Algorithms in the Stacks: Investigating automated, for-profit diversity audits in public libraries

Melanie Walsh
University of Washington
Seattle, USA
melwalsh@uw.edu

Connor Rey
San Diego State University
San Diego, USA
cfranklinrey@sdsu.edu

Chang Ge
University of Michigan
Ann Arbor, USA
changgge@umich.edu

Tina Nowak
University of Washington
Seattle, USA
tnowak@uw.edu

Sabina Tomkins
University of Michigan
Ann Arbor, USA
stomkins@umich.edu

Abstract

Algorithmic systems are increasingly being adopted by cultural heritage institutions like libraries. In this study, we investigate U.S. public libraries' adoption of one specific automated tool—automated collection diversity audits—which we see as an illuminating case study for broader trends. Typically developed and sold by commercial book distributors, automated diversity audits aim to evaluate how well library collections reflect demographic and thematic diversity. We investigate how these audits function, whether library workers find them useful, and what is at stake when sensitive, normative decisions about representation are outsourced to automated systems designed by private vendors. Our analysis draws on an anonymous survey of U.S. public librarians (n = 99), interviews with 14 librarians, a sample of purchasing records, and vendor documentation. We find that many library workers view these tools as convenient, time-saving solutions for assessing and diversifying collections under real and increasing constraints. Yet at the same time, the audits often flatten complex identities into standardized categories, fail to reflect local community needs, and further entrench libraries' infrastructural dependence on commercial vendors. We conclude with recommendations for improving collection diversity audits and reflect on the broader implications for public libraries operating at the intersection of AI adoption, escalating anti-DEI backlash, and politically motivated defunding.

CCS Concepts

• **Social and professional topics** → **Automation**; • **Information systems** → **Digital libraries and archives**.

Keywords

audit, diversity, DEI, diversity audit, metadata, algorithmic accountability, cultural heritage, libraries, books

ACM Reference Format:

Melanie Walsh, Connor Rey, Chang Ge, Tina Nowak, and Sabina Tomkins. 2025. Algorithms in the Stacks: Investigating automated, for-profit diversity audits in public libraries. In *The 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25), June 23–26, 2025, Athens, Greece.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3715275.3732140

FAccT '25, Athens, Greece

ACM ISBN 979-8-4007-1482-5/25/06
https://doi.org/10.1145/3715275.3732140

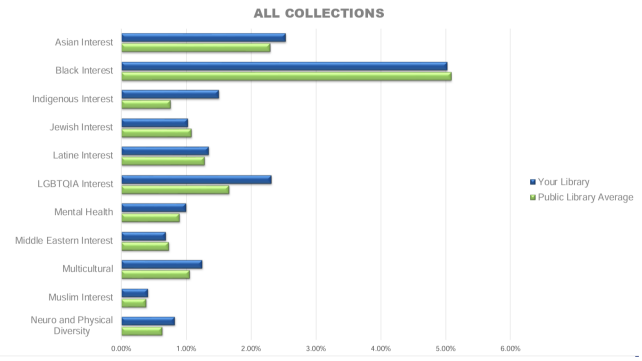


Figure 1: A library collection diversity audit report from the vendor Ingram, showing representation of books per diversity category.

1 INTRODUCTION

Algorithmic systems are increasingly being adopted by cultural heritage institutions like libraries, museums, and archives. In U.S. libraries, these tools now include recommendation systems that personalize borrowing suggestions [3, 43, 85, 99]; chatbots that assist patrons with questions and search [7, 102, 103, 118]; predictive algorithms that forecast book circulation and checkout hold times [42, 43]; and automated tools that measure the "diversity" of thousands or even millions of library books (Figure 1). Like algorithmic systems adopted by other public institutions—such as courts, schools, and hospitals [51, 108]—these developments offer potential advantages and risks. While automated tools are often embraced to perform tasks faster, at scale, and with a veneer of objectivity, extensive research at FAccT and elsewhere has shown that they can also reinforce and exacerbate biases, even when the goal is to root them out [45, 95, 101].

In this paper, we focus on the growing use of automated collection *diversity audits* in public libraries—tools that quantitatively

measure how well library holdings reflect demographic and thematic diversity. These tools are typically developed and sold by third-party vendors, and are rapidly becoming normalized in library practice. While we focus on U.S. public libraries, collection diversity audits are increasingly being used in school and academic libraries, and outside the U.S. [76, 97]. Broadly, we ask: **What happens when sensitive, value-driven cultural work is partly outsourced to automated systems developed by commercial vendors? Can these tools meaningfully support the values they claim to advance?**

Central to these questions is the term *diversity* itself. While at its most basic, diversity means difference within a group, the term has become widely used in U.S. institutional contexts to refer to structurally marginalized human identities and backgrounds, and especially Diversity, Equity, and Inclusion (DEI) efforts [75]. We acknowledge the complexity and long history of the term *diversity* in the U.S. [10, 74, 91]. For clarity, we broadly use *diverse* and *diversity* as they are deployed by vendors—to signal books that represent historically marginalized identities or subjects—and we examine how their systems define and operationalize those terms.

Collection diversity audits have emerged in response to longstanding inequities in both librarianship and book publishing [5, 13, 113]. While 61% of the U.S. population identified as white in the most recent census, over 81% of librarians and 72% of publishing industry employees identified as white as of 2023 [5, 27, 82]. The disparities are even greater in published books. Between 1950 and 2018, over 95% of the most widely held titles released by mainstream U.S. publishers were authored by white writers, according to a study by So and Wezerek [113]. Because public libraries rely heavily on mainstream publishers, their collections are likely similarly skewed, limiting the kinds of stories and perspectives that are available to readers. Diversity audits aim to assess and address these gaps. They can prompt libraries to seek out more queer love stories; more narratives with Muslim protagonists; more books by and about Black and Indigenous people and other people of color.

Yet in the increasingly hostile and reactionary political climate of the U.S.—marked by rising book bans and anti-DEI legislation [4, 9, 11, 14, 54, 60, 116]—diversity audits are not only used as tools to improve equity. They have also been used defensively: to demonstrate that library collections are not ideologically skewed toward progressive agendas and to push back against censorship efforts [44, 106].

While manual audits date back almost a decade [66], the practice accelerated in 2021 when commercial vendors like Baker & Taylor, Ingram, and OverDrive introduced large-scale audit tools powered by proprietary metadata and automated classification [62, 114, 117]. Where manually auditing just the children's section of a library might take months, these tools promise to audit entire collections in weeks or days. The companies also provide "shopping lists" and recommended books to fill diversity gaps, which are typically books sold by the same vendor performing the audit.

Despite their increased prevalence and popularity, little academic work has examined the effectiveness, accuracy, or implications of automated collection diversity audits [123]—and few FAccT studies have explored algorithmic interventions in cultural heritage settings [61, 67]. To address these gaps, we draw on an anonymous survey of U.S. public librarians (n=99), in-depth interviews with 14 librarians, purchasing records from a sample of public libraries, and an analysis of vendor documentation and promotional materials. We ask:

- **RQ1:** How prevalent are automated collection diversity audits in libraries, and who provides them at what cost?
- **RQ2:** How do automated collection diversity audits actually work, and what counts as a diverse book?
- **RQ3:** Do library workers find these audits nuanced, trustworthy, and useful?

We find that automated audits from commercial vendors are widely used and, for many library workers, offer a convenient tool for assessing collections under real and increasing constraints. Their appeal lies not only in their speed and scale, but also in their framing as neutral or objective—a framing that becomes especially valuable amid heightened political scrutiny of library content and DEI initiatives. Yet these tools often flatten the meaning of diversity, ignore structural drivers of inequality, and obscure the community-specific needs that shape thoughtful collection development. While the for-profit nature of these tools is not inherently problematic, we find that commercial incentives discourage methodological transparency, promote proprietary data hoarding, and deepen customer lock-in. Book vendors now play a central role not only in supplying materials for collections but also in providing the data and analytics through which libraries interpret their collections. This "double-dip" with data assets resembles information consolidation dynamics described by Lamdan [72].

Ultimately, we suggest that library collection diversity audits are a critical case study for understanding the benefits and risks of using automated tools in cultural heritage settings, especially tools provided by commercial vendors. While these tools do not currently (to our knowledge) incorporate Artificial Intelligence (AI) techniques, they exemplify the dynamics increasingly emerging with the application of AI to sensitive, subjective, and context-dependent cultural work. We conclude with recommendations for improving these audits and reflect on the broader implications for libraries navigating the fraught intersection of AI adoption, escalating anti-DEI movements, and politically motivated defunding.

2 RELATED WORK

2.1 Audits

Audits, formal evaluations of systems, are a foundational concept in FAccT research, where they are most often applied to algorithmic and AI systems [24, 84, 98, 100]. In recent years, FAccT scholars have turned their attention not only to algorithms but to the audit process itself, interrogating its assumptions, politics, and institutional uses [37, 56, 121]. We draw on both strands of this work because collection diversity audits operate as automated tools and audits. In the field of Library and Information Science (LIS), scholars and librarians have begun to document and analyze collection diversity audits [49, 52, 55, 111, 123, 126]. We build on their work with a systematic analysis of how these tools operate in practice.

2.2 Race, Diversity, and Algorithmic Classification

A substantial body of research has shown that technology and algorithmic systems often replicate and exacerbate social inequalities,

particularly along lines of race, gender, and other axes of identity [20, 26, 31, 94]. These harms are not always the result of explicitly encoded categories but often emerge through data, design decisions, and institutional assumptions.

In the field of algorithmic fairness, race and social identity are often explicitly operationalized—that is, converted into measurable categories—in order to evaluate whether systems produce discriminatory outcomes. While this is frequently done in the name of justice and anti-discrimination, the process of categorizing and quantitatively measuring social identity can reproduce harmful histories of classification, especially when underlying assumptions go unstated and unexamined [8]. Further, Hanna et al. [57] argue that the operationalization of race often obscures structural causes of discrimination: "treating race as an attribute, rather than a structural, institutional, and relational phenomenon...serves to minimize the structural aspects of algorithmic unfairness."

For decades, feminist and critical race theorists have also interrogated the institutional uses of *diversity*, sometimes critiquing its ambiguity, depoliticization, and neoliberal cooption, and sometimes acknowledging its strategic utility [10, 74, 90, 91]. These critiques resonate with concerns about diversity in recommendation systems, in which scholars have questioned how diversity is abstractly operationalized and argued that it must be defined in relation to the specific goals and contexts of each system [124]. Scholarship from algorithmic fairness, recommendation systems, and feminist and critical race theory thus highlights problems and complexities inherent in operationalizing race and diversity as collection diversity audits do.

2.3 Measuring Race and Identity in Books

Scholars in the digital humanities (DH) and LIS have long grappled with the challenges of categorizing race and other aspects of social identity in books, especially with computational methods or metadata. In DH, scholars have emphasized that computation and racist discrimination are historically and structurally entangled [53, 89]. While researchers have used computational approaches to explore race, gender, and social identity in literary and cultural texts [29, 30, 112], many foreground the myriad challenges involved in doing so [86]. In fiction, characters' identities are not always stated outright; they can be fluid or ambiguous; they can be contingent on historical period, geography, or context. Much of this is also true for real-life authors [112]. Researching an author's background is not only time-consuming but often leads to findings that defy neat categorization—as So and Roland [112] point out, the American author Nella Larsen is typically identified as Black today but in the 1920s (when she was writing) was often referred to as "mulatta," a term that is outdated, offensive, and reflective of a different racial classification system.

Within LIS scholarship, similar concerns have been raised about the classification of race, ethnicity, gender, sexuality, and Indigeneity through metadata like *subject headings*. In particular, scholars have critiqued the U.S. Library of Congress Subject Headings (LCSH) for outdated, inconsistent, or offensive treatment of marginalized identities [19, 21, 22, 39, 40, 71, 81, 127, 129]. Our work builds on these DH and LIS traditions to critically examine how diversity categories are constructed and measured with collection audit tools.

3 BACKGROUND

3.1 Major Library Vendors in the U.S.

Public libraries rely on a range of commercial vendors for everything from furniture and security systems to digital platforms and book distribution. This study focuses on three major players in the public library market—Baker & Taylor, Ingram, and OverDrive—that sell books as well as automated and data-driven tools, including diversity audit tools. Companies like Midwest Tape and LibraryIQ also offer tools for collection diversity audits [78, 119]. Some free, nonprofit tools exist, as well, such as Diverse BookFinder, which focuses on picture books, and a tool from the Cooperative Computer Services (CCC) Consortium [49]. We focus on these three vendors because they are widely used and exemplify the broad entanglement of book distribution, metadata infrastructure, and automated systems.

- **Baker & Taylor** (founded 1828): The largest U.S. library supplier, offering book distribution, cataloging, metadata, and collection development services. Acquired *collectionHQ* in 2011 to provide data-driven collection analysis and algorithmic tools.
- **Ingram** (founded 1970): The largest overall U.S. book distributor, serving both libraries and retailers like Barnes & Noble. Offers cataloging, processing, and AI-driven tools. Owns the self-publishing platform IngramSpark.
- **OverDrive** (founded 1986): The leading distributor of digital content (ebooks, audiobooks) to libraries. Operates the Libby app and videostreaming app Kanopy. Serves 95%+ of U.S. public libraries [25].

3.2 From Book Distribution to Data Analytics

The role of commercial library vendors like book distributors has expanded significantly over time—particularly with the rise of digital materials and computing technologies, and the decline of public funding for libraries. Rather than buying books directly from publishers, libraries have long relied on wholesalers like Baker & Taylor (a company that has existed for almost 200 years), to supply books in bulk. When catalog systems were computerized in the 1970s and 1980s—replacing physical card catalogs that listed each book's title, author, subjects, and location in the library—vendors gained a new foothold because they could ship books with machine-readable metadata (known as MARC records), allowing libraries to automatically populate catalog entries at scale [16, 69, 88, 104, 105].

Vendors' early foothold in metadata allowed them to expand further into the business of data and algorithmic tools in the twenty-first century—especially as collections have grown and become more digital, and as libraries have faced staffing shortages, technical skill gaps, and funding issues [35, 68, 87]. In 2011, for example, Baker & Taylor acquired the Scottish startup product *collectionHQ*, which helps libraries manage their collections by analyzing borrowing patterns and predicting future circulation with machine learning. These tools are used both to guide new book acquisitions and to recommend which titles should be "weeded," or removed, from the

collection—a core task for libraries with limited shelf space and budgets.

These developments resemble the broader trajectory of information conglomerates like RELX and Thomson Reuters, which have acquired more and more information companies (e.g., Elsevier, Westlaw), and which can "double-dip" with their massive data assets—selling both raw data and structured analytics derived from that data [72]. While vendors like Baker & Taylor differ from these conglomerates in many ways, including the entities they contract with, they too have acquired more information companies (e.g., collectionHQ) and more library data, and now sell back analytics and tools built from this aggregated data.

3.3 Collection Diversity Audits

Library collection diversity audits emerged as a coherent trend in the 2010s with predominantly manual processes [65, 66, 123]. Of course, attention to the lack of diverse books in library collections and the publishing industry pre-dates the 2010s [28, 73]. Many librarians have long been conscientious about developing collections that offer "mirrors, windows, and sliding glass doors" for their communities [23]. However, in the past, librarians mostly lacked formalized processes to measure the diversity of their collections comprehensively.

3.3.1 How Do (Manual) Collection Diversity Audits Work? There are many documented instances of libraries conducting manual audits, including at public, academic, and school libraries [32, 41, 92, 109]. In a manual collection audit, librarians typically define a list of diversity categories—often related to race, ethnicity, religion, disability, and/or LGBTQ+ representation. To measure how many books are in these categories, they often sample sections of the collection and physically handle the books, evaluating book covers, illustrations, publication dates, subject headings, and sometimes reading excerpts. Auditors may additionally consult book reviews, prize lists, or author biographies—although the identification of #OwnVoices titles (where author and protagonist share a minoritized identity) has faced criticism due to privacy concerns [65, 66, 123].

3.3.2 Rise of Automated Vendor Audits. In 2020, collection diversity audits became more prevalent in the U.S. Responding to the unjust murder of George Floyd, the 2020 Black Lives Matter (BLM) protests inspired organizations around the U.S. to invest in various DEI and anti-racism initiatives. Many interviewees reported that the confluence of BLM protests and the COVID-19 pandemic—which forced many libraries to close to the public and thus loosened some operational capacity—was a key reason they began a diversity audit. In summer 2021, Ingram launched a diversity audit service called *iCurate inClusive* [62], and Baker & Taylor introduced a *DEI Analysis* tool through their subsidiary, *collectionHQ* [117]. Soon after, other vendors like OverDrive [114], Midwest Tape (which owns Hoopla, provider of ebooks and digital content) [119], Follet [115], and Mackin [83] also developed diversity audit/analysis services.

3.4 Book Classification and Subject Headings

Libraries and vendors primarily rely on established classification systems to categorize books, and these systems play a central role in automated diversity audits. Standardized topical labels—often called *subject headings*—are typically assigned when a book is entered into a library system, and serve as a key way to indicate what a book is "about."

Two of the most commonly used systems are **Library of Congress Subject Headings (LCSH)** and **Book Industry Standards and Communications (BISAC)** codes. LCSH, maintained by the U.S. Library of Congress, is a controlled vocabulary that includes topics, geographies, time periods, genres/forms, ethnic groups, and more. While LCSH is widely used, there have been ongoing criticisms of its treatment of race, ethnicity, gender, sexuality, and Indigeneity, and substantive change has been slow [19, 22, 39, 40, 71, 81, 127, 129]. BISAC headings, developed by the Book Industry Study Group (BISG), are used primarily in commercial publishing and bookselling, though some libraries use this system, too.

Title	Library of Congress Subject Headings (LCSH)	Book Industry Standards and Communications (BISAC) Subject Headings
Fun Home (2006) Alison Bechdel	• Bechdel, Alison, 1960– — Comic books, strips, etc. • Cartoonists — United States — Comic books, strips, etc. • Graphic novels • Comics (Graphic works) • Nonfiction comics	• COMICS & GRAPHIC NOVELS: Nonfiction / Biography & Memoir • COMICS & GRAPHIC NOVELS: LGBTQ+ / General • COMICS & GRAPHIC NOVELS: Humorous
Beloved (1987) Toni Morrison	• African Americans – Ohio – History – 19th century – Fiction • African Americans – Social conditions – To 1964 – Fiction • African American women – Fiction • Enslaved women – Fiction • Slavery – United States – Fiction • Infanticide – Fiction • Ohio – Fiction • Post-Civil War	• Fiction / Literary • Literary Fiction • Popular Fiction
The Satanic Verses (1988) Salman Rushdie	• Airplane crash survival – Fiction • Life change events – Fiction • East Indians – England – Fiction • London (England) – Fiction	• Fiction / Literary • Fiction / Fantasy / Paranormal • Fiction / Psychological • Popular Fiction

Table 1: Subject headings for three well-known fiction titles. The Library of Congress Subject Headings (LCSH) are drawn from the Seattle Public Library, and the Book Industry Standards and Communications (BISAC) subject codes are drawn from publishers' websites.

Most libraries do not assign subject headings from scratch but instead receive cataloging records from vendors or central services like OCLC. Publishers assign BISAC headings during production. Catalogers at each library then decide whether to keep, modify, or supplement these headings based on local policies and community needs. Smaller libraries often rely heavily on vendor-supplied metadata.

To illustrate how these systems operate in practice, Table 1 presents subject headings for three well-known fiction titles. These examples show that subject headings vary in granularity and emphasis across classification systems—and that critical themes may

be inconsistently labeled or entirely absent depending on the source, such as the subject of Islam in Salman Rushdie's *The Satanic Verses* (1988), slavery in Toni Morrison's *Beloved* (1987), or queerness in Alison Bechdel's *Fun Home* (2006). Because vendor diversity audits often rely on subject headings, their effectiveness depends on how books are cataloged and whether the metadata accurately reflects the thematic and demographic dimensions under evaluation.

4 METHODS

4.1 Interviews

We interviewed 14 staff members from public libraries across the United States, including collection development librarians, library directors, and library assistants (see more details in Appendix A.1). We conducted five initial interviews that helped inform our later survey (two in July 2023, three in late 2024–early 2025). Then, in January 2025, seven survey participants elected to participate in follow-up interviews, with two additional interviewees identified via snowball sampling.

All interviews were conducted virtually using video conferencing software and transcribed through speech-to-text software. The first, second, and fourth authors conducted an iterative thematic analysis [33] of interview transcripts using Atlas.ti. We began with an open coding phase, collaboratively identifying emergent themes. Through ongoing discussion and refinement, we consolidated tags into a smaller, structured set of themes.

4.2 Survey

We developed a survey to evaluate librarians' experiences with and perspectives on diversity audits with a focus on public libraries. We distributed the survey to relevant mailing lists and online communities for public libraries and collection management (e.g., American Library Association Connect; mailing lists for the Association of Rural and Small Libraries, Urban Libraries Council, etc.). We received 99 quality responses from public library workers; 90 respondents filled out the entire survey, including basic demographic information about their library (see more details in Appendix A.3).

Region	Respondents
Northeast	37 (45%)
West	21 (26%)
Midwest	14 (17%)
Southeast	8 (10%)
Southwest	2 (2%)

(a)

Library Pop. Size	Respondents
Fewer than 1,000	2 (2%)
1,000–4,999	2 (2%)
5,000–9,999	6 (7%)
10,000–24,999	14 (16%)
25,000–49,999	19 (21%)
50,000–99,999	17 (19%)
100,000–249,999	15 (17%)
250,000–499,999	9 (10%)
500,000 or more	6 (7%)

(b)

Table 2: (a): Public library survey respondents by region of library. (b): Public library survey respondents by library population size. (n = 90)

4.3 Purchasing Data

We obtained public library purchasing records from a platform called GovSpend[1], which tracks federal, state, and local government procurement for over 8.8 million entities, including many public libraries. GovSpend is typically used by contractors and government entities to understand the market for their services or historical contract costs, but it has also been used to study government spending more broadly [36].

4.3.1 GovSpend Search. We query the GovSpend database for records where the item price is over $20, the agency name contains "library," and the description includes the words "diversity," "DEI," or the specific product names "iCurate" and "collectionHQ." We manually review the results to exclude irrelevant items (e.g., "pencils in *diverse* colors"). This process yields 171 entries broadly related to diversity initiatives, including 35 that directly pertain to collection diversity audits.

To contextualize diversity-related purchases within broader library spending, we collect records over $100 for two large library systems within multi-year time frames (2021-2024). Library A serves a Midwestern city with a population exceeding 900K, including approximately 650K active library cardholders. Library B is located in another Midwestern city with a population of around 300K, with over 100K active cardholders.

4.4 Vendor Materials and Background Research

To better understand how automated collection audits work, we also review available vendor documentation, including product descriptions, marketing materials, webinars, and case studies. Additionally, we conduct informal background interviews with vendor representatives.

4.5 Limitations

Multiple considerations limit the extent to which our findings can be generalized. The GovSpend platform does not include purchasing records for all public libraries, and its data—collected through public records requests and web scraping—may be incomplete or inconsistently labeled. Our survey was distributed through professional networks and online communities, resulting in a self-selecting and self-reported sample. Because it was anonymous, we also cannot rule out the possibility that multiple respondents came from the same institution. Nonetheless, responses span 26 states and 55 unique state/library population size combinations, indicating substantial geographic and institutional diversity. Several interviewees were identified because they or their libraries had publicly discussed conducting diversity audits, which may bias the sample toward more engaged perspectives. Lastly, because our study was conducted before Donald Trump took office as president in 2025 and before the latest wave of anti-DEI legislation, we do not capture the most recent political attacks and institutional constraints now being imposed on libraries. Future work in this area is essential.

[1] https://govspend.com/govspend-platform/

Features	Baker & Taylor	Ingram	OverDrive
Classification Method & Sources	Automated matching on: subject headings (LCSH, BISAC), proprietary lists of books mentioned in library media sources or identified by staff; user self-classification (some crowdsourcing).	Automated matching on: subject headings (LCSH, BISAC), proprietary lists of books identified by staff (curated for 20+ years); manual staff review.	Automated matching on: subject headings (BISAC), lists of books that won diversity awards (e.g., Stonewall Book Awards, Transgender Voices).
Audit Scope	Books sold by any vendor; any language (as long as it matches classification criteria).	English-language print books sold by Ingram (19+ million titles); includes some out-of-print books.	Digital books sold by OverDrive; any language (as long as it matches classification criteria).
Recommendations	Purchasable recommended titles.	"Shopping lists" of recommended Ingram books ranked by library sales popularity.	Recommended OverDrive titles.
Speed	Monthly updated results.	2 weeks.	At least 5 days.
Price	Free with collectionHQ subscription; $6-13,500+ for standalone for first year.	$1,500 (single age group); $4,200 (all age groups); Additional fee for follow-up audit.	Free with OverDrive subscription.
Reports	Interactive dashboard with gap analysis; exportable CSVs with identified diverse books.	Spreadsheet results; Slideshow presentation; Peer (average public library) benchmarking.	Live presentation; PDF report; Supplemental materials including all BISAC headings used.

Table 3: Comparison of vendor-provided diversity audit tools. This information is drawn from public vendor documentation and background interviews with vendor staff [17, 18, 34, 63, 64, 96, 114, 120].

5 KEY FINDINGS

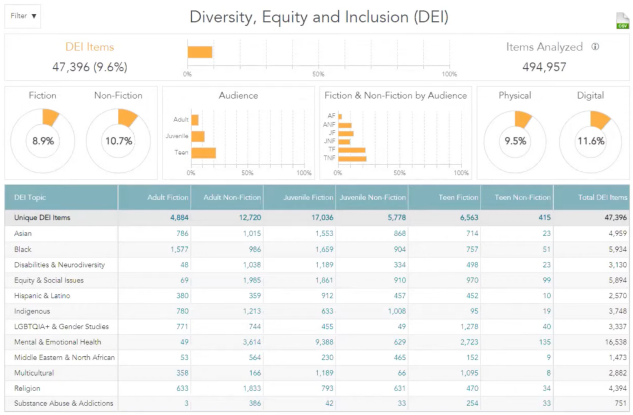
5.1 Vendor Audits Use Standardized Diversity Categories and Proprietary Metadata (RQ2)

We find that vendors operationalize diverse books by devising standardized categories, and then classifying books by automatically matching on a combination of subject headings and proprietary data sources, such as title lists curated by vendor staff (Table 3). Specific diversity categories vary slightly by vendor, but they all follow a similar structure and exhibit a degree of conceptual incongruity (Table 4). Vendors combine demographic markers (e.g., "Black", "Asian Interest"), thematic social issues (e.g., "Substance Abuse & Addictions"), and umbrella terms (e.g., "Multicultural"), resulting in an expansive but loosely organized schema. Many library workers raised concerns about the breadth and ambiguity of these frameworks (Section 5.4).

To assign books to categories, vendors typically draw on a library's catalog metadata and computationally match book identifiers—such as ISBNs or EANs[2] —against their own proprietary databases of diverse titles, and apply other automated classification rules. Because ISBNs track specific *editions*, this approach can introduce inconsistencies. One interviewee, for example, received a recommendation for a different edition of a book their library already owned.

Baker & Taylor provides examples of how subject headings are mapped to categories: a title tagged "Cooking, Middle Eastern; Arab American" would be classified as "Middle Eastern & North African," and one with "People & Places / Canada / Indigenous" would be classified as "Indigenous" [6]. While seemingly straightforward, such classifications sometimes drew criticism from respondents, who noted the recommendation of cookbooks over more substantive works and the flattening of cultural and geographic specificity (Section 5.4). Most systems allow books to appear in multiple categories.

[2]The International Standard Book Number (ISBN) is a unique identifier for books. The European Article Number (EAN) is a broader identifier used for retail items including but beyond books, and is typically encoded in barcodes. For books, the ISBN is often embedded in the EAN.



DEI Topic	Adult Fiction	Adult Non-Fiction	Juvenile Fiction	Juvenile Non-Fiction	Teen Fiction	Teen Non-Fiction	Total DEI Items
Unique DEI Items	4,884	12,720	17,036	5,778	6,563	415	47,396
Asian	786	1,015	1,553	868	714	23	4,959
Black	1,577	986	1,659	904	757	51	5,934
Disabilities & Neurodiversity	48	1,038	1,189	334	498	23	3,130
Equity & Social Issues	69	1,985	1,861	910	970	99	5,894
Hispanic & Latino	380	359	912	457	452	10	2,570
Indigenous	780	1,213	633	1,008	95	19	3,748
LGBTQIA+ & Gender Studies	771	744	455	49	1,278	40	3,337
Mental & Emotional Health	49	3,614	9,388	629	2,723	135	16,538
Middle Eastern & North African	53	564	230	465	152	9	1,473
Multicultural	358	166	1,189	66	1,095	8	2,882
Religion	633	1,833	793	631	470	34	4,394
Substance Abuse & Addictions	3	386	42	33	254	33	751

Figure 2: A dashboard featuring results from Baker & Taylor's CollectionHQ DEI Analysis tool, showing representation of books per DEI category.

Table 3 includes a full comparison of audit methods, price points, and scopes. Each audit yields a quantitative report with metrics for overall diversity and specific categories, commonly presented through pie and bar charts (Figure 2).

Vendors often claim their audits are context-sensitive and designed to avoid reinforcing stereotypes [64]. In other words, a book featuring an inaccurate, negative, or harmful portrayal of a demographic group would not be counted as "diverse." For these reasons, collectionHQ's *DEI Analysis* allows libraries to remove books from categories deemed inappropriate (and if enough do so, the book is excluded system-wide).

Yet it remains difficult to assess how such contextual sensitivity is applied at scale. For example, Sharri Markson's *What Really Happened in Wuhan* (2021), a journalistic account of COVID-19's origins, was tagged as "Asian Interest" in one vendor report. Markson is a Jewish Australian journalist, and the book's BISAC headings include: "HEALTH & FITNESS / Diseases / Contagious (incl. Pandemics)," "POLITICAL SCIENCE / Geopolitics," "SOCIAL SCIENCE / Conspiracy Theories," and "TRAVEL / Asia / East / China." While

the terms "Asian" and "China" appear in the BISAC headings, this book is conceptually distinct from the kinds of titles libraries often seek when addressing racial representation or community demographics. This example illustrates both the limitations of automated systems that rely heavily on metadata and the consequences of not making methodological choices more transparent.

Baker & Taylor Categories	Ingram Categories	Overdrive Categories
Asian & Pacific Islander (with subcategories)	Asian Interest	Asian
Black	Black Interest	African American
Disabilities & Neurodiversity	Neuro and Physical Diversity	Mental & Physical Differences
Hispanic & Latino	Latine Interest	Hispanic & Latino
Indigenous	Indigenous Interest	Indigenous
Sexuality & Gender (with subcategories)	LGBTQIA+ Interest	LGBTQ+
		Women
Mental and Emotional Health	Mental Health	Mental Health
Middle Eastern & North African	Middle Eastern Interest	Non-US Geography
		African
Multicultural	Multicultural	Multicultural Studies
Religion (with subcategories)	Muslim Interest	Religion
Equity & Social Issues	Jewish Interest	Immigration
Substance Abuse & Addictions		Urban
		Class
		Alternative Family

Table 4: Comparison of diversity audit categories by vendor. Note: OverDrive uses these categories for visualization purposes; they measure 500 BISAC subjects. Baker & Taylor subcategories include: Asian & Pacific Islander (Central Asian, East Asian, Pacific Islander, South Asian, Southeast Asian); Religion (Agnostic & Atheist, Buddhist, Hindu, Muslim, Jewish); Gender & Sexuality (Bisexual, Gay, Lesbian, Transgender).

Additionally, while OverDrive shares a report with all 500 BISAC subjects used in their analysis, the other vendors do not disclose the exact details or criteria of their approach. This gives the audits an aura of "smoke and mirrors," according to one interviewee. Library workers are often left to trust that audits are conducted sensitively and responsibly on tens of thousands or even millions of titles.

5.2 Collection Diversity Audits Are Prevalent (RQ1)

We find that overall adoption of collection diversity audits seems to be increasing. According to our survey, 67 respondents (68%) across 24 states reported that their library had conducted a collection diversity audit. Of the 32 who had not, 21 respondents (70%) said that their library will or might conduct one in the future. While our survey is limited in scope, these findings align with trends reported elsewhere. *Library Journal* found that the percentage of public libraries conducting diversity audits rose from 5.5% in 2019 to 46% in 2022 [122, 128], suggesting that the practice has become increasingly common nationwide.

In the spending data, we see collection diversity audit purchases across at least 13 states, and adjacent purchases such as seminars, webinars, and e-courses *about* diversity audits, and guidebooks *for* conducting diversity audits (A.5). Libraries spend more on other

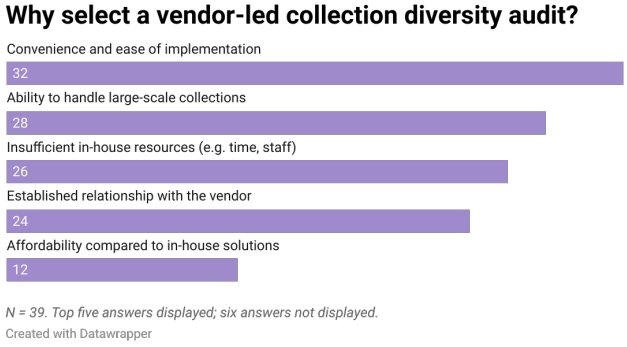


Figure 3: Reasons libraries chose to conduct a vendor-provided collection diversity audit, according to survey respondents.

diversity-related purchases (A.5)—such as staff trainings—but overall, they spend more on vendors who provide collection audits (Table 5 and 6).

5.3 Vendor Audits Increase Vendor Dependence (RQ1, RQ3)

Vendor-provided diversity audits have become a popular and often appreciated tool but we find that they contribute to deeper forms of financial and technical dependence on vendors. In our survey, 40 respondents (60%) reported that their library had used a commercial audit, either exclusively or in conjunction with internal assessments. The most commonly used vendors include Ingram (24), Baker & Taylor (18), OverDrive (12), Midwest Tape (2), and the non-profit Diverse BookFinder (4), with 15 respondents (37.5%) reporting use of more than one audit. Using multiple vendor audits in a single year is not uncommon, as is confirmed by interviewees (e.g., P5, P8, P9) and public records [77].

Twenty-four respondents (60%) said they chose a commercial audit because of an existing relationship with the vendor. Audits are sometimes offered at no additional cost as part of broader service contracts, like an OverDrive or CollectionHQ subscription. This bundling makes vendor audits especially attractive for libraries with limited staff capacity or technical expertise to conduct in-house assessments.

These "free" audits are rarely without cost, however. Purchasing records from two large Midwestern library systems show that they spent hundreds of thousands—and in some cases, millions—of dollars annually on materials and services from vendors such as Ingram, OverDrive, and Baker & Taylor (Table 5, Table 6). These audits function as a value-added feature that reinforces existing vendor relationships and incentivizes continued purchasing. In some cases, libraries conduct the same audit multiple times: Library A paid Ingram an average of $3.6K annually over three years for diversity audits.

Vendors typically recommend purchasing their own titles to fill identified diversity gaps. The audit thus serves as both evaluator and sales driver, and sales can be substantial. Most respondents (28%) estimated their library spent between $5K and $9K per audit

Company Name	2021	2022	2023	2024	GovSpend Total
OverDrive	**$3.1M**	**$108K**	**$4.2M**	**$2.5M**	**$9.9M**
Midwest Tape	$235K	$250K	$230K		$715K
Kanopy (OverDrive)	**$110K**	**$120K**	**$115K**		**$345K**
Baker & Taylor	**$89K**	**$34K**	**$39K**	**$6K**	**$168K**
ProQuest	$39K	$41K	$54K		$134K
Ingram Library Services	**$77K**	**$4K**			**$82K**
Total GovSpend	$3.9M	$839K	$4.9M	$2.7M	
Total Library Collection Spending†	$8.9M	$9.7M	$11.7M		

Table 5: Spending by a large Midwestern library (Library A) with select vendors between January 1, 2021 and August 9, 2024, according to GovSpend. Records include items over $100. Since GovSpend records are incomplete, figures represent the minimum amount spent. †Total library collection expenditures were obtained from state audit reports.

Company Name	2022	2023	2024	GovSpend Total
OverDrive	**$104K**	**$673K**	**$372K**	**$1.1M**
Midwest Tape	$81K	$560K	$292K	$932K
Ingram Library Services	**$55K**	**$31K**	**$17K**	**$103K**
GovSpend Total	$240K	$1.3M	$681K	**$2.2M**

Table 6: Spending by a large Midwestern library (Library B) with select vendors between October 11, 2022 and August 15, 2024, according to GovSpend. Records include items over $100. Since GovSpend records are incomplete, figures represent the minimum amount spent.

cycle, including new books purchased, with three (7%) reporting totals exceeding $30K. After conducting three vendor audits in 2023, the Seattle Public Library acquired 1.5K new titles, totaling nearly 4.7K print and digital copies [77].

Fifty respondents (50%) expressed that they were *somewhat* or *very* concerned that vendors might have a conflict of interest when auditing and recommending their own titles for purchase. Others were less concerned, though sometimes due to incomplete information. As one participant noted, "The vendor we used, collectionHQ, does not sell books. We use them to run reports that tell us stats about the items in our collection. So they have no reason to have a conflict of interest." In fact, collectionHQ is owned by Baker & Taylor—one of the largest book distributors in the U.S. This response points to a gap in awareness about the increasingly complex and opaque relationships among vendors, especially in an era of industry consolidation.

5.4 Vendor Audits Homogenize Diversity—And Potentially Collections (RQ2, RQ3)

Many librarians expressed frustration with the overly broad and opaque nature of the diversity categories in automated diversity audits (Table 4). These standardized labels often failed to capture local demographic nuances. Thirty-two respondents (41%) said they were *not too* or *not at all* effective at being responsive to community needs. Some librarians perceived them as reductive or even counterproductive. For example, P14, a librarian serving a Tribal Nation, shared dissatisfaction with the umbrella term "Indigenous":

> [Vendors] assume that a book about a Cherokee experience has any contact with the Mandan people in North Dakota. So that's where the diversity labeling gets into a sticky mess. The granular level is the better level because you're much more localized. But that's not how the big companies do it.

The same critique was also raised in the survey responses: "To bundle 'Indigenous' materials as a single concept is akin to bundling 'European' as a concept: inadequate and superficial." These frustrations echo scholarly critiques about the flattening of Indigenous identities and knowledge within LCSH [40]. Respondents also flagged the use of the category "Black" as overly broad—failing, for example, to distinguish between African American and African or African diasporic literatures (one respondent noted the audit neglected "a HUGE collection of African works of literature").

Interviewees wished for more customization related to specific local communities. For example, P7 conducted an in-house diversity audit to assess how well their collection was supporting their community's large and growing Bengali population, but vendor-driven audits did not allow for this level of granularity in their diversity categories.[3]

Other respondents took issue with the actual books that were recommend in these broad categories, reporting that they seemed to be superficially or even offensively related. One librarian said that they were looking for Indigenous and Middle Eastern nonfiction but ended up getting recommended mostly cookbooks. Another asked why "one of the big vendors put[s] books about basketball only into their 'Black interest' category and not other categories?"

5.5 Vendor Audits Lack Transparency—and Sometimes Trust (RQ2, RQ3)

We find that most audits lack methodological transparency, which undermines trust in results for some library workers. Interviewees consistently exhibited confusion, skepticism, or lack of knowledge about how audit results were generated. While most survey respondents thought vendor audits were *about the same*, *more*, or *much more* trustworthy than those conducted by internal staff, thirty-six (40%) answered that they were *less* or *much less* trustworthy.

These concerns were compounded by broader misgivings about metadata quality. Several librarians observed that subject headings—the core of many vendor audit systems—are often incomplete, inconsistently applied, or outdated. P7, for example, raised questions about whether metadata could accurately reflect the diversity of a book's content:

> I've worked with metadata. I know how bad it can be... a lot of the MARC [Machine-Readable Cataloging] records for books in ILS [Integrated Library System] systems or in book vendors, those subject headings are assigned by publishers who aren't necessarily librarians. And then now we're running a diversity audit on that? So what are we missing?

[3]Some audits, like Baker & Taylor's, have recently introduced more subcategories–such as "South Asian," which includes Bengali—but the granularity is still limited overall.

These metadata issues were frequently linked to a broader decline in vendor service quality following the COVID-19 pandemic. Interviewees reported that supply chain disruptions, staff turnover, and rising costs all contributed to a reduction in cataloging quality [1]. As vendor audits increasingly depend on automated classification and bibliographic metadata, the deterioration in catalog record quality raises serious questions about the reliability of these tools.

5.6 Vendor Audits Offer Convenience, Benchmarking, and Evidence (RQ3)

Despite the critiques around transparency and category design, many interviewees emphasized the convenience, time savings, and institutional utility of these tools. Seventy respondents (90%) said they were *somewhat* or *very* effective at delivering quick, scalable results. Sixty-four (82%) said they were effective at producing evidence to communicate with stakeholders. Nine interviewees similarly mentioned time saved and convenience as a major draw. "There's no way we would've had time to manually go through even a portion of our collection," P9 shared. Librarians reported that vendor-driven audits not only streamline the evaluation process, but also provide evidence that can guide collection development, demonstrate commitments to administrators, and support broader DEI initiatives.

Participants also found that comparative benchmarks—showing their results compared to "average" public library results—were useful not only for highlighting areas of improvement but for demonstrating to critics that their collections were not abnormally imbalanced. One motivation of their diversity audit, P4 said, was to prove "we're not that crazy. We're not that different than other libraries. It's not like we're... just offering wildly crazy books that no one else is offering." This sentiment has also been expressed by other library workers. "The data shows the pushback is not founded in reality," one New York librarian told journalists after a diversity audit. "Our collections aren't what they think they are" [44].

6 DISCUSSION

Our findings reveal that for many library workers, automated collection diversity audits provided by commercial vendors offer a useful and time-saving tool for addressing a real and urgent problem: the lack of diversity in library collections. As we found in Section 5.6, library workers reported that these tools deliver quick results, generate stakeholder-facing reports, and provide curated lists for acquisition—filling a gap for many libraries that lack staff capacity to conduct in-house assessments. The standardization of diversity categories also facilitates comparisons across institutions and provides evidence that libraries can present to stakeholders and the broader community—including advocacy groups bent on limiting access to diverse books.

We affirm that the lack of various kinds of diversity in library collections is an urgent problem. We believe that automated tools, when thoughtfully designed and deployed, may serve as effective mechanisms for addressing large collections that cannot be manually reviewed title-by-title. As political attacks on diversity and libraries escalate, it is important to recognize that what may seem inadequate from a distance may nonetheless offer vital support to library workers on the ground [44, 106].

However, our study surfaces several major concerns in existing vendor collection diversity audits, in the broader framing of diversity they promote, and in public libraries' deepening financial and technical dependence on conglomerate vendors. When considered in light of related work on algorithmic bias, fairness, critical race theory, and library privatization, additional concerns emerge.

First, as detailed in Sections 4.5 and 5.4, library workers expressed concerns about the homogeneity of diversity categories and the disconnect between standardized labels and the specific needs of their local communities. One respondent compared the bundling of "Indigenous" materials to labeling all "European" books under one heading, pointing to the loss of nuance, history, and identity in such taxonomies. Thirty-two respondents (41%) said they were *not too* or *not at all* effective at being responsive to community needs.

Second, as discussed in Section 5.3, vendor audits reinforce libraries' financial and infrastructural dependence on commercial providers. Public libraries fundamentally rely on commercial vendors for a wide range of services, but these audits may deepen vendor lock-in and disincentivize libraries from developing their own automated or manual assessment tools, or from investing in human time and labor. As libraries cede more authority over content, analytics, and infrastructure to commercial providers, they may lose the capacity—and bargaining power—to evaluate and shape their own collections independently.

Third, market competition discourages vendors from fully disclosing how their tools work, limiting transparency and contributing to confusion or mistrust among library workers (Section 5.5). These concerns were compounded by broader skepticism about the quality and provenance of bibliographic metadata. In total, 36 (40%) respondents said they trusted vendor-conducted audits *less* or *much less* than audits conducted by internal staff.

Notably, many of these critiques are not new. For example, in the late 1990s, librarians in Hawai'i mounted a grassroots campaign against Baker & Taylor after the state librarian signed a large outsourcing contract for collection development. Opponents argued that the vendor's privatized, centralized approach failed to meet the needs of Native Hawaiian communities [70, 125]. As one librarian warned, "It's like putting Safeway in charge of the school lunch program. The kids will be fed what makes a profit" [48, 58, 70, 125]. Today's critiques mirror these earlier tensions. Automated diversity audits may ultimately extend rather than resolve the longstanding problem of vendor-driven and profit-driven library collection standardization.

These findings also resonate with broader critiques about the operationalization of race and the cooption of diversity. As Hanna et al. [57] argue, "treating race as an attribute, rather than a structural, institutional, and relational phenomenon… minimize[s] the structural aspects" of inequality. In the context of libraries, these structural forces include systemic discrimination in the publishing industry, the commercial consolidation of the book industry, and the chronic underfunding of public institutions. In a similar vein, Melamed [90] argues that frameworks of diversity were coopted and commodified after World War II, and that they mask and enable ongoing racial exploitation; she even specifically highlights diverse literature as a central instrument in this process. These perspectives suggest that diversity audits risk turning numerical targets into an end in themselves—pursued without regard for the quality,

relevance, or political significance of the selected materials, and without prompting meaningful changes to collection development practices or engagement with community knowledge

Moreover, the widespread adoption of automated audits risks hollowing out the critical, context-specific labor that collection development traditionally entails. Without careful intervention, libraries may find themselves increasingly dependent on opaque external systems, alienated from the intellectual and social responsibilities at the heart of their mission.

6.1 Recommendations

Automated collection diversity audits represent just one among many data-driven, algorithmic, and AI tools now being introduced to libraries. Based on our findings, we present the following recommendations for automated diversity audits, emphasizing that many of these principles extend to libraries' adoption of automated and AI systems more broadly.

- **Enhance Transparency in Automated Vendor Tools**: To improve trust and understanding, vendors should fully disclose classification criteria, data sources, and results. As shown in Section 5.5, library workers often do not know exactly how audit classifications are made. More than a third of respondents also found them less trustworthy than those performed by library staff. As AI systems become more embedded in library workflows, opaque infrastructures will only deepen unless transparency becomes a collective demand.
- **Promote Flexibility and Contextual Awareness**: Baker & Taylor's tool currently offers the most customization, but vendor tools should allow more customization and crowdsourcing. Section 5.4 shows that libraries want tools that reflect local demographic complexity.
- **Invest in Staffing and Manual Review Processes**: Libraries should secure funding and create positions for manual assessment, framing this labor as core—not supplemental—to ethical data practices. Throughout our interviews, librarians consistently emphasized the importance of human judgment and review.
- **Support Open Alternatives and Build Internal Data Capacity**: As discussed in Sections 5.3, vendor audits further cement library relationships with vendors. Reducing dependence on proprietary systems requires investment in free, open-source tools and broader data capacity. For example, the Institute of Museum and Library Services (IMLS) has funded tools like Diverse BookFinder, which supports audits of diversity in picture books. Graduate programs in library science might also incorporate more training in data literacy and computational methods.
- **Organize Collectively to Influence Vendor Practices**: Because most libraries lack bargaining power on their own, regional consortia, professional associations, and advocacy organizations—such as the American Library Association (ALA), SPARC, and Library Futures—should collaborate to demand greater transparency, flexibility, and accountability in automated vendor tools. These groups have already laid important groundwork by promoting open infrastructures, equitable access, and public-interest technology policy.

6.2 The Future of Collection Diversity Audits—And Public Libraries

The future of diversity audits remains uncertain as a wave of reactionary anti-DEI and anti-library action has overtaken institutions across the U.S. In early 2025, the Trump administration made drastic cuts to the Institute of Museum and Library Services (IMLS)—a key source of federal funding for libraries, programs, and innovation nationwide—as well as to several state libraries [59, 79]. In May, Trump fired Carla Hayden, the first Black Librarian of Congress, and his administration cited her "pursuit of DEI" and supposed placement of "inappropriate books in the library for children" as justification [80, 93].

Even before Trump's second election, conservative state governments had forced libraries to withdraw their membership from ALA [15], banned DEI initiatives [38], and threatened librarians with $5,000 or jail time for "distribution of harmful materials to minors" [107]. This is among 128 other bills levied against libraries in 2024, with almost as many levied five months into 2025 [46, 47]. Additionally, Moms for Liberty, a nonprofit with over 130,000 volunteer members in 49 states, continues to organize book challenges and bans [50]. Lastly, after supporting DEI initiatives in the wake of 2020, many mainstream publishers have backpedaled, and several prominent Black women in the industry have been fired, causing outcry and calling into question the availability of diverse books moving forward [12, 110].

Whether public libraries will be able to continue conducting collection diversity audits, at least under this name, is unclear. We speculate that budget cuts may heighten libraries' need to outsource labor to automated tools, and that anti-DEI legislation will hamper efforts to improve collection diversity.

7 CONCLUSION

We explore the benefits and limitations of automated library collection diversity audits by drawing on a survey, interviews, purchasing records, and an analysis of vendor documentation. We find that these audits arise from a real need: many library systems are so large and resource-constrained that some form of computational assistance is necessary to assess the inclusivity of their collections. Vendor-led audits offer a convenient but homogenizing solution—one that risks flattening the complexities of identity into standardized categories while deepening libraries' infrastructural dependence on opaque, commercial systems. As algorithmic tools become more embedded in public knowledge institutions, we advocate for increased transparency and flexibility, greater investment in open-source tools and human expertise, and collective organizing.

This work is especially urgent in a political moment marked by renewed attacks on libraries, education, and marginalized communities. In the context of rising authoritarianism and right-wing censorship efforts, we need genuine engagement with and protection of diversity. We must support libraries in resisting tech cooption and in ensuring automation serves the public good.

Acknowledgments

We thank the library workers who took the time to share their experiences and perspectives with us. We are grateful for feedback from our anonymous reviewers and from Lorcan Dempsey, Anna Preus, and others. This paper also benefited from conversations with UCLA's Critical Data Lab in Fall 2024. This research was partly supported by a Strategic Research Fund award from the Information School at the University of Washington.

References

[1] 2021. Ingram Warns of 4th-Quarter Supply Chain Disruptions. *Shelf Awareness Newsletter* (24 Aug. 2021). https://www.shelf-awareness.com/issue.html?issue=4056#m53553

[2] 2022. *Public Libraries Survey.* Technical Report. Institute of Museum and Library Services. https://www.imls.gov/research-evaluation/data-collection/public-libraries-survey

[3] OCLC 2023. *OCLC Introduces AI-generated Book Recommendations in WorldCat.Org and WorldCat Find Beta.* OCLC. https://www.oclc.org/en/news/releases/2023/20230621-ai-book-recs-worldcatorg.html

[4] 2024. 2024 Preliminary Book Ban Data. *American Library Association's Office of Intellectual Freedom* (March 2024). https://www.ala.org/bbooks/book-ban-data

[5] Department for Professional Employees, AFL-CIO 2024. *Library Professionals: Facts, Figures, and Union Membership.* Department for Professional Employees, AFL-CIO. https://web.archive.org/web/20250429180949/https://www.dpeaflcio.org/factsheets/library-professionals-facts-and-figures

[6] Baker and Taylor 2025. *Diversity, Equity & Inclusion Resources.* Baker and Taylor. https://web.archive.org/web/20250429145154/https://www.baker-taylor.com/library-solutions/books-av-content/dei-sel

[7] Ex Libris Knowledge Center 2025. *Getting Started with Primo Research Assistant.* Ex Libris Knowledge Center. https://web.archive.org/web/20250429180155/https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/020Primo_VE/Primo_VE_(English)/015_Getting_Started_with_Primo_Research_Assistant

[8] Amina A. Abdu, Irene V. Pasquetto, and Abigail Z. Jacobs. 2023. An Empirical Analysis of Racial Categories in the Algorithmic Fairness Literature. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT '23).* Association for Computing Machinery, New York, NY, USA, 1324–1333. https://doi.org/10.1145/3593013.3594083

[9] Char Adams and Nigel Chiwaya. 2024. Map: See Which States Have Introduced or Passed Anti-DEI Bills. *NBC News* (March 2024). https://www.nbcnews.com/data-graphics/anti-dei-bills-states-republican-lawmakers-map-rcna140756

[10] Sara Ahmed. 2012. *On Being Included: Racism and Diversity in Institutional Life.* Duke University Press, Durham.

[11] Kira Alfonseca. 2024. Map: The Impact of Anti-DEI Legislation. *ABC News* (April 2024). https://abcnews.go.com/US/map-impact-anti-dei-legislation/story?id=108795967

[12] Alexandra Alter and Elizabeth Harris. 2024. 'A Lot of Us Are Gone': How the Push to Diversify Publishing Fell Short. *The New York Times* (Aug. 2024).

[13] Alexandra Alter and Elizabeth A. Harris. 2024. Publishing Pledged to Diversify. Change Has Been Slow. (2024). https://www.nytimes.com/2024/02/28/books/publishing-books-poc-dei.html Accessed: 2025-01-22.

[14] PEN America. 2024. Banned in the USA: Beyond the Shelves. https://pen.org/report/beyond-the-shelves/.

[15] Andrew Atterbury. 2023. Florida joins conservative states severing ties with national library group. (31 Oct. 2023). https://www.politico.com/news/2023/10/31/florida-conservative-national-library-ala-00124516

[16] Henriette D. Avram. 1975. *MARC; Its History and Implications.* Technical Report. Superintendent of Documents, U.

[17] Baker & Taylor. 2022. Introducing Baker & Taylor's Diversity Analysis Tool. https://www.youtube.com/watch?v=ldybXE2dozM

[18] Baker and Taylor. 2025. Baker & Taylor | Diversity, Equity & Inclusive | Social and Emotional Learning. https://www.baker-taylor.com/library-solutions/books-av-content/dei-sel.

[19] Jill Baron and Sawyer Broadley. 2019. Change the Subject. https://collections.dartmouth.edu/archive/object/change-subject/change-subject-film

[20] Ruha Benjamin. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code.* Polity, Cambridge, UK Medford, MA.

[21] Sanford Berman (Ed.). 2013. *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People* (reprint edition ed.). McFarland & Company, Jefferson, N.C.

[22] Amber Billey, Elizabeth Nelson, and Rebecca Uhl (Eds.). 2024. *Inclusive Cataloging: Histories, Context, and Reparative Approaches.* ALA Editions, Chicago.

[23] Rudine Sims Bishop. 1990. Mirrors, Windows, and Sliding Glass Doors. *Perspectives: Choosing and Using Books for the Classroom* 6, 3 (1990). https://scenicregional.org/wp-content/uploads/2017/08/Mirrors-Windows-and-Sliding-Glass-Doors.pdf

[24] Glencora Borradaile, Brett Burkhardt, and Alexandria LeClerc. [n. d.]. Whose Tweets Are Surveilled for the Police: An Audit of a Social-Media Monitoring Tool via Log Files. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, 2020-01-27) *(FAT* '20).* Association for Computing Machinery, 570–580. https://doi.org/10.1145/3351095.3372841

[25] Marhsall Breeding. 2019. *OverDrive's New Owners: What It Means.* American Libraries Magazine. https://americanlibrariesmagazine.org/blogs/the-scoop/overdrives-new-owners-what-means/ Accessed: 2025-01-22.

[26] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency.* PMLR, 77–91.

[27] US Census Bureau. 2021. 2020 Census Illuminates Racial and Ethnic Composition of the Country. https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html.

[28] Otis A Chadley. 1992. Addressing cultural diversity in academic and research libraries. *College & research libraries* 53, 3 (1992), 206–214. https://doi.org/10.5860/crl_53_03_206

[29] Kent K. Chang. 2023. The Queer Gap in Cultural Analytics. *Debates in the Digital Humanities 2023* (2023), 105–119.

[30] Kent K. Chang, Anna Ho, and David Bamman. 2024. Subversive Characters and Stereotyping Readers: Characterizing Queer Relationalities with Dialogue-Based Relation Extraction. In *Computational Humanities Research.* arXiv. https://doi.org/10.48550/arXiv.2410.14978 arXiv:2410.14978 [cs]

[31] Wendy Hui Kyong Chun. 2009. Introduction: Race and/as Technology; or, How to Do Things to Race. *Camera Obscura: Feminism, Culture, and Media Studies* 24, 1 (70) (May 2009), 7–35. https://doi.org/10.1215/02705346-2008-013

[32] Matthew P Ciszek and Courtney L Young. 2010. Diversity collection assessment in large academic libraries. *Collection Building* 29, 4 (2010), 154–161. https://doi.org/10.1108/01604951011088899

[33] Victoria Clarke and Virginia Braun. 2017. Thematic Analysis. *The Journal of Positive Psychology* 12, 3 (May 2017), 297–298. https://doi.org/10.1080/17439760.2016.1262613

[34] CollectionHQ. 2025. Diversity, Equity and Inclusion Analysis. https://www.collectionhq.com/diversity-analysis-2/.

[35] Tim Collins. 2012. The Current Budget Environment and Its Impact on Libraries, Publishers and Vendors. *Journal of Library Administration* 52, 1 (Jan. 2012), 18–35. https://doi.org/10.1080/01930826.2012.630643

[36] Jack Corrigan, Sergio Fontanez, and Michael Kratsios. 2022. Banned in DC: Examining Government Approaches to Foreign Technology Threats. (Oct 2022). https://doi.org/10.51593/20220007

[37] Sasha Costanza-Chock, Inioluwa Deborah Raji, and Joy Buolamwini. 2022. Who Audits the Auditors? Recommendations from a Field Scan of the Algorithmic Auditing Ecosystem. In *2022 ACM Conference on Fairness, Accountability, and Transparency.* ACM, Seoul Republic of Korea, 1571–1583. https://doi.org/10.1145/3531146.3533213

[38] Jaclyn Diaz. 2023. Florida Gov. Ron DeSantis signs a bill banning DEI initiatives in public colleges. *NPR* (15 May 2023).

[39] Emily Drabinski. 2013. Queering the Catalog: Queer Theory and the Politics of Correction. *The Library Quarterly* 83, 2 (April 2013), 94–111. https://doi.org/10.1086/669547

[40] Marisa Elena Duarte and Miranda Belarde-Lewis. 2015. Imagining: Creating Spaces for Indigenous Ontologies. *Cataloging & Classification Quarterly* 53, 5-6 (July 2015), 677–702. https://doi.org/10.1080/01639374.2015.1018396

[41] María Evelia Emerson and Lauryn Grace Lehman. 2022. Who Are We Missing? Conducting a Diversity Audit in a Liberal Arts College Library. *The Journal of Academic Librarianship* 48, 3 (2022), 102517. https://doi.org/10.1016/j.acalib.2022.102517

[42] Matt Enis. 2014. Baker & Taylor, collectionHQ Launch ESP, a Predictive Collection Development Tool. https://www.libraryjournal.com/story/baker-taylor-collectionhq-launch-esp-a-predictive-collection-development-tool. https://www.libraryjournal.com/story/baker-taylor-collectionhq-launch-esp-a-predictive-collection-development-tool

[43] Matt Enis. 2024. AI on the Horizon. *Library Journal* (Oct 2024). https://www.libraryjournal.com/story/ai-on-the-horizon Accessed: 2025-01-21.

[44] Shayla Escudero. 2023. How Albany Is Adding Diversity to the Shelves at Its Library. (2023). https://democratherald.com/news/local/article_d3b2735e-deed-11ed-9cd7-c3306e7d1d76.html

[45] Virginia Eubanks. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor.* St. Martin's Press, New York, NY.

[46] EveryLibrary. 2024. Legislation of Concern in 2024. https://www.everylibrary.org/billtracking2024

[47] EveryLibrary. 2025. Legislation of Concern. https://www.everylibrary.org/billtracking

[48] Roxanne Farmanfarmaian. 1997. Hawaii libraries vs. Baker & Taylor: better times ahead? *Publishers Weekly* 244, 9 (1997), 16–17.

[49] Rachel Fischer. 2023. Automating the Diversity Audit Process. *Information Technology and Libraries* 42, 3 (Sept. 2023). https://doi.org/10.5860/ital.v42i3.16925

[50] Moms for Liberty. 2024. Who We Are. https://www.momsforliberty.org/about/

[51] Daniel Freund and Wentao Weng. 2024. The Dedicated Docket in U.S. Immigration Courts: An Analysis of Fairness and Efficiency Properties. In *Proceedings of the 25th ACM Conference on Economics and Computation* (New York, NY, USA, 2024-12-17) *(EC '24)*. Association for Computing Machinery, 346. https://doi.org/10.1145/3670865.3673541

[52] Christina Fuller-Gregory. 2022. DEI Audits: The Whole Picture | Equity. https://www.libraryjournal.com/story/DEI-Audits-The-Whole-Picture-Equity. Accessed: 2025-01-14.

[53] Kim Gallon. 2016. Making a Case for the Black Digital Humanities. In *Debates in the Digital Humanities 2016*, Matthew K. Gold and Lauren F. Klein (Eds.). University of Minnesota Press. https://doi.org/10.5749/j.ctt1cn6thb jstor:10.5749/j.ctt1cn6thb

[54] Dustin Gardiner. 2025. Judge Blocks Trump Admin from Pulling San Francisco Housing Funds over Anti-DEI Rules. https://www.politico.com/news/2025/05/08/trump-san-francisco-housing-dei-00336454.

[55] Anitra Gates, Amberlee McGaughey, Celia Mulder, and Sarah Voels. 2022. Using Data From Collection Diversity Audits. *Computers in Libraries* 42, 9 (Nov. 2022).

[56] Lara Groves, Jacob Metcalf, Alayna Kennedy, Briana Vecchione, and Andrew Strait. 2024. Auditing Work: Exploring the New York City Algorithmic Bias Audit Regime. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*. Association for Computing Machinery, New York, NY, USA, 1107–1120. https://doi.org/10.1145/3630106.3658959

[57] Alex Hanna, Emily Denton, Andrew Smart, and Jamila Smith-Loud. [n. d.]. Towards a Critical Race Methodology in Algorithmic Fairness. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, 2020-01-27) *(FAT* '20)*. Association for Computing Machinery, 501–512. https://doi.org/10.1145/3351095.3372826

[58] Barbara Hoffert. 2007. Who's selecting now? As Phoenix Public Library boldly passes on selection responsibilities to its vendors, some libraries follow–and others dig in. *Library Journal* 132, 14 (2007), 40–44.

[59] The White House. 2025. Continuing the Reduction of the Federal Bureaucracy. (14 March 2025). https://www.whitehouse.gov/presidential-actions/2025/03/continuing-the-reduction-of-the-federal-bureaucracy/

[60] The White House. 2025. Ending Radical And Wasteful Government DEI Programs And Preferencing. https://www.whitehouse.gov/presidential-actions/2025/01/ending-radical-and-wasteful-government-dei-programs-and-preferencing/.

[61] Han-Yin Huang and Cynthia C. S. Liem. 2022. Social Inclusion in Curated Contexts: Insights from Museum Practices. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT '22)*. Association for Computing Machinery, New York, NY, USA, 300–309. https://doi.org/10.1145/3531146.3533095

[62] Ingram. 2021. *Ingram Announces A New Service to Help Libraries Effectively Diversify Their Print Collection.* https://web.archive.org/web/20250429195151/https://www.ingramcontent.com/news/ingram-content-group-announces-a-new-service-to-help-libraries-effectively-diversify-their-print-col

[63] Ingram. 2025. iCurate. https://www.ingramcontent.com/libraries/collection-development/icurate.

[64] Ingram. 2025. *iCurate inClusive for Schools Frequently Asked Questions.* https://web.archive.org/web/20250429171336/https://www.ingramcontent.com/libraries-page/icurate-inclusive-schools-faqs

[65] Karen Jensen. 2018. *Diversity Auditing 101: How to Evaluate Your Collection.* School Library Journal. https://www.slj.com/story/diversity-auditing-101-how-to-evaluate-collection

[66] Karen Jensen and MLS 8 comments. 2017. *Doing a YA Collection Diversity Audit: Understanding Your Local Community (Part 1).* Teen Librarian Toolbox. https://teenlibrariantoolbox.com/2017/11/01/doing-a-diversity-audit-understanding-your-local-community/

[67] Eun Seo Jo and Timnit Gebru. 2020. Lessons from Archives: Strategies for Collecting Sociocultural Data in Machine Learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 306–316. https://doi.org/10.1145/3351095.3372829

[68] Rebecca Kemp. 2008. MARC Record Services: A Comparative Study of Library Practices and Perceptions. *The Serials Librarian* 55, 3 (Sept. 2008), 379–410. https://doi.org/10.1080/0361526080205636 7

[69] Frederick G. Kilgour. 1970. History of Library Computerization. *Information Technology and Libraries* 3, 3 (Sept. 1970), 218–229. https://doi.org/10.6017/ital.v3i3.5256

[70] Rebecca Knuth and Donna G Bair-Mundy. 1998. Revolt over outsourcing: Hawaii's librarians speak out about contracted selection. *Collection management* 23, 1-2 (1998), 81–112. https://doi.org/10.1300/J105v23n01_04

[71] Cataloging Lab. 2025. Problem LCSH.

[72] Sarah Lamdan. 2022. *Data Cartels: The Companies That Control and Monopolize Our Information.* Stanford University Press.

[73] Nancy Larrick. 1965. The All-White World of Children's Books. 48, 11 (1965), 63–65.

[74] Charles R. III Lawrence. 1997. Each Other's Harvest: Diversity's Deeper Meaning. *University of San Francisco Law Review* (1997). http://hdl.handle.net/10125/65961

[75] Nancy Leong. 2025. Diversity Messaging After Affirmative Action. *Minnesota Law Review* (Feb. 2025).

[76] Canadian School Libraries. 2025. *Diversity Audits.* https://www.canadianschoollibraries.ca/collection-diversity-toolkit/diversity-audits/

[77] The Seattle Public Library. 2023. Board of Trustees Meeting. https://www.spl.org/Seattle-Public-Library/documents/about-us/board-meetings/2023%20library%20board%20meeting%20materials/2023%20July%2027%20Library%20Board%20Meeting%20Packet.pdf

[78] libraryIQ. 2025. Diversity. https://www.libraryiq.com/diversity.

[79] Andrew Limbong. 2025. Entire staff at federal agency that funds libraries and museums put on leave. *National Public Radio* (31 March 2025). https://www.npr.org/2025/03/31/nx-s1-5334415/doge-institute-of-museum-and-library-services

[80] Andrew Limbong. 2025. Librarian of Congress Firing Is Latest Move in Upheaval of U.S. Cultural Institutions. *NPR* (May 2025). https://www.npr.org/2025/05/09/nx-s1-5393737/carla-hayden-fired-library-of-congress-trump

[81] Sandra Littletree and Cheryl A. and Metoyer. 2015. Knowledge Organization from an Indigenous Perspective: The Mashantucket Pequot Thesaurus of American Indian Terminology Project. *Cataloging & Classification Quarterly* 53, 5-6 (July 2015), 640–657. https://doi.org/10.1080/01639374.2015.1010113

[82] Lee & Low. 2023. Diversity Baseline Survey 3.0. https://www.leeandlow.com/about/diversity-baseline-survey/dbs3/

[83] Mackin. 2025. Tag Report: Analyze & Diversify Your Library Collection. https://web.archive.org/web/20250429203910/https://home.mackin.com/library/collection-management/tag-report/

[84] Yaaseen Mahomed, Charlie M. Crawford, Sanjana Gautam, Sorelle A. Friedler, and Danaë Metaxa. 2024. Auditing GPT's Content Moderation Guardrails: Can ChatGPT Write Your Favorite TV Show?. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, 2024-06-05) *(FAccT '24)*. Association for Computing Machinery, 660–686. https://doi.org/10.1145/3630106.3658932

[85] Elissa Malespina. 2024. AI for the Library: Readers. *School Library Journal* (July 2024). https://www.slj.com/story/ai-for-the-library-readers-advisory-tasks

[86] Laura Mandell. 2019. Gender and Cultural Analytics: Finding or Making Stereotypes? *Debates in the Digital Humanities 2019* (2019), 3–26.

[87] Kristin E. Martin and Kavita Mundle. 2010. Cataloging E-Books and Vendor Records. *Library Resources & Technical Services* 54, 4 (2010), 227–237. https://doi.org/10.5860/lrts.54n4.227

[88] Joseph R. Matthews. 1982. The Automated Circulation Marketplace: Active and Heating Up. *Library Journal* 107, 3 (Feb. 1982), 233–35.

[89] Tara McPherson. 2012. Why Are the Digital Humanities So White? Or Thinking the Histories of Race and Computation. In *Debates in the Digital Humanities*, Matthew K. Gold and Lauren F. Klein (Eds.). University of Minnesota Press, 139–160. jstor:10.5749/j.ctttv8hq.12

[90] Jodi Melamed. 2011. *Represent and Destroy: Rationalizing Violence in the New Racial Capitalism.* Univ Of Minnesota Press.

[91] Jodi Melamed. 2020. Diversity. In *Keywords for American Cultural Studies*. New York University Press, 93–97. https://doi.org/10.18574/nyu/9781479867455.003.0027

[92] Annabelle Mortensen. 2019. Measuring diversity in the collection. *Library Journal* 144, 4 (2019), 28–30. https://www.libraryjournal.com/story/Measuring-Diversity-in-the-Collection

[93] Sophia Nguyen and Herb Scribner. 2025. Trump Fires Library of Congress Chief Carla Hayden, Citing DEI. *The Washington Post* (May 2025).

[94] Safiya Umoja Noble. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism.* NYU Press, New York.

[95] Ziad Obermeyer and Sendhil Mullainathan. [n. d.]. Dissecting Racial Bias in an Algorithm That Guides Health Decisions for 70 Million People. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, 2019-01-29) *(FAT* '19)*. Association for Computing Machinery, 89. https://doi.org/10.1145/3287560.3287593

[96] OverDrive. 2025. Diversity Audits - OverDrive Resource Center. https://resources.overdrive.com/library/libby-features/diversity-audit/.

[97] Jordan Pedersen. 2022. Measuring Collection Diversity via Exploratory Analysis of Collection Metadata. *The Serials Librarian* 82, 1-4 (May 2022), 186–193. https://doi.org/10.1080/0361526X.2022.2028499

[98] Brooke Perreault, Johanna Hoonsun Lee, Ropafadzo Shava, and Eni Mustafaraj. 2024. Algorithmic Misjudgement in Google Search Results: Evidence from Auditing the US Online Electoral Information Environment. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, 2024-06-05) *(FAccT '24)*. Association for Computing Machinery, 433–443. https://doi.org/10.1145/3630106.3658916

[99] Gary Price. 2023. *New Beta Test From OCLC: AI-Generated Book Recommendations in WorldCat.Org and WorldCat Find.* Library Journal infoDOCKET. https://www.infodocket.com/2023/06/21/new-beta-test-from-oclc-ai-generated-book-recommendations-in-worldcat-org-and-worldcat-find/

[100] Evani Radiya-Dixit and Gina Neff. 2023. A Sociotechnical Audit: Assessing Police Use of Facial Recognition. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, 2023-06-12) *(FAccT '23)*. Association for Computing Machinery, 1334–1346. https://doi.org/10.1145/3593013.3594084

[101] Manish Raghavan, Solon Barocas, Jon Kleinberg, and Karen Levy. 2020. Mitigating bias in algorithmic hiring: evaluating claims and practices. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain) *(FAT* '20)*. Association for Computing Machinery, New York, NY, USA, 469–481. https://doi.org/10.1145/3351095.3372828

[102] Sharesly Rodriguez and Christina Mune. 2022. Uncoding Library Chatbots: Deploying a New Virtual Reference Tool at the San Jose State University Library. *Reference Services Review* 50, 3-4 (Nov. 2022), 392–405. https://doi.org/10.1108/RSR-05-2022-0020

[103] Sharesly Rodriguez and Christina Mune. 2025. LIBRARY CHATBOTS: Easier Than You Think. *Computers in Libraries* 41, 8 (May 2025). https://www.infotoday.com/cilmag/oct21/Rodriguez-Mune--Library-Chatbots-Easier-Than-You-Think.shtml

[104] James Rush. 1988. The Library Automation Market: Why Do Vendors Fail? A History of Vendors and Their Characteristics. *Library Hi Tech* 6, 3 (March 1988), 7–33.

[105] James E. Rush. 1982. Library Automation Systems and Networks. In *Advances in Computers*, Marshall C. Yovits (Ed.). Vol. 21. Elsevier, 333–422. https://doi.org/10.1016/S0065-2458(08)60572-0

[106] Zachary Schermele. 2022. A Cultural Power Struggle at an Iowa Library Casts a 'dark Cloud' over a Small Town. *NBC News* (Aug. 2022). https://www.nbcnews.com/nbc-out/out-news/small-town-library-shut-say-culture-wars-closed-rcna39816

[107] The Georgia State Senate. 2024. Senate Bill 154. https://www.legis.ga.gov/api/legislation/document/20232024/214457

[108] Mark Sendak, Madeleine Clare Elish, Michael Gao, Joseph Futoma, William Ratliff, Marshall Nichols, Armando Bedoya, Suresh Balu, and Cara O'Brien. 2020. "The Human Body Is a Black Box": Supporting Clinical Decision-Making with Deep Learning. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, 2020-01-27) *(FAT* '20)*. Association for Computing Machinery, 99–109. https://doi.org/10.1145/3351095.3372827

[109] Christi Sevits and Alex D McAllister. 2024. Developing an Introductory Diversity Audit at Liberty Public Library: An Action Research Approach. *Public Library Quarterly* 43, 3 (2024), 286–306.

[110] Dan Sinykin and Richard Jean So. 2024. Has the DEI Backlash Come for Publishing? https://www.theatlantic.com/books/archive/2024/06/diversity-publishing-backlash-study/678734/

[111] Carrie Smith. 2023. Building Diverse Collections. https://americanlibrariesmagazine.org/2023/06/01/building-diverse-collections/. *American Libraries Magazine* (June 2023).

[112] Richard Jean So and Edwin Roland. 2020. Race and Distant Reading. *PMLA* 135, 1 (Jan. 2020), 59–73. https://doi.org/10.1632/pmla.2020.135.1.59

[113] Richard Jean So and Gus Wezerek. 2020. Just How White Is the Book Industry? *The New York Times* (Dec. 2020). https://www.nytimes.com/interactive/2020/12/11/opinion/culture/diversity-publishing-industry.html

[114] OverDrive Staff. 2022. Gain Greater Insight into Your Collection with an OverDrive Diversity Audit. https://web.archive.org/web/20250429201829/https://company.overdrive.com/2022/06/14/gain-greater-insight-into-your-collection-with-an-overdrive-diversity-audit/

[115] RILINK Staff. 2022. Using the Follett Titlewave Diversity Report: Session Overview. https://guides.rilink.org/diversityreport/overview.

[116] Jon Swaine and Jeremy B. Merrill. 2025. Amid Anti-DEI Push, National Park Service Rewrites History of Underground Railroad. *The Washington Post* (April 2025). https://www.washingtonpost.com/investigations/2025/04/06/national-park-service-underground-railroad-history-slavery/

[117] Fiona Syme. 2021. *Available Now: DEI Analysis.* collectionHQ. https://web.archive.org/web/20241203103844/https://www.collectionhq.com/available-now-dei-analysis/

[118] Nick Szydlowski. 2025), note = Accessed: 2025-01-21, abstract = Dr. Martin Luther King, Jr. Library: Kingbot: Home, langid = english. *Dr. Martin Luther King, Jr. Library: Kingbot: Home.* https://library.sjsu.edu/c.php?g=1425539&p=10573855

[119] Midwest Tape. 2022. Midwest Tape Black Stories November 2022. https://issuu.com/midwesttape/docs/mwt_nov2022_blackstories_digital/8

[120] Baker & Taylor. 2025. Collective Purchasing Agreements - Detailed Information. https://www.olservice.ca/index.php/collective-purchasing-agreements-detailed-information?product=110.

[121] Petros Terzis, Michael Veale, and Noëlle Gaumann. 2024. Law and the Emerging Political Economy of Algorithmic Audits. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (New York, NY, USA, 2024-06-05) *(FAccT '24)*. Association for Computing Machinery, 1255–1267. https://doi.org/10.1145/3630106.3658970

[122] Christina Vercelletto. 2019. *How Diverse Are Our Books?* Library Journal. https://www.libraryjournal.com/story/How-Diverse-Are-Our-Books

[123] Sarah Voels. 2022. *Auditing Diversity in Library Collections.* Libraries Unlimited, an imprint of ABC-CLIO, LLC, Santa Barbara, California.

[124] Sanne Vrijenhoek, Savvina Daniil, Jorden Sandel, and Laura Hollink. 2024. Diversity of What? On the Different Conceptualizations of Diversity in Recommender Systems. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*. Association for Computing Machinery, New York, NY, USA, 573–584. https://doi.org/10.1145/3630106.3658926

[125] Patricia D Wallace. 1997. Outsourcing book selection in public and school libraries. *Collection building* 16, 4 (1997), 160–166. https://doi.org/10.1108/01604959710187679

[126] William H. Walters. 2023. Assessing Diversity in Academic Library Book Collections: Diversity Audit Principles and Methods. *Open Information Science* 7, 1 (Jan. 2023). https://doi.org/10.1515/opis-2022-0148

[127] Brian Watson. 2019. Homosaurus and The Digital Transgender Archive. *The American Archivist Reviews Portal* (June 2019).

[128] Neal Wyatt. 2022. *Collection Rebalance | 2022 Materials Survey.* Library Journal. https://www.libraryjournal.com/story/Collection-Rebalance-2022-Materials-Survey Accessed: 2025-01-22.

[129] Katrina Zwaaf. 2020. The Homosaurus http://homosaurus.org. *Technical Services Quarterly* 37, 2 (April 2020), 207–208. https://doi.org/10.1080/07317131.2020.1728140

A APPENDIX

A.1 Interviews

We interviewed 14 staff members from public libraries across the United States, including collection development librarians, library directors, and library assistants. We conducted five initial interviews that helped inform our later survey (two in July 2023, three in late 2024–early 2025). Then, in January 2025, seven survey participants elected to participate in follow-up interview. A recruitment email with a booking link to schedule a virtual meeting over a video conferencing software was sent to these survey respondents. Two interviewees were recruited via snowball sampling.

ID	US State or Tribal Nation	Community Type	Diversity Audit Tool(s)
P1	Alabama	Town, Fringe	In-house
P2	Georgia	City, Mid-size	Ingram
P3	Illinois	Suburb, Large	Ingram, In-house
P4	Indiana	City, Small	Baker & Taylor (collectionHQ), In-house
P5	Louisiana	City, Large	Baker & Taylor (collectionHQ), Ingram
P6	Massachusetts	Suburb, Large	In-house
P7	Massachusetts	Suburb, Large	In-house
P8	Michigan	Suburb, Large	Baker & Taylor (collectionHQ), Ingram, In-house, Overdrive
P9	Missouri	City, Mid-size	Baker & Taylor (collectionHQ), In-house, Overdrive
P10	North Carolina	City, Large	Baker & Taylor (collectionHQ), In-house
P11	Oregon	Rural, Remote	Ingram
P12	Texas	City, Large; Rural, Distant	In-house
P13	Washington	Rural, Remote	Has not conducted a diversity audit
P14	Native American Reservation	Rural, Remote	Follett

Table 7: List of interviewees and participant IDs, with corresponding information about diversity audit tools and library communities. "Community Type" describes the size of the service community of the library that the interviewee worked at at the time of conducting diversity audit, or at the time of the interview (in the case of P13, who did not conduct a diversity audit). Community size categories were drawn from the FY 2022 Institute of Museum and Library Services Public Libraries Survey[2].

See Table 7 for a list of interviewees. Our interview protocol was reviewed by the IRB and determined to be exempt.

A.2 Interview Instrument

We provide a PDF copy of our interview instrument. The linked questions were used as prompts for interviews via Zoom, with additional follow-up questions asked based on responses. Interviewees were first asked for their consent to participate in the interview and to record the interview for the purposes of transcription and internal review.

A.3 Survey

We developed an online survey about librarians' experiences with and perspectives on collection diversity audits, which we shared with relevant mailing lists and online communities. The survey was open from January 14 to January 21. We offered a drawing where respondents could win one of three $100 Visa gift cards. On one evening during the survey window, we received a large number of fraudulent responses from the same IP address, which prompted us to put the survey behind a password. We used Qualtrics to process survey responses, and we filtered out the fraudulent responses and some other spam by cross-referencing IP addresses, timestamps, and the platform's fraud detection metrics. Responses were also reviewed for coherence and relevance to the survey topic. While efforts were made to identify and exclude low-quality responses, some fraudulent entries may have gone undetected. This represents a potential limitation of our study. Our survey protocol was reviewed by the IRB and determined to be exempt.

A.4 Survey Instrument

We provide a PDF copy of our survey instrument.

A.5 Spending Data

We use GovSpend to explore public libraries' diversity-related purchases. We search for any item purchased by a public library that exceeds $20 and includes the words "diversity", "dei", "icurate", or "collectionhq". The search results capture partial spending from 24 U.S. public libraries. We summarize and plot these data in Figure 4 and Figure 5.

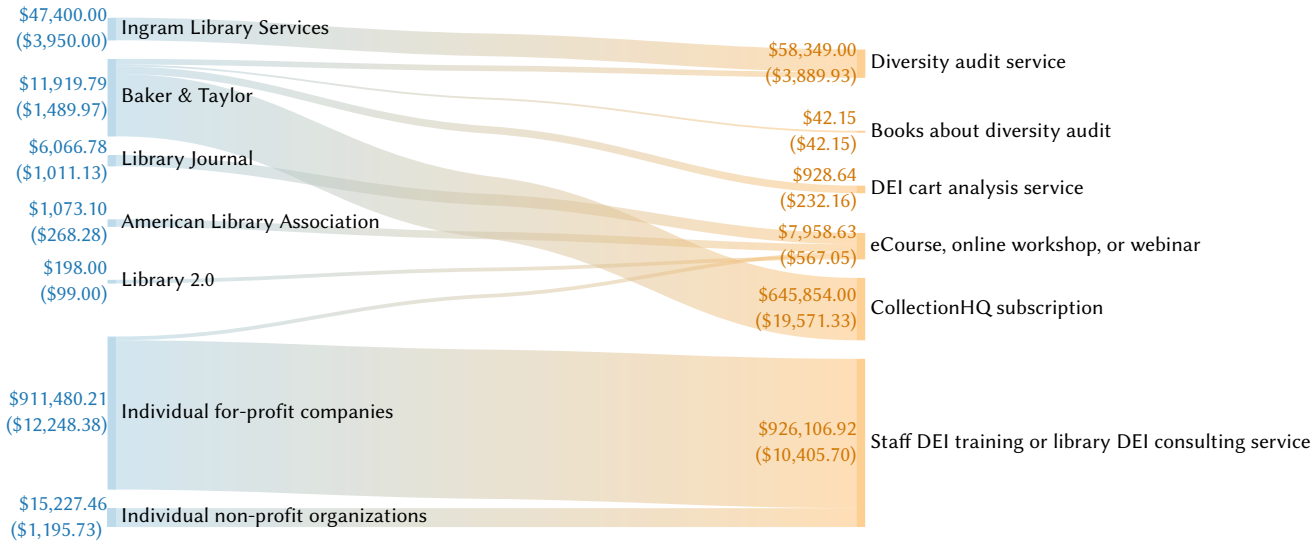


Figure 4: This figure shows diversity-related purchases for 24 U.S. public libraries, drawn from GovSpend (for any item exceeding $20 and including the words "diversity", "dei", "icurate", or "collectionhq"). Average item price is listed in parenthesis. Diversity audit services were provided by Ingram and Baker & Taylor, and the price of such services cost several thousand dollars. A wider range of organizations provide services that aim to educate library employees on the concept of diversity. Libraries appear to spend more on staff DEI training and consulting than collection audits, averaging around $10,000. However, as Table 5 and 6 show, libraries spend more with vendors that provide collection audits overall.

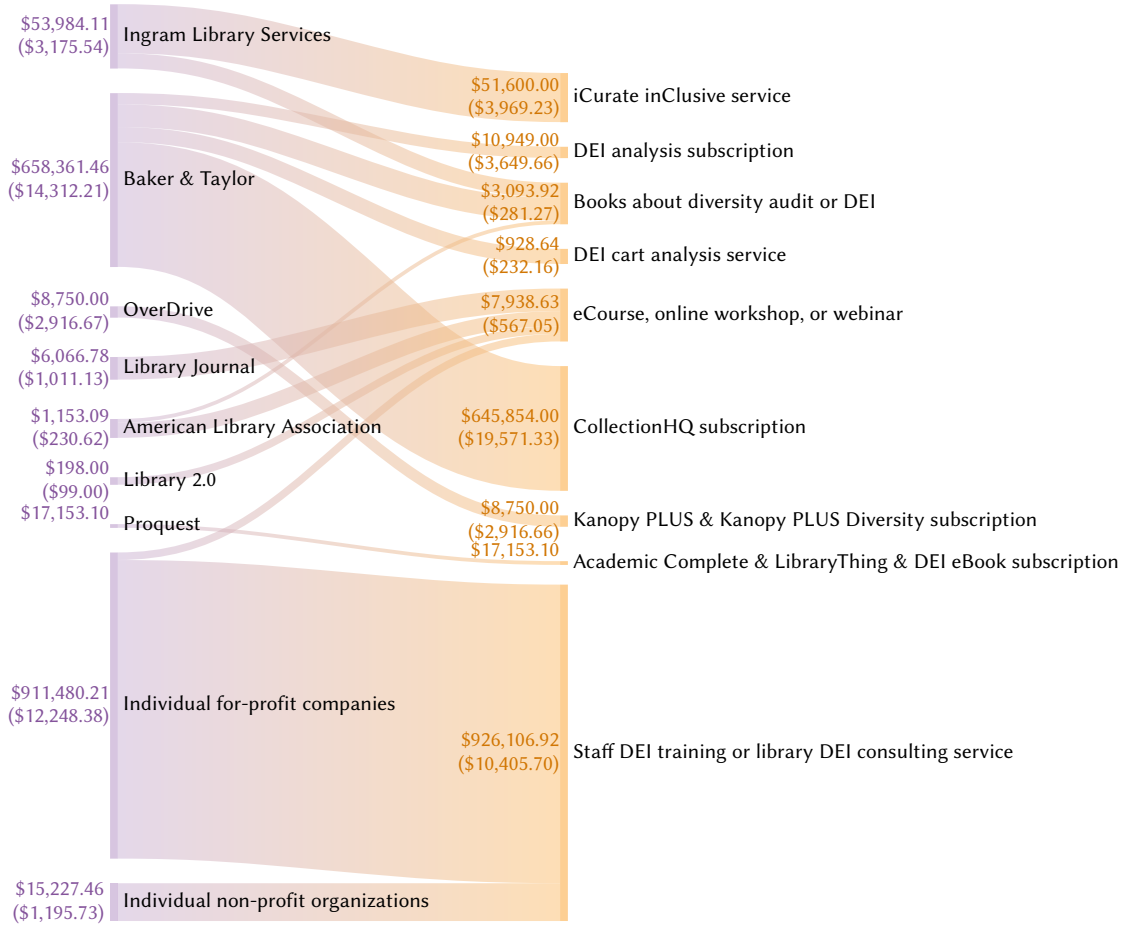


Figure 5: This figure shows diversity-related purchases from 57 U.S. public libraries. Purchase records are drawn from GovSpend and aggregated by type of seller and purchase; we include average item price in parenthesis. These results reflect any purchases on GovSpend that include "diversity", "dei", "icurate", or "collectionhq" in the item description; include "library" in their agency name; and exceed $20. These figures represent the minimum amount libraries spent during this period.