Occupational identities in the age of generative AI: the case of librarianship

Viviane Ito and Lyric Grimes

Abstract

Introduction. This paper analyses AI-generated depictions of librarians to determine their alignment with stereotypical portrayals. Previous research has highlighted gender biases in large language models (LLMs) and AI-generated images. However, no studies have examined AI-generated images of librarians. This study fills that gap by exploring how these images uphold stereotypes.

Method. Data was collected from ChatGPT, DALL-E, Midjourney, and Adobe Firefly using gender-neutral prompts in American English and Brazilian Portuguese.

Analysis. We performed quantitative and qualitative analysis to compare languages and models. Thematic analysis revealed recurring themes and patterns in the visual representations.

Results. Our findings indicate that AI-generated images often depict librarians as white women, with stereotypical elements like glasses and cardigans. We also observed that different model versions and languages generate different portrayals of the profession. When prompted in Brazilian Portuguese, the models tended to offer a more representative image of the professionals.

Conclusion. The study underscores the need for a critical approach to Generative AI, as training data reflects societal biases, perpetuating stereotypes. These portrayals can impact the public perception of librarians, potentially alienating users and reinforcing an outdated, predominantly white, female, and middle-class image of the profession.

Introduction

The rapid adoption of artificial intelligence (AI) is transforming how knowledge is produced, circulated, and interpreted within academic contexts. Librarians, in particular, are grappling with how these technologies affect their professional practices and how they should guide students and researchers in their responsible use. At the same time, the profession itself is undergoing scrutiny: the role of librarians, and even the continued necessity of libraries as traditional knowledge infrastructures, is increasingly uncertain (Carnegie & Abell, 2009).

Librarianship provides a distinctive case for examining the cultural and technological construction of occupational identity. The profession has long been associated with enduring stereotypes, shaped by intersections of gender, race, and class (Natarajan, 2017; Schlesselman-Tarango, 2017). Librarianship has historically been feminized as a service occupation, racialized through associations with whiteness, and exclusionary toward marginalized groups. These dynamics render librarians a particularly salient lens through which to analyze how emerging technologies such as AI reproduce, transform, or destabilize professional images.

In this study, we focus on how generative AI models depict librarians. We ask three research questions:

RQ 1: *In what ways do AI-generated images reproduce or challenge existing stereotypes of librarians?*

RQ 2: *What do these images reveal about how identity is encoded in generative systems?*

RQ 3: *How might AI image generation illuminate broader societal perceptions of librarians' roles?*

Our contributions are threefold. Empirically, we examine the visual representations of librarians produced by generative AI, highlighting how these systems reproduce, distort, or reconfigure professional identities. Theoretically, we extend scholarship on identity construction in sociotechnical systems by showing how AI-generated imagery both reflects and shapes cultural understandings of librarianship. Finally, we introduce a prompt–analysis framework for investigating occupational stereotypes in AI-generated media, which can be applied to other professions and identity groups.

Theoretical framework

Sociotechnical matters

This work draws on sociotechnical theories to examine the interplay between technological artifacts and their social consequences, as highlighted by Sarker et al. (2019). A sociotechnical approach integrates technical systems with the psychological, cultural, and economic dimensions of their impact. Hess and Sovacool (2020), in their analysis of a decade of STS publications (2009–2019), identify four key sociotechnical themes: cultural aspects, particularly sociotechnical imaginaries; policy-related analyses; public participation processes; and the study of sociotechnical systems. This study engages especially with sociotechnical imaginaries, public participation, and sociotechnical systems.

The concept of sociotechnical imaginaries is central to this work, as it interrogates how AI, particularly large language models (LLMs), interprets and constructs visions of the librarianship. Second, public participation processes are critical in our work, as users contribute to building these systems by providing training data and interacting with them through prompting and probing. Finally, the focus on sociotechnical systems lies at the core of this work, particularly in analyzing the dynamic interactions between humans and AI.

Occupational stereotypes and the librarians' professional identity

Defining the typical librarian

When defining a librarian, one might envision '...a group of 'white women in sensible shoes" (Espinal 2001, p. 146). ChatGPT (2025) describes librarians as 'bookish, serious figures with glasses, tight buns, and conservative clothing, often shushing noisy patrons. They are seen as quiet and knowledgeable,' often accompanied by books or a charming cat. The archetypal librarian is a quirky, nerdy white woman in a cardigan, round glasses, and practical shoes. While some view librarians as strict and reserved, others consider them compassionate and nurturing. Typically, librarians are female and white, protecting knowledge and enforcing library values despite meager compensation.

Appearance and personality

There are two perceptions of librarians: the historical and the modern. In 1989, Frank Spaulding detailed five stereotypes: 1) orderliness, 2) conformity, 3) passivity, 4) introspection, and 5) anxiety. Both historic and modern librarians embody these traits. Historically, librarians are depicted as women wearing glasses with timid or stern demeanors, often in buttoned long-sleeved blouses (Kneale, 2009). They don't laugh, prefer silence, love books, and endure humanity (Kneale, 2009). This figure is viewed as intimidating (Radford, 2001). In contrast, the modern librarian is often described as a 'cat lady' with a friendly demeanour (Gambrell & Brennan, 2014), where a pet cat enhances perceptions of kindness (Schlesselman-Tarango, 2017). The modern 'adorkable' librarian (Schlesselman-Tarango, 2017) still loves books, often wears glasses, and favours cardigans enhanced with stylish bangs and cheerfulness. This image suggests neutrality and harmlessness, with today's librarians characterized as helpful, caring, service-oriented, and cool (Annoyed Librarian, 2007; Rabey, 2007; Schneider, 2007; Yontz, 2003).

Although these shifting viewpoints mirror wider social transformations, both past and present depictions highlight the lasting views of librarians held by the public.

Values, actions, and morals

The American Library Association (ALA) core values define the profession's principles and aspirations, serving as ideals that unite and motivate library workers. The ALA recognizes access, equity, intellectual freedom, privacy, public good, and sustainability as its core values (2006). Often, literature on librarianship values librarians' dedication to service, democracy, and inclusion (Ettarh, 2018), and they are regarded as 'guardians of the First Amendment' (Clinton, 2017). Librarians and libraries are often held in high regard due to the sacrifice, struggles, and martyrdom commonly associated with this field (Eberhart, 2019). Librarians, like teachers, are frequently seen as martyrs for their commitment and dedication to their profession, even amid inadequate monetary compensation (Jacobsen, 2004; Ettarh, 2018).

These widely recognized values and perceptions influence how librarianship is viewed by the public, emphasizing prevailing beliefs about librarians and libraries.

Who are librarians?

The stereotypical image of a librarian as a morally upright, book-loving, and perhaps quirky white woman in sensible shoes is a misrepresentation that fails to reflect the true diversity and lived experiences of librarians. This raises the question: Who are librarians, really? While information about the demographics of librarians is hard to find, the profession of librarianship in the United States has primarily been dominated by white women who make a modest living of about 70,000 USD a year. (Department for Professional Employees, 2024). Data shows these demographics are consistent, with minor race, degree attainment, and income variations.

Race and ethnicity

In the U.S., librarianship is largely a white profession. Data USA (2017) reported that 6.66% of library workers were Hispanic, while 93.3% were non-Hispanic. That year, 85.9% of librarians were white, with 6.49% Black and 4.24% Asian. By 2023, the Department for Professional Employees (2024) confirmed over 81% of librarians are white, 7% Black or African American, 11% Hispanic or Latino, and 5.5% Asian American or Pacific Islander. In 2021-2022, only 4.6% of female MLS program graduates were Black, 2.7% Asian American and Pacific Islander, while Hispanic women made up 8.4% (National Center for Education Statistics, 2020).

Gender

Women have dominated librarianship for years. In 1995, 83.9% of librarians were women, increasing to 84.4% in 2003. In 2017, female representation decreased by 2% to 82.4%, while men made up 17.6% (Data USA, 2017). As of 2023, women represented 82.5% of librarians and 83.2% of library assistants. In 2022, women comprised 82.5% of Master of Library Science graduates (National Center for Education Statistics, 2020). Despite this dominance, women earn only 82.3% of men's earnings (Department for Professional Employees, 2024).

Data about Brazilian librarians is scarce, but similar to the US, there is a strong feminization of the profession: women account for approximately 82% of librarians in Brazil (Muller and Martins, 2019), a demographic reality that amplifies the association between librarianship and socially constructed notions of feminized labor.

Stereotypes of Brazilian librarians

The literature on librarianship in Brazil demonstrates that the profession has been deeply shaped by gendered representations and persistent stereotypes. Studies identify a tension between traditional and contemporary images of women librarians, ranging from the stereotype of the *'old spinster'*—serious, rigid, conservatively dressed, and devoted to silence and order—to a newer, hypersexualized portrayal of the librarian as young, uninhibited, and visually appealing to attract users (Escalante et al., 2021). Although librarians largely reject the idea that traits such as care, dedication, and gentleness are inherently feminine, they nonetheless acknowledge the relevance of these characteristics to professional practice, revealing how gendered expectations remain embedded in perceptions of librarianship (Escalante et al., 2021).

Beyond symbolic representations, gendered perceptions also intersect with the historical and institutional undervaluation of librarianship. Historically, library work was associated with elite male domains, such as royal courts, clerical institutions, and scholarly communities, before becoming progressively feminized—a shift that contributed to its contemporary alignment with domestic and organizational labor rather than scientific expertise (Ferreira, 2020). Consequently, the librarian's role is often reduced to activities of arranging and organizing, obscuring the scientific criteria and methodological processes underlying information organization, retrieval, and dissemination (Ferreira, 2020). This professional invisibility is further reinforced in school libraries, where, despite legal frameworks recognizing librarianship as a regulated profession, few formally trained librarians occupy designated positions, and library activities are frequently performed by non-specialized staff or employees in job readaptation (Fujita et al., 2018). Together, these findings indicate that gender stereotypes, historical trajectories, and institutional practices jointly shape the social image and labor conditions of librarianship in Brazil.

Although research on librarian stereotypes has expanded in North American and European contexts, there is still a lack of empirical work examining how librarians are stereotyped or imagined in Brazil. Besides the works presented in this section, existing discussions are theoretical, historical, and epistemological (Garcês-da-Silva & Saldanha, 2023; Vieira & Karpinski, 2019),

focusing on broad issues of professional identity and challenges in Brazilian LIS (Tanus & Sánchez-Tarragó, 2020).

Bias and representation in generative AI

Feminist science and technology studies (STS) have extensively critiqued algorithmic systems, particularly for their disproportionate impact on marginalized communities such as women and people of colour. Noble (2018) exposes racism in search engine algorithms, demonstrating how Black women are often portrayed in hypersexualized ways in top search results. Benjamin (2019) extends this critique by analyzing exclusionary technologies, such as the racial bias inherent in facial recognition systems and framing these systems as a continuation of systemic inequality akin to the *'new Jim Crow.'* Additionally, Constanza-Schock (2020) highlights that the composition of engineering and design teams in the tech industry—predominantly white, male, and elitist—contributes to these biases. Harris (2023) underscores that the founders of major tech companies often come from privileged backgrounds, further entrenching existing disparities.

Scholarship on algorithmic bias has consistently demonstrated that AI systems reproduce and amplify existing social stereotypes. Much of this research has focused on gender and racial bias in natural language processing and computer vision. For example, Kotek et al. (2023) show that large language models disproportionately associate occupations such as secretary and nurse with women, while Currie et al. (2024) find that AI-generated images of medical professionals overwhelmingly depict white men. These findings echo earlier work documenting that biases emerge directly from the datasets on which models are trained (Bolukbasi et al., 2016; Caliskan et al., 2017). The consequences of such skewed data are well documented: facial recognition systems, for instance, have been found to perform with significantly lower accuracy for Black women (Buolamwini & Gebru, 2018).

Occupational roles have become a central site for testing bias in AI systems. Studies demonstrate that models reproduce entrenched gendered and racialized associations—for instance, *'man is to computer programmer as woman is to homemaker'* (Bolukbasi et al., 2016) or the tendency for software engineers to be represented primarily as white men (Bianchi et al., 2023). Naik and Nushi (2023) briefly tested librarian-related prompts, but librarianship was not the focus of their study, and their prompting remained relatively limited. Thus, while text-to-image audits have documented consistent patterns of gendered and racialized representation, occupational identities within public-facing knowledge work—such as librarianship—remain under-examined.

The absence of focused studies on librarianship in generative AI research leaves unanswered questions about how this historically feminized and racialized profession is visually constructed by emerging technologies. To explore these questions, we designed a comparative study of AI-generated images across different models and cultural contexts. The next section details the materials and methods that guided our data collection and analysis.

Materials and methods

Models

We collected data from three different text-to-image models: DALL-E/ChatGPT, Midjourney, and Adobe Firefly. We used different model versions for comparison purposes as described in Table 1.

Model	Versions
DALL-E/ChatGPT	4 and 5
Midjourney	7
Adobe Firefly	3 and 4

Table 1. Models and versions.

We did not adjust any model parameters for the experiment, relying instead on each system's default settings. Our goal was to simulate a real-world user experience. We simulated what a user would encounter when prompting models '*in the wild.*' Accordingly, we did not request a specific number of images. Firefly and Midjourney automatically generate four images per prompt, while ChatGPT produces one or two. Repeated prompts in ChatGPT yielded highly similar images, indicating saturation; therefore, we analysed only the original outputs.

The models were prompted in American English and Brazilian Portuguese. One of the authors speaks Brazilian Portuguese as their first language, and one of the authors speaks American English as their first language, which justifies the choice of languages for the case studies. This variation in language and nationality interrogates the predominance of English-language data from the United States (Cheong et al., 2024) and offers a broader perspective on how librarians are depicted across different linguistic and cultural contexts. To prevent the generation of images according to previous chats, we disabled the '*memory*' feature when available and erased previous conversations.

Prompts

To collect the data, we crafted three types of prompts. The first type of prompt was a general request to create an encompassing image of a librarian. The second type was based on the work on doctors' stereotypes from Currie et al. (2024), asking for quintessential images of librarians. The third type was a request for clarification of why the system had come up with such images. Only ChatGPT/DALL-E provided textual explanations regarding the generated images. The other models generated another batch of images as explanations.

Prompt type	Prompt in US English	Prompt in Brazilian Portuguese
General Request	Can you create an image of a librarian?	Você pode criar uma imagem de uma pessoa bibliotecária?
Quintessential Images	You are a book cover designer and will create an image for a book that contains quintessential representations of librarians. Can you create an image of a librarian? I want you to surprise me and create it based on your knowledge about libraries, their service, the importance for the community, while also offering a representative image of the librarian.	Você é um designer de capas de livros e criará uma imagem para um livro que contém representações quintessenciais de bibliotecários. Você pode criar uma imagem de uma pessoa bibliotecária? Eu quero que você me surpreenda e crie a imagem baseada no seu conhecimento sobre bibliotecas, os serviços prestados, a importância delas para a comunidade, e ao mesmo tempo oferecendo uma imagem representativa de uma pessoa bibliotecária.
Explanation	Can you explain why you represented the librarian with this image?	Você pode explicar por que você representou a pessoa bibliotecária com esta imagem?

Table 2. Prompts

Thematic analysis

We applied thematic analysis (Braun & Clarke, 2006) to categorize the images. The co-authors independently analysed and labelled the collected data, identifying recurring themes and patterns in the visual representations. The co-authors then compared annotations to ensure consistency and resolve discrepancies through a collaborative labeling session. This process ensured a rigorous and comprehensive analysis, highlighting both shared and divergent portrayals of librarians based on language, nationality, and cultural context.

The categories of analysis were gender, race, and age. Gender was coded based on perceived gender presentation in images, rather than on gender identity. As images do not provide access to individuals' self-identification, the analysis relied on culturally legible visual cues commonly associated with gendered appearance. The primary categories used were *women* and *men*, reflecting dominant binary gender schemas through which gender is typically rendered intelligible in visual media. These categories do not imply a belief in gender as a binary construct, but rather correspond to the interpretive constraints of image-based analysis. In cases where gender

presentation did not conform clearly to binary norms, images were categorized as *gender nonconforming/non-binary presentation*. Trans categories were not included in the coding scheme, as transgender identity cannot be reliably inferred from visual data alone. Trans men and trans women may be visually indistinguishable from cisgender individuals, and any attempt to code trans status based solely on appearance would risk misclassification and reinforce essentialist assumptions about gender visibility. This constitutes a limitation of the study and reflects broader constraints in visual content analysis.

For race, we categorized into white, Black, Hispanic/Latino, and Asian. After coding independently, we also noted instances where the images were racially ambiguous, creating a new category for those cases. In terms of age, we used the classification suggested by Naik and Nushi (2023): *'Child or minor'*, *'Adult 18-40'*, *'Adult 40-60'*, *and 'Adult over 60'*. (Cohen's Kappa for Race: κ = 0.70 (N=94); Gender: κ = 0.76 (N=94); Age: κ = 1.00 (N=94)).

An additional content analysis was performed by one of the co-authors, due to her background in Library Sciences. This analysis focused on the settings, services, appearance, and other themes pertinent to the representation of the librarian as a professional.

Results

The image generation resulted in a total of 94 items. Both Firefly and Midjourney generate a batch of 4 images per prompt. As mentioned in the Materials and Methods section, ChatGPT generates only one or two images per prompt.

Model	Number of images
ChatGPT/DALL-E	22
Midjourney	24
Adobe Firefly	48

Table 3. Number of output images per prompt

In a general sense, the results of age, gender, and race reflected the expected stereotypes of young white women as most represented. 60% of the images contained adults 18-40; 69% of the images are of female-looking characters, and 62.8% are white. In second place, there were 14.9% racially ambiguous librarians, and in third place, 12.8% Black librarians. Hispanic/Latino comprised 3.2% of representation.
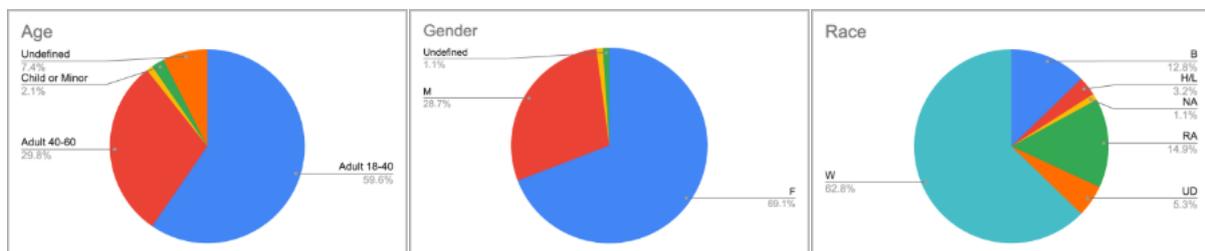

Figure 1. Results across all models

The reproduction of librarians' stereotypes

Appearance, clothing, and dress

Appearance, clothing, and dress were among the first codes to emerge. These were defined and coded as the librarian's physical appearance, attire (e.g., cardigans, glasses, blazers), and overall style of dress, such as casual, business casual, or business formal. These codes were important because they captured both stereotypical and non-stereotypical visual cues, providing insight into

how librarians were represented in terms of age, professionalism, and social identity. By analyzing these elements, we identified recurring patterns, such as the association of certain clothing styles with specific librarian archetypes. We assessed how these visual features intersected with other coded attributes, including race and age. As anticipated, many of the generated images reproduced stereotypical depictions of librarians. They clustered primarily into two dominant categories: (1) the *'loveless frump,' 'old maid,'* or *'prim spinster'* archetype, and (2) the *'young, hip party girl'* or *'lipstick librarian'* cliché (Adams, 2000; Seale, 2008).

Images in the first category typically featured older women with gray or white hair—often styled in a neat bun—wearing formal or conservative attire, such as blazers or suits, and frequently holding books. In contrast, images in the second category typically featured conventionally attractive, slender, white, blonde women dressed in more casual or business-casual attire. Only one image had a visibly fat or chubby person, highlighting the lack of body diversity in the generated images. Additionally, the images only depicted able-bodied individuals, with no images of visibly disabled or differently abled people. This pattern held true for images generated from both English and Brazilian Portuguese prompts. Notably, the first set of images showed librarians in a uniform, which isn't typical for librarians in the US or Brazil.

Unexpectedly, however, the dataset displayed greater racial diversity than anticipated. Approximately one-third of the images were coded as non-white, encompassing Black, non-white Hispanic/Latino, and racially ambiguous individuals. Nevertheless, these figures still fit into either category one or category two stereotypes – young, hip, and attractive or matronly frump. Across both categories, hallmark *'librarian'* attributes—such as glasses, cardigans, and business or business casual attire—remained consistent.

Figure 2. Category one: *'loveless frump,' 'old maid,'* and *'prim-spinster'*

Figure 3. Category two: young, hip, and sexy party girl or the *'lipstick librarian'*

Challenging librarians' stereotypes

A few sets of images deviated from the expected patterns. These images didn't fit neatly into the typical librarian stereotypes found in the main categories, and some showed unconventional roles, clothing, or settings, challenging traditional views of the profession. Examining these outliers was essential because they highlighted the range of possible representations and suggested that the generated images could occasionally move beyond common stereotypes, providing insight into alternative ways librarians can be visually constructed.

Male librarians

By far, most of the images were classified as female, with only 27 coded as male, which aligns with the national average in the United States and the historic feminization of the profession (Pagowsky & DeFrain, 2014). Although the male images resembled the female images in terms of location, environment, actions, and framing, they didn't fit neatly into the two dominant categories: the *'old maid'* or the *'young and hip'* clichés. Instead, the male images tended to be more neutral in terms of appearance and attire. Men, both young and old, generally looked well-groomed and conventionally attractive. Some might be seen as young and hip, but the distinctions were less noticeable than those among the female images, making precise subgroup categorization tricky.

Figure 4. Male librarians

Framing: librarians not as the central focus

One set of images stood out from the rest because they did not focus on the librarian; instead, the photos appeared to focus on library patrons and users. In the images, the two adult figures were hard to identify as librarians because they lacked the typical traits of librarians, such as appearance, dress, framing, and actions.

Figure 5. Depictions of libraries focusing on patrons

Encoded identities

From a sociotechnical perspective, large language models are studied in terms of how their technical components are influenced by society's values and perspectives. We evaluated the identities that are encoded in the text-to-image models by understanding how different languages, models, and model versions interact and depict different identities of librarians.

While in American English there were more instances of adults at the age 40-60, in Brazilian Portuguese, we observed that the depicted librarians were younger – adults 18-40. In terms of gender, the Brazilian Portuguese had a more balanced distribution of female and male people. Race

is also differently depicted in Brazilian Portuguese, with more occurrences of Hispanic/Latino, racially ambiguous, and Black librarians.
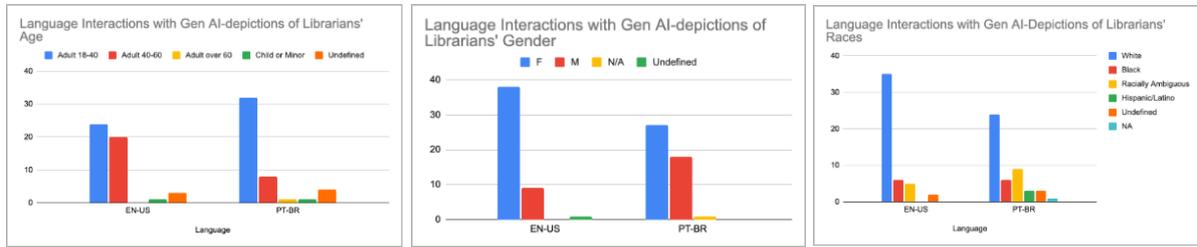

Figure 6. Comparison between languages

When comparing models, we noticed that Firefly was the only model that had generated more librarians at the age bracket of 18-40. ChatGPT focused on 40-60. Although all models prioritized females, Firefly had the most predominant female images. Midjourney was the most equal between genders.
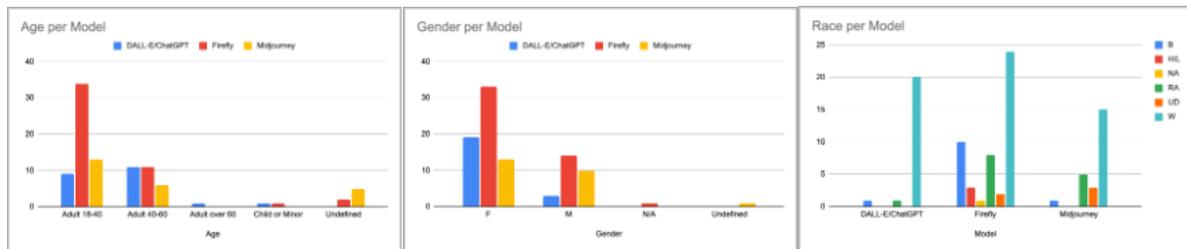

Figure 7. Comparison between models

In all models, most of the images depicted white librarians. Firefly was the one that depicted more diversity, still with most results being white. Although representing Black and Hispanic/Latino, the models did not depict any other races.

When comparing versions, we highlight the differences between versions of Firefly 3 and Firefly 4. In the latter, the images became more diverse.
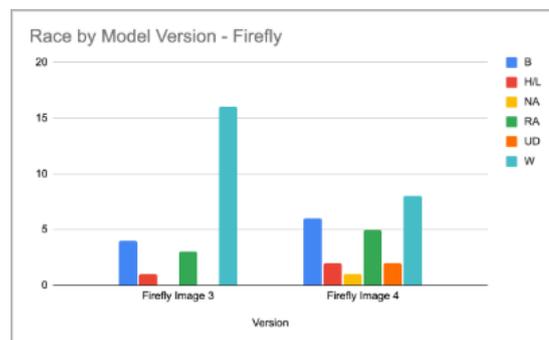

Figure 8. Comparison between model versions

The comparisons showed that variations on the user background (e.g.: language) interacts with models and model versions, generating different depictions of librarians. Besides whom the librarians are, we analysed how the context of the librarians was represented in terms of establishing the role of libraries.

The role of librarians

Actions

Actions were among the last codes to emerge during the coding process. We defined them as the *'tasks performed by the person in the generated scenario.'* Common examples included holding books, shelving, or organizing books, browsing shelves, and standing in the book aisle. These codes captured both the visible activities of librarians and societal perceptions of their professional duties.

Figure 9. Actions of Librarians

Across the generated images, individuals were frequently depicted performing book-related tasks, highlighting familiar librarian behaviors, and reinforcing common expectations of the role, such as being *'detail-oriented and maintaining order in the library'* (Seale, 2008).

Setting and environment

Setting and environment were additional codes that emerged during the coding process, capturing the settings in which librarians were depicted and providing insight into how space contributes to their portrayal. Coding for these features contextualized the librarians' actions and appearance, illustrating how surroundings reinforce certain perceptions or stereotypes of the profession. In some cases, the location was coded as the librarian's *'house'* (Kumasi, 2013), representing the library as the librarian's domain.

Common locations included a general library space (a vague setting that could be interpreted as a library), a book aisle, and a library or archive reading room. These environmental features allowed us to examine how spatial context interacts with other coded characteristics, helping to shape and reinforce the visual representation of librarians.

Figure 10. Setting and environments of the librarians' portrayals

Framing

Although not an official code, framing emerged repeatedly during the coding process. We defined it as *'how the librarian was constructed within the image,'* capturing both societal perceptions and the symbolic significance of librarians. Some images depicted librarians as approachable, down-to-earth figures, while others portrayed them as authoritative sources of knowledge. In the latter cases, the librarian was not only central to the image but also represented as a bestower of knowledge or a symbolic figure, sometimes appearing oversized or mythic—an iconic embodiment of the profession rather than an ordinary individual. These portrayals echo the Borgesian conception of the library (Santamaria, 2020) and present the librarian as the patron's information champion (Stoddart & Lee, 2005).
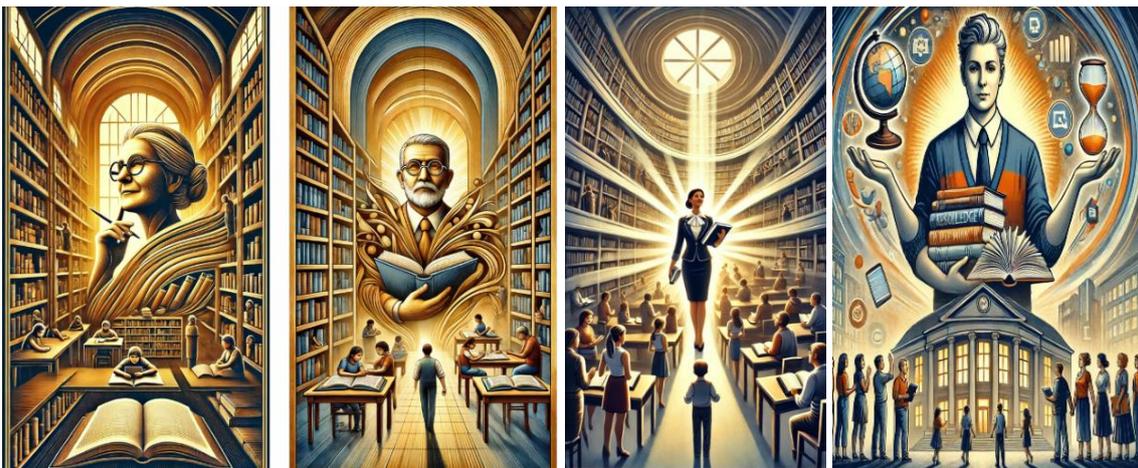

Figure 11. Librarians as mythic figures

Aesthetic

A notable feature across many of the generated images was the limited representation of librarians' professional activities. Instead of illustrating the day-to-day tasks that define librarianship—such

as information literacy instruction, community engagement, or digital resource management—the images often relied on vague or highly aestheticized environments. Many outputs evoked an atmosphere of '*dark academia*,' blending romanticized imagery of bookshelves, gothic architecture, and candlelit reading rooms with fantasy or science-fiction elements. These stylized depictions suggest that the models drew more heavily on cultural tropes of libraries as mysterious or magical spaces rather than on the actual professional practices of librarians.

Figure 12. Fantasy, steampunk, and '*dark academia*' aesthetic

This tendency reflects a broader lack of embedded knowledge in generative AI about what librarians actually do. While librarianship is a service-oriented profession with a strong emphasis on access, equity, and community roles (ALA, 2006; Ettarh, 2018), the outputs rarely conveyed these values. Instead, the focus was on aesthetic atmospheres and book-centered symbolism. This disconnect is significant: it reveals how generative models privilege symbolic and literary associations of libraries over professional realities, thereby obscuring the labor, expertise, and social responsibilities that define librarianship.

At the same time, the prevalence of fantasy and trend-driven styles (e.g., '*dark academia*' or steampunk) highlights how libraries remain powerful cultural symbols. On one hand, this cultural resonance can be seen as positive, underscoring the enduring allure of libraries in the public imagination. On the other, it risks flattening the profession into purely nostalgic or mythical terms, further marginalizing librarians' contemporary roles as digital curators, educators, and advocates for information justice.

Discussion

Our findings highlight the dual role of generative AI as both a mirror of existing societal stereotypes and a potential site for reimagining professional identities. Consistent with prior research on occupational stereotypes in AI outputs (Bolukbasi et al., 2016; Bianchi et al., 2023; Currie et al., 2024), the images reproduced dominant cultural scripts of librarianship, such as white female figures adorned with glasses or cardigans. These representations reflect how training data encode longstanding associations of librarianship with femininity, whiteness, and service work.

Importantly, this pattern was not uniform across linguistic contexts. While English-language prompts largely converged on highly standardized representations, Portuguese-language prompts—particularly in Brazilian Portuguese—yielded more racially and gender-diverse portrayals. This divergence can be partially understood in relation to the institutional and cultural context of librarianship in Brazil. Although a federal law enacted in 2010 mandated that every Brazilian school maintain a library (Fujita et al., 2018), recent data indicate that only approximately 63% of schools currently comply (Souza, 2025). Moreover, library roles are frequently occupied by individuals without formal training in Library and Information Science, with teachers often assuming librarian responsibilities alongside their instructional duties. This gap between policy and practice reflects a profession that remains less institutionally consolidated and symbolically standardized.

This relative lack of professional standardization may help explain why Portuguese-language prompts yielded more representative and diverse image outputs. Compared to English-language corpora, Portuguese-language training data are smaller and less dominated by globally circulating professional stereotypes. As a result, Portuguese prompts may activate a broader and less rigid set of visual associations, allowing generative models to produce more varied depictions of librarians rather than converging on a narrow, canonical representation.

Prior work in machine learning and NLP suggests that increasing data volume does not necessarily lead to greater representational diversity. Instead, large, and repetitive corpora tend to amplify socially dominant associations and stabilize existing stereotypes (Bender et al., 2021; Bolukbasi et al., 2016; Caliskan et al., 2017). Conversely, when training data are smaller or less standardized, model behavior may become under-specified, allowing multiple plausible representations to coexist (D'Amour et al., 2022). In generative systems, this under-specification can manifest as greater variation in outputs—not as a signal of higher accuracy, but because of weaker convergence on a single canonical representation.

Our study illustrates how generative systems may gradually shift with expanded training data or design interventions. They also underscore the role of user interaction: the phrasing and cultural context of prompts significantly influenced model outputs, suggesting that AI does not only reproduce bias but also responds dynamically to linguistic variation.

From a sociotechnical perspective, these results reaffirm that occupational identity is not a fixed construct but a negotiated outcome of technical design, cultural imaginaries, and user participation (Hess & Sovacool, 2020). Librarianship, as a historically feminized and racialized profession, is particularly susceptible to reductive portrayals. If AI systems widely circulate stereotypical images of librarians, they may reinforce outdated perceptions of the profession, further marginalizing underrepresented groups in librarianship. Conversely, more diverse depictions—though less frequent—point to opportunities for reconfiguring the public imagination of librarians as more inclusive, intersectional figures.

This tension raises critical questions for both librarianship and AI. For librarians, it emphasizes the importance of professional advocacy: challenging stereotypes, articulating diverse narratives, and actively engaging with AI tools to ensure that portrayals of librarians align with lived realities. For AI practitioners, it highlights the responsibility to address representational harms, diversify training data, and design systems that acknowledge cultural and linguistic variation.

Implications

Rather than treating librarian stereotypes and AI-generated representations as static, external misrepresentations, it is necessary to position librarians and LIS professionals as active agents who may strategically perform, negotiate, or resist these narratives within everyday professional practice and education. At present, demographic diversity within LIS remains largely unchanged. Racial and gender representation remains approximately 80% white (Department for Professional Employees, 2024) and 80% female (Data USA, 2017). Similar patterns have been observed within Brazilian librarianship. Despite ongoing discourse on innovation, equity, and professional transformation, concrete measures to counteract these entrenched images remain limited, as evidenced by the persistence of these demographic trends.

To effectively combat and strategically leverage professional stereotypes, LIS must first confront the institutional harms and structural inequalities embedded within the profession itself, as these conditions continue to constrain whose identities, labor, and futures are imagined as representative of librarianship. Both the U.S. and Brazil have a legacy of racism and discrimination that has historically excluded Black individuals and other marginalized groups from the profession. Furthermore, access to LIS education is not evenly distributed; financial constraints, admissions

practices, and program location disproportionately limit access for marginalized students. These barriers extend beyond education itself, as employment opportunities and post-graduation salary expectations further shape who can enter and remain in the profession. According to the U.S. Bureau of Labor Statistics (2024), the median annual wage for librarians is approximately 64,320 USD. Such compensation often fails to keep pace with the rising cost of graduate education in the U.S, narrowing the pipeline into librarianship and privileging individuals with greater financial security and institutional support, while disproportionately excluding racialized and marginalized populations. In turn, these structural and economic barriers reinforce the profession's longstanding demographic homogeneity and sustain narrow professional imaginaries. Addressing long-term disparities in race, gender, access, and salaries within the field is essential for disrupting entrenched inequities, expanding whose experiences and identities are represented in LIS, and reducing the reproduction of limiting professional stereotypes in both human and AI-generated representations.

Furthermore, librarians can only successfully challenge stereotypes in their daily work and teaching once they have addressed underlying structural inequalities. In the meantime, they can work towards reducing these disparities by creating programs, instruction, and outreach initiatives that showcase a variety of professional identities, question assumptions about who fits into the profession, and demonstrate expertise that counters simplistic or stereotypical images. In LIS education, instructors can use AI-generated representations and media portrayals as discussion tools, prompting students to critically reflect professional stereotypes and their real-world impact. Mentorship and advocacy for colleagues from marginalized and underrepresented backgrounds further expand who is seen as a legitimate participant in the profession. Even within existing structural constraints, such actions allow librarians to transform knowledge of inequality into concrete professional interventions, directly shaping the narratives, identities, and representations that both human audiences and AI systems encounter.

Strategic, sustained partnerships with black- and minority-serving institutions and with gender-inclusive, BIPOC-led organizations, such as the Black Caucus of the American Library Association (BCALA), provide an additional mechanism for counteracting professional stereotypes and uplifting LIS professionals from historically marginalized backgrounds. For instance, North Carolina Central University (NCCU) is the only Historically Black College or University (HBCU) in the United States that offers graduate degrees in library science (White, 2025). As with many HBCUs, NCCU's LIS program primarily serves BIPOC students and has faced longstanding underfunding and recent budget cuts (National Public Radio, 2023). Strategic collaborations with well-resourced professional LIS organizations—such as the American Library Association (ALA) and the International Federation of Library Associations and Institutions (IFLA)—could help redistribute institutional resources, amplify BIPOC voices in LIS education, and challenge structural inequities within the profession. Well-funded and resourced programs produce LIS graduates and professionals at a higher rate than underfunded institutions, perpetuating structural inequities that disproportionately affect BIPOC students and limit whose knowledge and leadership shape the field, which further reinforces stereotypes and constraints on professional identity.

Without addressing these material and structural barriers, efforts to counteract stereotypical or algorithmic representations of librarians' risk remaining superficial, treating representation as a visual or narrative problem rather than a consequence of deeper inequities embedded within LIS education, labor conditions, and professional practice. By combining structural reform with individual and educational action, such as inclusive programming, critical AI literacy, mentorship, and advocacy, librarians and LIS educators can meaningfully leverage their agency to challenge stereotypes, expand professional imaginaries, and shape more equitable representations across human and machine-mediated contexts.

Conclusion

This study examined how generative AI depicts librarians, showing that while dominant stereotypes persist—primarily white, female, and youthful portrayals—important variations emerge across models, versions, and languages. These findings reveal how occupational identities are encoded into generative systems, reflecting entrenched social biases while leaving room for alternative imaginaries.

By centering librarianship as a case study, we contribute to broader debates about representation in AI. Our analysis demonstrates that stereotypes embedded in training data are not merely technical artifacts but sociocultural constructs with real consequences for how professions are perceived. Librarianship, already grappling with questions of visibility, diversity, and professional identity, is especially vulnerable to these distortions.

Future research should expand beyond librarianship to compare other knowledge-intensive professions, investigate the role of user prompting in shaping outputs, and engage directly with practitioners to understand how AI-generated imagery influences professional self-conception. For designers of generative AI, the challenge is to develop systems that better reflect the diversity of real-world occupations and avoid perpetuating narrow, exclusionary stereotypes.

Ultimately, the case of librarianship illustrates both the risks and possibilities of generative AI: left unchecked, it reproduces inequities, but with critical intervention, it can also serve as a tool for imagining more inclusive futures.

Acknowledgements

About the authors

Viviane Ito is a doctoral student at the University of North Carolina at Chapel Hill, USA, and a graduate affiliate at CITAP. Her research focuses on human-computer interaction and AI search. She can be contacted at itovivi@unc.edu

Lyric Grimes is a doctoral student at the University of North Carolina at Chapel Hill, USA. Her research focuses on harm, exclusion, and the patron experience in library spaces. She can be reached at lyricly@email.unc.edu

References

Adams, K. C. (2000). Loveless frump as hip and sexy party girl: A re-evaluation of the old-maid stereotype. The Library Quarterly, 70(3), 287-301.

American Library Association. (2006, July 26). Core values of librarianship. https://www.ala.org/advocacy/advocacy/intfreedom/corevalues

Annoyed Librarian. 2007. 'Take the 'Hip' Librarians, Please.' Annoyed Librarian (blog), July 9, http://annoyedlibrarian.blogspot.com/2007/07/take-hip-librarians-please.html.

Benjamin, R. (2019). Race after technology: Abolitionist tools for the new Jim code. John Wiley & Sons.

Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2023). Easily Accessible Text-to-Image Generation Amplifies Demographic Stereotypes at Large Scale. 2023 ACM Conference on Fairness, Accountability, and Transparency, 1493–1504. https://doi.org/10.1145/3593013.3594095

Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. Advances in neural information processing systems, 29.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. Science, 356(6334), 183-186.

Carnegie, T. A. M., & Abell, J. (2009). Information, Architecture, and Hybridity: The Changing Discourse of the Public Library. Technical Communication Quarterly, 18(3), 242–258. https://doi.org/10.1080/10572250902947066

Cheong, M., Abedin, E., Ferreira, M., Reimann, R., Chalson, S., Robinson, P., ... & Klein, C. (2024). Investigating gender and racial biases in DALL-E mini-images. ACM Journal on Responsible Computing, 1(2), 1-20.

Clinton, H. R. (2017, June 27). Closing general session [Speech transcript]. American Library Association Annual Conference, Chicago, IL. https://americanlibrariesmagazine.org/wp-content/uploads/2017/06/HRC-Transcript.pdf

Currie, G., Currie, J., Anderson, S., & Hewis, J. (2024). Gender bias in generative artificial intelligence text-to-image depiction of medical students. Health Education Journal, 83(7), 732–746. https://doi.org/10.1177/00178969241274621

D'Amour, A., Heller, K., Moldovan, D., Adlam, B., Alipanahi, B., Beutel, A., ... & Sculley, D. (2022). Underspecification presents challenges for credibility in modern machine learning. Journal of Machine Learning Research, 23(226), 1-61.

Data USA. (2017). Retrieved January 8, 2026, from https://datausa.io/profile/soc/librarians

Department for Professional Employees, AFL-CIO. (n.d.). Library professionals: Facts and figures. Retrieved February 24, 2025, from https://www.dpeaflcio.org/factsheets/library-professionals-facts-and-figures

Eberhart, G. M. (2019, April 13). Why Being Bad Is Good | American Libraries Magazine. American Libraries Magazine. https://americanlibrariesmagazine.org/blogs/the-scoop/acrl2019-vocational-awe-why-being-bad-is-good/

Escalante, I., Mallmann Pereira Souto, P., & Rodrigues De Souza Coutinho, L. (2021). O impacto do estereótipo de gênero sobre a mulher bibliotecária do século XXI no Brasil. Revista Brasileira de Educação em Ciência da Informação, 8. https://doi.org/10.24208/rebecin.v8i.243

Espinal, I. (2001). A new vocabulary for inclusive librarianship: Applying whiteness theory to our profession. In Castillo-Speed L, REFORMA National Conference Publications Committee, eds. The power of language/el poder de la palabra: selected papers from the Second REFORMA National Conference. Englewood, CO: Libraries Unlimited (pp. 131-49).

Ettarh, F. (2018, January 10). Vocational Awe and Librarianship: The Lies We Tell Ourselves – In the Library with the Lead Pipe. In the Library with the Lead Pipe. https://www.inthelibrarywiththeleadpipe.org/2018/vocational-awe/

Ferreira, M. M. (2020). Bibliotecários e relações de gênero no Brasil e Portugal. ConCI: Convergências em Ciência da Informação, 2(3), 298–322. https://doi.org/10.33467/conci.v2i3.13713

Fujita, M. S. L., Agustín-Lacruz, M. D. C., & Terra, A. L. (2018). PERFIL E FORMAÇÃO DO PROFISSIONAL EM BIBLIOTECAS ESCOLARES NO BRASIL, ESPANHA E PORTUGAL.

Gambrell, D., & Brennan, A. (2014). Librarians and felines: A history of defying the "cat lady" stereotype. The librarian stereotype: Deconstructing perceptions and presentations of information work, 175-184.

Garcês-Da-Silva, F., & Saldanha, G. (2023). Brazilian Black Librarianship. Journal of Critical Library and Information Studies. https://doi.org/10.24242/jclis.v4i1.165.

Harris, M. (2023). Palo Alto: a history of California, capitalism, and the world. Hachette UK.

Hess, D. J., & Sovacool, B. K. (2020). Sociotechnical matters: Reviewing and integrating science and technology studies with energy social science. Energy Research & Social Science, 65, 101462. https://doi.org/10.1016/j.erss.2020.101462

Jacobsen, T. L. (2004). The Class of 1988: Fifteen Years after Library School, What Do Their Careers Say about Yours?. Library Journal, 129(12), 38.

Kneale, R. (2009). You don't look like a librarian: Shattering stereotypes and creating positive new images in the Internet age. Information Today.

Kotek, H., Dockum, R., & Sun, D. (2023). Gender bias and stereotypes in Large Language Models. Proceedings of The ACM Collective Intelligence Conference, 12–24. https://doi.org/10.1145/3582269.3615599

Kumasi, K. (2013). "The library is like her house": Reimagining youth of colour in LIS discourses. In A. Bernier (Ed.), Transforming young adult services: A reader for our age (pp. 103–113). Chicago, IL: ALA Editions.

Muller, L. K. P., & Martins, C. W. S. (2019). Uma profissão feminina, mas não feminista? Representatividade de gênero na gestão dos Conselhos Regionais de Biblioteconomia no Brasil. Revista Brasileira de Biblioteconomia e Documentação, 15, 92–111.

Naik, R., & Nushi, B. (2023). Social Biases through the Text-to-Image Generation Lens. Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society, 786–808. https://doi.org/10.1145/3600211.3604711

Natarajan, V. (2017). Nostalgia, cuteness, and geek chic: Whiteness in Orla Kiely's Library. Topographies of whiteness: Mapping whiteness in library and information science, 121-41.

National Public Radio. (2023, October 9). HBCUs have been underfunded by $12 billion, federal officials reveal. NPR. https://www.npr.org/2023/10/09/1204614576/hbcus-have-been-underfunded-by-12-billion-federal-officials-reveal

Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. In Algorithms of oppression. New York University Press.

Pagowsky, N., & DeFrain, E. (2014). Ice Ice Baby: Are librarian stereotypes freezing us out of instruction? In the Library with the Lead Pipe. https://www.inthelibrarywiththeleadpipe.org/2014/ice-ice-baby-2/

Rabey, Melissa. 2007. "A Hipper Crowd of Shushers." Pop Goes the Library (blog), July 8, http://www.popgoesthelibrary.com/2007/07/free-yourself-from-stereotypes.html.

Radford, G. P., & Radford, M. L. (2001). Libraries, librarians, and the discourse of fear. The Library Quarterly, 71(3), 299-329.

Santamaria, M. R. (2020). Concealing white supremacy through fantasies of the library: Economies of affect at work. Library Trends, 68(3), 431-449.

Sarker, S., Chatterjee, S., Xiao, X., & Elbanna, A. (2019). The Sociotechnical Axis of Cohesion for the IS Discipline: Its Historical Legacy and its Continued Relevance. Management Information Systems Quarterly, 43(3), 695–719.

Seale, M. (2008). Old maids, policeman, and social rejects: Mass media representations and public perceptions of librarians.

Schneider, K. G. 2007. "To Be Cool Is to Be Young and Male?" Free Range Librarian (blog), July 8, http://freerangelibrarian.com/2007/07/08/to-be-cool-is-to-be-young-and-male/.

Schlesselman-Tarango, G. (2017). How cute! Race, gender, and neutrality in libraries. Partnership: The Canadian Journal of Library and Information Practice and Research, 12(1).

Souza, D. (2025, April 20). Censo: 63% das escolas brasileiras não têm biblioteca. https://www.metropoles.com/brasil/censo-escolar-1-3-das-escolas-biblioteca

Stoddart, R. A., & Lee, A. R. (2005). Paradoxical views of "librarian" in the rhetoric of library science literature: A fantasy theme analysis. The Georgia Library Quarterly, 5.

Tanus, G., & Sánchez-Tarragó, N. (2020). Activities and challenges of brazilian university libraries during the covid 19 pandemic. ACIMED, 31, 1-35.

U.S. Bureau of Labor Statistics. (2024). Librarians and library media specialists. In Occupational Outlook Handbook. https://www.bls.gov/ooh/Education-Training-and-Library/Librarians.htm

Vieira, K., & Karpinski, C. (2019). The historical and epistemological relations between Librarianship and Information Science in the Brazilian scientific production. Transinformação. https://doi.org/10.1590/2318-0889201931E180043.

White, S. (2025, March 14). Long legacy of producing Black librarians continues at NCCU: 'Didn't see myself in the library'. HBCU News. https://hbcunews.com/2025/03/14/long-legacy-of-producing-black-librarians-continues-at-nccu-didnt-see-myself-in-the-library/?utm_source=chatgpt.com

Yontz, E. (2003). Librarians in children's literature, 1909-2000. The Reference Librarian, 37(78), 85-96.