision cycles, more responsive patron communication, and clearer data governance compliance. Such arguments, grounded in workflow efficiency and institutional risk management rather than novelty, tend to be more persuasive than capability demonstrations alone. Indeed, the data privacy argument, namely that a paid institutional plan provides contractual protection that a free consumer account does not, is often the argument that moves library administration and IT from caution to action.

Data privacy: what each tool does with your inputs

Data privacy is the dimension of tool selection that most directly implicates professional library ethics, and it deserves more attention than it typically receives in informal librarian conversations about AI. In order to understand the implications, it is useful to know what "using inputs for training" actually means. When a user submits a prompt and receives a response, the tool provider may log both the prompt and the response, use them to evaluate model performance, and incorporate them into future training rounds. For example, if a librarian's prompt describes a patron's research question in identifying detail, that information has potentially left the librarian's control, and, depending on the tool's terms, may have left the institution's jurisdiction.

The general rules as of 2025 are as follows. Free consumer tiers, including ChatGPT free, Claude free, and Gemini personal account, typically reserve the right to use interaction data for training, with opt out available in some cases through account settings. Paid consumer plans, specifically ChatGPT Plus and Claude Pro, generally offer stronger opt out terms but not always contractual protections. Paid institutional plans, including ChatGPT Team/Enterprise, Claude Team/Enterprise, and Google Workspace Business/Enterprise with Gemini, typically include contractual commitments not to use customer data for training. Such commitments carry legal weight in ways that opt out checkboxes in account settings do not.

Additionally, data storage and jurisdiction matter for institutions with specific compliance requirements. For example, some institutions subject to HIPAA, FERPA, or state level student privacy laws may face restrictions on which AI tools are permissible for certain data types, regardless of the tool's general privacy policy. The appropriate resource for institution specific guidance is the library's IT department, general counsel, or compliance office, not the AI vendor's marketing materials. Librarians who develop working relationships with their IT colleagues around AI governance will be better positioned to make responsible tool recommendations than those who navigate privacy questions independently.

When not to use AI

There is no doubt that AI is inappropriate for certain library tasks, and identifying those tasks clearly is as important as knowing when AI is useful. Such clarity is part of what distinguishes professional AI use from uncritical adoption, and it is the practical expression of the skepticism that ACRL identifies as a core competency mindset.

AI is not appropriate in the following situations:

When the task requires verified citations. AI fabricates references, as discussed in Module 01. A librarian should never use AI to find sources; AI may be used to process, summarize, or discuss sources that the librarian has already verified through authoritative channels.

When the input involves patron identifiable information in an unapproved tool. Patron privacy obligations apply to what a librarian puts into a prompt. A detailed reference question that could identify an individual patron should not go into a free tier consumer tool, and should not go into any tool that the institution has not reviewed and approved for that use case.

When the task requires high stakes factual accuracy. Statistics, legal information, medical information, accreditation standards, and institutional policy, any claim where being wrong

has significant professional or institutional consequences, must be verified from primary, authoritative sources rather than accepted from AI output. For example, a librarian advising a patron on FMLA eligibility or a faculty member on fair use should consult authoritative legal resources, not an AI summary.

When the patron expects human judgment and empathy. Some reference interactions involve personal circumstances that require discretion, sensitivity, and professional judgment that AI cannot replicate. For example, a student navigating an academic integrity process, a patron researching a sensitive health situation, or a faculty member in a difficult publication dispute: these interactions require a human librarian. AI may inform the librarian's preparation, but it should not mediate the interaction itself.

When the institution has not approved the tool. Many institutions have developed or are developing AI acceptable use policies. Using an unapproved tool for professional work, even for a seemingly low risk task, creates compliance and liability exposure that the librarian bears.

A framework for evaluating new tools as they emerge

New AI tools appear at a rate that makes it impossible to evaluate each one thoroughly as it launches. In order to stay current without being overwhelmed, it is useful to work through a structured set of questions before adopting or recommending any new tool. Such a framework makes evaluation repeatable, defensible, and faster than approaching each new tool from scratch.

The questions to apply are as follows. First: who built this tool and what is their business model? Free tools often monetize user data, and tools without a clear business model may disappear or change terms without notice, a meaningful risk for any library workflow that depends on them. Second: what does the privacy policy say, specifically, about training data use, data retention, and data jurisdiction? Vendor privacy pages vary enormously in clarity and specificity, and the details that matter most for library use are rarely in the top level summary.

Third: is there library community discussion about this tool? LTI (LibTech Insights) from Choice360, Library Technology Reports from ALA TechSource, and LITA/CORE discussion forums are the most reliable sources for library contextualized assessment, and they tend to identify practical limitations that vendor marketing does not mention. Fourth: does my institution have a policy about this tool, or is it subject to a category exclusion in the

acceptable use policy? Checking with IT before adopting is significantly more efficient than explaining an unauthorized tool adoption after the fact.

Fifth: is there a genuine library specific use case, or is this a general tool being marketed into library contexts without library relevant features? Such marketing is common. Sixth: can I pilot it with low stakes tasks such as internal drafts, brainstorming, and non patron facing content before using it for anything that affects professional output or patron services? Additionally, piloting with a colleague rather than alone produces faster and more reliable assessment, since different task types surface different limitations.

Matching tool to task: a working guide

The practical question most librarians face is not which tool is best in the abstract but which tool is most appropriate for the specific task at hand. For example, the tool best suited for summarizing a forty page policy document is not necessarily the same tool best suited for drafting patron facing instruction emails, and neither is the same tool best suited for a quick factual lookup where source citations matter. Understanding this task and tool relationship is the applied expression of the professional judgment this module develops.

The following principles guide tool selection by task type. For drafting and editing tasks, including emails, instruction content, LibGuides, policy documents, and committee reports, Claude is generally preferable for longer documents or anything requiring sustained tone consistency, and ChatGPT is useful for shorter tasks where its large community of examples supports rapid iteration. For tasks requiring Google Workspace integration, such as drafting in Docs, managing in Sheets, and summarizing in Drive, Gemini is the practical choice given its native integration with those environments. For institutions running Microsoft 365, such as generating reports in Word, drafting emails in Outlook, and taking meeting notes in Teams, Copilot is worth exploring even if it is less discussed in library AI circles, since it may already be available through the institutional Microsoft agreement.

For factual questions requiring source citations (quick research overviews, policy summaries with verifiable sources), Perplexity is a useful complement to, but not a replacement for, library database search. Such a tool is appropriate for orientation to a topic, not for authoritative information. For bibliographic and catalog level tasks, including discovery search, citation recommendation, cataloging, and digital collections description, the platform specific tools described above, Primo Research Assistant, Summon Research Assistant, Alma's AI Metadata Assistant, OCLC's WorldShare and Connexion cataloging suggestions, Specto, and the database level assistants in ProQuest Central, JSTOR, EBSCO, Ebook

Central, and Scopus, are more appropriate than general purpose AI, because they retrieve from controlled sources rather than generating from training data.

In order to develop genuine judgment about tool selection, deliberately varying the tools used for low stakes tasks over a period of several weeks, noting where each tool's outputs differ meaningfully, and building that experiential knowledge before applying it to higher stakes professional work produces the most reliable results. Such deliberate practice is the most reliable path to tool discernment, the kind of calibrated, task specific judgment that ACRL subcompetencies 3.4 and 4.5 describe as the professional standard.

PRACTITIONER NOTE

One approach to developing an institutional tool policy is through conversation rather than top down mandate: reference and instruction work uses Claude Pro under an institutional account for anything involving patron context, and ChatGPT free is acceptable for internal drafts and brainstorming that do not involve patron specific information. Such a division of use typically takes about three months of informal experimentation to develop, and the most useful forcing function is having everyone attempt the same task on two different tools and compare the results directly. In order to develop the same kind of institutional calibration, identifying two or three recurring task types and running a structured comparison across available tools produces immediate, visible differences and helps the team build shared vocabulary for discussing and justifying tool choices.

KEY TAKEAWAYS

- Tool selection is a professional judgment, not a popularity contest. Evaluate by task type, privacy terms, and institutional fit.
- Free tiers are genuinely useful but lack the contractual data privacy protections that paid institutional plans provide.
- AI is inappropriate for verified citations, patron identifiable data in unapproved tools, and any high stakes factual claim.
- Match tool to task: Claude for long documents, Gemini for Google Workspace, Copilot for Microsoft 365 environments.

- Apply a structured evaluation framework, including business model, privacy policy, community feedback, and institutional approval, before adopting any new tool.
- AI native library tools, including Primo Research Assistant, Summon Research Assistant, Alma's AI Metadata Assistant, OCLC's WorldShare and Connexion cataloging suggestions, Specto, ProQuest Research Assistant, and database level assistants in JSTOR, EBSCO, Ebook Central, and Scopus, retrieve from controlled sources and carry lower hallucination risk for bibliographic and metadata work. General purpose LLMs remain the tools librarians reach for day to day drafting and ideation; the two categories complement rather than replace each other.

References

APA 7TH EDITION

1. Dakan, R., & Feller, A. (2025). *AI fluency: Framework & foundations* [Online course]. Anthropic Academy. https://anthropic.skilljar.com/ai-fluency-framework-foundations

ACRL AI Competencies covered

Analysis & Evaluation | Knowledge & Understanding

Sub-competencies: 3.4, 4.5, 2.3 · ACRL AI Competencies (2025)