

Help or Hazard? Patrons' Checkout History Retention Choices and Relations to Trust and Campus Role

Craig E. Smith and Kenneth J. Varnum

We explored the motivations of 588 university library patrons who chose to either have the library retain their checkout histories or not. We also examined associations between checkout history choice and both general data privacy concerns, and campus role. Over 90% opted to retain their histories. This group, compared to the no-retain group, had fewer privacy concerns, had greater trust in the library relative to the university, and were less likely to be librarians, archivists, or curators. We discuss how these findings add to the literature on privacy concerns of library users, and their possible implications for privacy policies.

Introduction

In recent years, a great deal of conversation and research has centered on the potential benefits and risks of data collection and analysis by organizations making their services available via the Internet (e.g., Gutierrez, 2023). In the public sphere, these conversations often focus on social media sites such as Facebook, X (formerly Twitter), and TikTok, online marketplaces such as Amazon, and online platforms where users interact with a vendor or with each other, leaving digital footprints behind. Many commercial sites have offered (sometimes limited) ways for individuals to opt out of some data collection or to see what data the site has collected. Apple's iOS, which powers hundreds of millions of smartphones, has a setting that can block third-party user tracking completely, causing some advertisers to rethink their approaches to customizing and optimizing campaigns (Loveless, 2022). Public perceptions of the risks and benefits of these data collection practices run across a wide spectrum, from those who profess no particular concern about what data is collected, to those who are strongly opposed to leaving any trace of where they travel online (McClain et al., 2023).

Libraries have long been concerned that records of library transactions, if stored and associated with individuals, have the potential to harm individuals (Matz, 2008). These concerns were elevated following the enactment of the USA Patriot Act in 2001. Concerns about the disproportionate harm that could be caused to individuals from at-risk communities (e.g., those from minority populations and those with undocumented immigrant status) have led

* Craig E. Smith, Assessment Specialist, Dean's Office, University of Michigan Library, email: craigsm@umich.edu; Kenneth J. Varnum, Senior Program Manager and Discovery Strategist, Library Information Technology, University of Michigan Library, email: varnum@umich.edu. ©2025 Craig E. Smith and Kenneth J. Varnum, Attribution-NonCommercial (<https://creativecommons.org/licenses/by-nc/4.0/>) CC BY-NC.

libraries to, in general, proactively remove library transactions that could identify both the resource and the individual using it from library systems. To this end, patron privacy is addressed explicitly in the American Library Association's Code of Ethics (2021). The present study explores how library patrons feel about their checkout data being stored, the choices patrons make when given a choice to retain their history or not, and the reasoning supplied by patrons for their choices. In an era in which many academic libraries are retaining and using patron data for endeavors such as learning analytics studies, and in which people are increasingly accustomed to having their data held by a range of institutions, knowing more about how patrons weigh the benefits and risks of library data retention is important. The findings of this and similar studies can inform library decisions about whether to retain user data, and how to communicate and offer choices about data retention.[†]

Background

In higher education, student concerns and beliefs surrounding choice in data collection have been studied by, among others, the IMLS-funded Data Doubles project (<https://datadoubles.org>). Among this project's findings is that most students are aware that their use of the library generates data that are not wholly private; only 22% of students considered their use of physical library resources to be completely private (Asher, 2022). Nonetheless, students are typically not alert to the wide range of library interactions—for example, with physical and electronic materials—that generate or store data about them that is personally identifying (Asher, 2022). At the same time, regulations such as the European Union's General Data Protection Regulation (GDPR) directed that web content providers explicitly seek permission from users before conducting any tracking of their activities. This significantly raised many users' awareness of the fact that data are tracked (MacDonald & Klebe, 2018).

The Data Doubles project also found that students typically accepted university practices such as analytics and data mining, but wanted these activities to be paired with more transparency and the ability to consent versus opt out (Jones, 2023). These findings mirror similar findings from earlier research (e.g., Ifenthaler & Schumacher, 2016). Students expressed trust in librarians, more than in other campus actors, in part because students perceived librarians as having relatively little agency in students' educational outcomes (Jones, 2023). Other research has also found that librarians are highly trusted, relative to people in most other professions (Ipsos MORI, 2021; Lockwood & Ritter, 2016).

When it comes to libraries' collection and use of transaction data, most of the literature focuses on this in the context of library learning analytics. Library learning analytics can be considered within the broader scope of campus learning analytics activities. Some academic libraries have explored the relationships between the resources and services they provide students and student success (ACRL, 2010; Oakleaf, 2021). We note that a recent metanalysis indicated that such studies have found little to no statistical relationship between academic library use and academic outcomes (e.g., GPA and retention), and these types of studies are not able to shed light on causality (Robertshaw & Asher, 2019). In contrast to these studies, ours focused on individual concerns and feelings about one form of library data collection rather than using library data to make inferences about the impact of academic libraries.

[†] This study was reviewed by the U-M IRB (HUM#00200146) and was determined to be exempt due to the lack of serious risk and vulnerable populations.

Collecting identified user data to build library-specific services has received much less attention, compared to studies focused on analytics. Another Data Doubles project output, “Transparency and Consent: Student Perspectives on Educational Data Analytics Scenarios,” highlights the importance of consent but focuses more on research-oriented processes, rather than service-oriented (Jones et al., 2023). The present study takes this next step, asking how patrons feel about their data being retained to support a library service.

Additional Context for the Present Study

Until early 2016, the University of Michigan Library was typical of most public and academic libraries (Harper & Oltmann, 2017): as soon as a book was returned, the data connecting the library user and the item that was circulated were deleted. In 2016, the library dean advocated for greater library involvement in campus-wide learning analytics efforts and led a process to update the library’s privacy policy and practices. After extensive discussions and debates about data and privacy, the decision was made to preserve the connection between users and the items they circulated; some library employees were supportive of this decision, and some were not.

This new practice allowed the library to provide a checkout history service for all users. Behind authentication, this checkout history was available to each individual and listed the items each person had checked out. At the time, technological barriers prevented the library from offering an opt-in or opt-out mechanism, and the service was automatically activated for all members of the university community who checked out an item. Interestingly, this led to a situation in which there was a considerable population of users who had a checkout history maintained but were not aware of it.*

The University Library changed its library management system in 2021. This technology update provided, for the first time, an opportunity to enable a mechanism allowing library users to opt in or out of the checkout history service. Conversations about the importance of allowing such a choice were initially led by the Library Diversity Council, a library-wide group concerned with questions of diversity, equity, and inclusion. The group expressed concerns about the potential harm an “always on” checkout history could create, particularly for marginalized and underrepresented communities across the campus.† In response to these concerns, and with the new library services platform enabling such an approach, the library decided to give users the ability to opt in or out of checkout history data collection through their library account page. The authors viewed this change of service model as a unique opportunity both to give our library’s users more power over a subset of data collected about them, but also to better understand the motivations individual users might have when it came to making that choice.

Because of our past decision to create a checkout history for all users, the library found itself in a somewhat challenging situation. Individuals who had checked out an item between 2016–2021 had a checkout history (regardless of whether they knew it existed or wanted it), and those who did not check out an item in that time period — perhaps because their academic work did not lend itself to use of physical library materials or they were new to campus — did not. To avoid a situation in which the library arbitrarily deleted previous checkouts for all users, including those for whom it was a useful service, a split process was offered.

* This also underscores the challenges of educating large user groups about library privacy policies and changes to those policies.

† For more on demographic-specific privacy concerns see, for example, Mathson & Hancks, 2007.

Starting with the switch to our new library management system in August 2021, if an individual had a checkout history, it was preserved and remained active. However, individuals without items in their checkout history would not have future checkouts added. Next, all individuals were given the option to change their current checkout history status, (i.e., to delete their history and stop future data collection if they had a checkout history, or to start collecting a checkout history if they did not have one). Library users could also download a copy of their checkout history at any time, and were invited to do so before deleting it, if they chose to opt out.[‡]

Of the 2,045 individuals who made an explicit choice about their checkout history between July 2021 and November 2022, 1899 (92.9%) opted in, while the remaining 146 (7.1%) opted out. The overall breakdown between choices was relatively consistent between those who had a checkout history before making the choice and those who did not: 91.7% of those with a checkout history chose to keep it, and 93.6% of those without a checkout history chose to start one (see Table 1).

TABLE 1
Frequencies and Percentages of Users Opting In or Opting Out of Checkout History
Between July 2021 and November 2022

Decision	Had Checkout History		Did Not Have History		Overall	
Opted In	738	91.68%	1,161	93.63%	1,899	92.86%
Opted Out	67	8.32%	79	6.37%	146	7.14%
Totals	805	100.00%	1,240	100.00%	2,045	100.00%

This situation presented an opportunity for the quasi experiment (Stevenson, 2020) reported here. The present study explored the benefits and concerns that were considered by library users who opted to have a checkout history and compared them to the considerations of the group that opted out. We also tested whether some groups of library users (e.g., as a function of campus role, demographics, etc.) were more likely to maintain a checkout history than others.

Methods

Participants

Invitations to participate in the study were sent to those who made a choice about their checkout histories (i.e., retain/start a checkout history versus delete/not start a checkout history). The survey was sent in five waves; each wave was timed to be within three months of when members of that wave had made a decision about their checkout histories. The goal was to ask about a decision that was in respondents' recent memory; a survey question to check on respondent memory indicated that only three people who started a survey did not remember the choice they had made about their checkout history (this was < 1% of all opened surveys).

[‡] The service checkout history described in this article is user-focused and data collected are explicitly for the individual's benefit, the principles of informed data collection and permission apply. The University Library's privacy statement (<https://www.lib.umich.edu/about-us/policies/library-privacy-statement>) describes how the library may use data, with more detail on data collection and use maintained by the central Information Technology organization, dubbed ViziBLUE, outlines data policies and uses across campus (<https://safecomputing.umich.edu/viziblue/library-data>).

TABLE 2
Counts of Survey Invitees and Participants

Survey Variant	Survey Wave	# of People Invited	# of Surveys Submitted	# Valid Surveys (Response Rate)
1	1	244	96	93 (38%)
1	2	431	158	155 (36%)
2	3	702	186	182 (26%)
2	4	226	65	61 (29%)
2	5	420	106	97 (23%)
Totals		2,023	611	588 (29%)

Note: In three submitted surveys, respondents had no memory of making a choice about checkout history and were thus not included in the count of valid surveys in the rightmost column above. The remaining non-valid surveys were characterized by large numbers of skipped questions.

Table 2 presents information about the number of library users who were invited to the study, the number of people who started a survey and submitted a survey, and the number of individuals whose survey responses were included in the data we present in this paper. Note that Table 2 refers to two variants of the survey; the only difference was that, in variant 2, we asked an additional question about whether respondents had vacillated in their checkout history decision.

Of the 588 participants in our final sample, 536 had elected to start or maintain a checkout history (91.2%), while 52 did not want to start or retain such a history (8.8%). These sample-level percentages are similar to those in the population of those who were invited to participate: 92.5% wanted a checkout history, 7.5% did not, respectively. Thus, regarding this important variable, our group of study participants was not different from the larger group of people who made choices about their histories.

Information about participants' primary campus role was obtained via the survey, as were data regarding whether respondents self-identified as a member of one or more groups that have been traditionally underrepresented and/or marginalized on university campuses. Further, information about many respondents' ethnicity and sex was available via the campus data warehouse. These data are quite flawed due to the omission of gender identities outside of the cis-male/cis-female binary, and due to the non-nuanced manner by which race and

TABLE 3
Frequencies: Participant Campus Role and Demographic Information

Primary campus role (self-report)	N	Member of marginalized group (self-report)	N	Person of color (univ. data)	N	Participant sex (univ. data)	N
Student	210	Yes	192	Yes	136	Female	297
Faculty	117	No	229	No	345	Male	234
LACs	46	Unknown	167	Unknown	107	Unknown	57
Staff	127						
Unknown	88						

Note: LACs are librarians, archivists, and curators. Unknown values are a result of skipped survey questions and/or missing university-held data. Participant sex is recorded as binary in the university's data warehouse; more inclusive data on gender identity was not available.

ethnicity data are collected. Nonetheless, Table 3, offers a useful snapshot of the identities of our participants; these data were used in some of the analyses we report below.

Survey Instrument and Administration

We chose to collect data via a survey because it allowed us to invite all possible participants into the study, and because we had only a small number of open-ended questions. The survey was developed by the authors and was then vetted by library colleagues and pilot-tested by a small number of library users. Once finalized, the survey questions were delivered to participants via Qualtrics online survey software.

The survey questions are presented in Appendix A. The first part of the survey asked about participants' campus roles and checked on participants' memories of opting in or out of having a checkout history; three participants were taken directly to the end of the survey because they did not remember their choice.

Next, participants were asked, in an open-ended fashion, why they made the opt-in vs. opt-out choice that they did. In survey waves one and two (see Table 2), the question was simply, "What reason(s) led you to make that choice?" A second variant of the survey was used for waves three to five, in which that initial question was retained, and a second question was added: "Did you give serious consideration to the opposite choice?" If participants in waves three to five responded affirmatively to this question, they were then asked, "What were those other considerations?" This was the only difference between the two variants of the survey; as such, aggregated results from waves one to five are largely reported together in the Results section.

The survey then asked questions about behavior and cognition. First, respondents whose checkout histories had been maintained for them prior to the survey were asked if they had been aware of their checkout histories and, if so, how often they looked at them. We were interested in whether people who used their checkout histories would reason differently about keeping or deleting their checkout histories. We then asked about participants' level of concern about data mismanagement at the levels of: 1. the University of Michigan library; 2. the University of Michigan more broadly;* and 3. internet-based companies such as Amazon, Apple, Netflix, etc. We expected that higher levels of concern with the library's and university's data management practices to be associated with a greater likelihood of opting out of having a checkout history. Further, we also expected that, on average, respondents would have the lowest levels of concern about data mismanagement regarding the library, based on previous research showing that library staff are highly trusted relative to most professions (Ipsos MORI, 2021; Lockwood & Ritter, 2016).

Finally, respondents were asked if they were a member of one or more groups that have been underrepresented or marginalized on college and university campuses; examples given included being underrepresented or marginalized on the basis of race, gender identity, disability status, sexual orientation, socioeconomic status, etc. We hypothesized that participants who are members of groups that have historically been underrepresented or excluded when data management practices were created might have less trust in such practices. The last survey question was open-ended and gave space for respondents to share other comments about library data.

* The University of Michigan suffered a large-scale data breach in August 2023. Our results reflect opinions before that event happened.

To protect the data of all survey invitees, both those who completed the survey and those who did not, study data were stored on password-protected computers and restricted, HIPAA-compliant cloud storage folders.

Coding of Qualitative Data

We employed an inductive approach to the qualitative data we collected in the survey, seeking coherent themes in the responses to each open-ended question without assumptions or an existing theoretical framework (Charmaz, 2003; Urquhart, 2013).

As a first step, the two authors repeatedly read and discussed the responses to the open-ended survey questions, making notes about essential themes that emerged in response to each. After gaining familiarity with the data, we created a provisional coding scheme for each question. The research team reviewed the open-ended responses again and paid attention to whether the coding system for each question needed to be adjusted to accommodate previously missed themes. The final coding schemes for each open-ended question are presented in Appendix B.

The two research team members then independently used the coding systems to code the open-ended data from survey waves one and two. Following this initial round of coding, the two sets of codes were compared. For each code in the scheme, the measure of interrater reliability (kappa) was above the commonly used threshold of 0.70 (the kappas ranged from 0.76 – 1.00). Discrepancies were easily resolved through discussion. After establishing that the coding system could be used reliably, the rest of the coding was completed by the two authors together; here again, agreements were very common, and disagreements were easily resolved. The themes that emerged from the coding process are reported on in the Results section.

Results

Checkout History Choice and Campus Role

A total of 536 respondents opted to have a checkout history, with 239 starting one, and 297 electing to keep their existing history. A total of 52 respondents had chosen to either not start a checkout history (23) or to delete an existing one (29). These two groups were of central interest. We first analyzed whether the opt-in/opt-out choice was associated with campus role (see Table 4).

TABLE 4				
Checkout History Choice as a Function of Broad Campus Role				
Choice	Students	Faculty	LACs	Staff
Wanted History	95.2% (n = 200)	91.5% (n = 117)	76.1% (n = 35)	89.0% (n = 113)
Did not want history	4.8% (n = 10)	8.5% (n = 10)	23.9% (n = 11)	11.0% (n = 14)
Note: People who identified as alums, community members, unaffiliated researchers, and 'other' were excluded from this analysis due to cell size considerations.				

An omnibus chi-square analysis indicated that there was at least one significant difference between the four campus role groups regarding their checkout history choice $\chi^2(3, 500) = 17.76$, $p < .001$, $\phi = .19$. Pairwise tests were used to clarify this result, and the following significant differences were found:

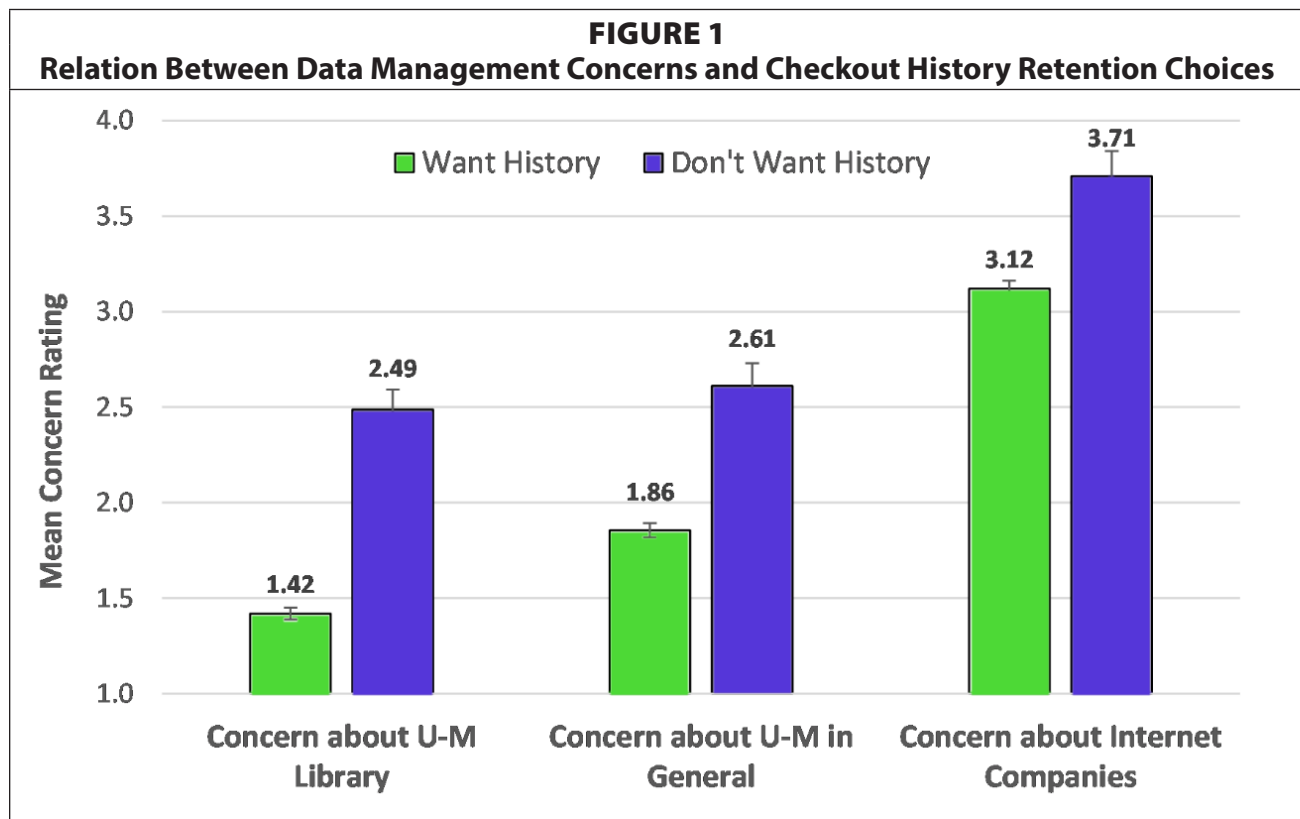
- Students (95.2%) were more likely than staff members (89.0%) to want a checkout history ($\chi^2 (1, 337) = 4.69, p = .03$).
- Students (95.2%) were more likely than librarians, archivists, and curators (LACs; 76.1%) to want a checkout history (Fisher's exact test $< .001^*$).
- Faculty members (91.5%) were more likely than LACs (76.1%) to want a checkout history ($\chi^2 (1, 163) = 6.95, p = .008$).
- Staff members (89.0%) were more likely than LACs (76.1%) to want a checkout history ($\chi^2 (1, 173) = 4.54, p = .03$).

The clear trend that emerged in this set of analyses was that—although three quarters of librarians, archivists, and curators wanted a checkout history—LACs were significantly less likely than the other three groups to want their checkout data retained.

Checkout History Choice and Concerns about Data Management

Participants were asked about their concerns about data mismanagement in the library, at the University of Michigan in general, and by internet corporations (scored as: 1 = not at all concerned; 2 = mildly concerned; 3 = moderately concerned; 4 = very concerned). Choices about checkout history were analyzed in relation to these three questions about data mismanagement using a 2 (choice groups) $\times$ 3 (target of concern) mixed-measures ANOVA, with choice as the between-subjects variable and the concern questions as the within-subjects variables (see Figure 1).

Figure 1 highlights that concern about data mismanagement was stronger in the group that didn't want a checkout history, compared to the group that did. The figure also indicates



* Fisher's exact test was used due to low expected cell counts ($n < 5$). Here and elsewhere, measures were taken to use appropriate analyses for the very uneven group sizes in our sample, with those electing to have a checkout history far outnumbering those who did not.

that data mismanagement concerns are least strong for the library, stronger for the University of Michigan in general, and stronger still for internet companies. This is in accord with the results of the ANOVA, which yielded significant main effects of both choice group ($F(1,468) = 67.15, p < .001, \eta^2 = .13$) and target of concern ($F(2,936) = 262.15, p < .001, \eta^2 = .36$).

Further, there was a significant choice group $\times$ target of concern interaction ($F(2,936) = 2.24, p = .002, \eta^2 = .01$). The nature of this interaction was clarified via analyses of simple effects. The results of these post hoc pairwise contrasts indicated that:

- The group that wanted to retain a checkout history rated their data mismanagement concerns as significantly lower than the opt-out group across all three questions (i.e., about the library, the University of Michigan, and internet corporations), all p -values $< .001$.
- The group that wanted a checkout history viewed the library with least concern ($M = 1.42$), the university with an intermediate level of concern ($M = 1.86$), and internet corporations with the most concern ($M = 3.12$); all differences were significant, and all p -values $< .001$.
- The group that did not want a checkout history viewed data mismanagement risks at internet companies with more concern ($M = 3.71$) than the University library ($M = 2.61$) or the University of Michigan in general ($M = 2.49$), both p -values $< .001$.

The group that did not want a checkout history, unlike the group that did, made no significant differentiation between the library and the university more broadly regarding data mismanagement risk. This led to the significant interaction effect reported above.

Checkout History Choice and Prior Use of Checkout History

We next examined potential links between checkout history choice and whether people were aware of and used their checkout histories prior to making their choice; these analyses were limited to people who had an existing history that they chose to retain or delete. Of the whole group of 286 people in this analysis, 186 (65%) had been aware of their checkout histories, compared to 100 (35%) who had not. Within the 'keep history' group, 65.0% were aware they had a history, and within the 'don't keep' group, 65.4% had this awareness. Thus, the groups were essentially the same regarding awareness, $\chi^2(1, 286) = 0.002, p = .97$. We also investigated whether awareness of a prior checkout history was associated with campus role; there were no differences that rose to significance in an initial chi square test, $\chi^2(3, 250) = 6.61, p = .09$.

Those who did have an awareness of their existing checkout histories were then asked how often they referred to it or used it in some way. We were interested in whether those who chose to keep their histories had made more use of them compared to those who did not to retain a checkout history. The frequency scale respondents used ranged from 1 (never) to 6 (about once a week or more); all the scale points were anchored with frequency labels, available in Appendix A. On average, those who wanted to retain their checkout histories used their histories 'no more than a few times' per year ($M = 2.27$), while the group that did not want their histories used them closer to 'never' ($M = 1.40$). Although neither group used their histories often, the group that wanted their histories retained had indeed made more use of them in the past, $t(237) = 4.81, p < .001$.

Associations with Aspects of Social Identity

We were also interested in whether members of groups that have been historically and are currently underrepresented and marginalized at colleges and universities would have different checkout history preferences and different levels of institutional trust, compared to those

who are often well represented on campus and in academic leadership positions (e.g., White people, straight people, cis-males, etc.). We found that 91.7% of respondents who identified as being in an underrepresented/marginalized group ($n = 192$) wanted a checkout history, and 93.0% of those not identifying as such wanted a history. This difference was not significant, $\chi^2(1, 421) = 0.27, p = .60$.

A 2 (broad social identity groupings) $\times$ 3 (target institution in question) mixed-measures ANOVA replicated the significant finding (previously reported) that level of trust depends significantly on the target institution (e.g., the library, the university, internet companies). There was no significant main effect of being underrepresented/marginalized or not ($p = .37$), nor was there a significant interaction effect ($p = .62$). Thus, concern about data privacy was similar across the two broad social-identity groupings, and the groups varied similarly as a function of the institution in question.

We then used the available institutional data to look specifically at the roles of race and sex (i.e., the binary variable in the University's data warehouse). First, although most people wanted a checkout history, respondents of color (98.5% of 136 people) were more likely than white respondents (88.1% of 345 people) to want a checkout history, $\chi^2(1, 481) = 12.99, p < .001$. When looking in a more nuanced way at race, the same pattern was found (see Table 5).

TABLE 5
Checkout History Choice as a Function of Race

Choice	Asian/Asian-American	Black	Hispanic	Native American	Multiracial	White
Wanted History	98.3% ($n = 59$)	100% ($n = 22$)	97.2% ($n = 36$)	100% ($n = 1$)	100% ($n = 17$)	88.1% ($n = 304$)
Did not want history	1.7% ($n = 1$)	0% ($n = 0$)	2.8% ($n = 1$)	0% ($n = 0$)	0% ($n = 0$)	11.9% ($n = 41$)

Note: Race data—including the terminology—were obtained from the University's data warehouse, and do not always reflect how people describe themselves when asked about their racial backgrounds.

Finally, there was not an association between the binary sex variable and checkout history choice. Those listed as female were as likely to want a checkout history (92.3% of 297 people) as those listed as male (89.7% of 234 people), $\chi^2(1, 531) = 1.02, p = .31$.

Reasoning about Checkout History Choice

Central to the present research were the open-ended questions about why people made the choices they did (asked in all five survey waves), and what consideration—if any—they gave to the opposite choice (asked in survey waves three to five). The process used to code the resulting data is described in the Methods section, and is further detailed in Appendix B.

As noted, all participants were asked why they made their checkout history choice, and 265 participants in data collection waves three through five were asked if they considered making the opposite choice. Of these 265, 31% ($n = 81$) reported that they gave serious consideration to both checkout history options. This included 45% (9 of 20) in the 'don't want history' group, and 29% (72 of 245) of the 'want history' group. This difference was not statistically significant, in part due to the reduced power associated with the small number of people who did not want a checkout history, $\chi^2(1, 265) = 2.12, p = .15$.

There was a difference as a function of campus role, however, about whether people gave serious consideration to both checkout history options during decision making. Among students in data collection waves three to five, 22.2% indicated that they seriously considered both options, compared to 36.7% of faculty members, and 53.8% of librarians, archivists, and curators (LACs), $\chi^2(2,161) = 7.50, p = .02$.^{*} In clarifying pairwise tests, a Fisher's exact test indicated that students (22.2%) were less likely than LACs (53.8%) to have considered both options carefully, $p = .037$. The faculty group occupied an intermediate position and did not differ significantly from either the student or the LAC group.

There were no significant differences as a function of race or gender with regard to giving serious consideration to both checkout history options or not.

Reasons Provided for Choice Made

Table 6 presents the themes that emerged from the coding of participants' open-ended responses about the bases of their checkout history choices.

TABLE 6 Frequencies of Considerations Underlying Checkout History Choices, as a Function of Choice Group		
Concerns considered during choice	Group: Keep checkout history	Group: Not keep checkout history
Data privacy concerns	0.6% (3 of 530)	44.2% (23 of 52)
Library shouldn't be keeping such data	0%	28.8% (15 of 52)
Don't want or need it, not useful	0%	25.0% (13 of 52)
Wasn't aware library was retaining these data	0.2% (1 of 530)	7.7% (4 of 52)
Lack of trust in the university	0%	5.8% (3 of 52)
Not doing much academic work at present	0%	1.9% (1 of 52)
Benefits considered during choice	Group: Keep checkout history	Group: Not keep checkout history
For use as a reading list (no mention of research)	49.2% (261 of 530)	0%
Simple desire to have it (no reason provided)	17.0% (90 of 530)	0%
Miscellaneous future uses	9.2% (49 of 530)	0%
For enabling future research	8.1% (43 of 530)	0%
Accessibility issues raised as part of reason	0.4% (2 of 530)	0%
Other issues considered during choice	Group: Keep checkout history	Group: Not keep checkout history
Response was uncodable	10.8% (57 of 530)	7.7% (4 of 52)
No privacy concerns in this situation	4.9% (26 of 530)	1.9% (1 of 52)
Expect this type of service from libraries	0.2% (1 of 530)	0%
Note: The question about issues underlying one's choice was asked in all five waves of data collection. Formal comparisons between the two choice groups—regarding the likelihood of providing each type of response—were not conducted due to the clearly group-dependent nature of the responses.		

^{*} Staff were not included in this analysis due to problems with small cell sizes.

As is evident in Table 6, respondents in the ‘keep’ and ‘don’t keep’ groups each considered very different issues when asked about why they made the choices they did. Common reasons for choosing to not keep a checkout history included: concerns about data privacy (44%), the opinion that the library should not be retaining such data (29%), and the lack of a personal or professional need for a checkout history (25%). Conversely, relatively common reasons for electing to maintain a checkout history included: its utility as a reading list (49%), a generic desire to have it (17%), and its utility in informing future research activities (9%).

As noted, in data collection waves three to five we asked whether people gave serious consideration to the checkout history option they didn’t ultimately choose; those who answered affirmatively were then asked what those other considerations were. This new question was added because we were concerned that the approach to question-asking used in data collection waves one and two might be obscuring a more nuanced decision-making process. The results in Table 7 highlight that some people did indeed consider both sides of the issue prior to making their choices.

TABLE 7
Frequencies of Alternative Considerations During Checkout History Choicemaking, as a Function of Choice Group

Issues mentioned when asked about considerations of alternative choice	Group: Keep checkout history	Group: Not keep checkout history
Concerns about privacy (without elaboration)	8.5% (26 of 307)	0%
Had privacy concerns related to specific materials being checked out	2.9% (9 of 307)	0%
Had privacy concerns that were outweighed by expected benefits of checkout history	2.3% (7 of 307)	0%
Respondents unsure they would ever review their histories	2.0% (6 of 307)	0%
Concern about data breach (e.g., hacking)	1.6% (5 of 307)	0%
Unsure how the checkout history feature works	1.6% (5 of 307)	0%
Concern about government access to records	1.3% (4 of 307)	0%
Had privacy concerns but trusted library to protect data	0.7% (2 of 307)	0%
Simple desire to have it (no reason provided)	1.0% (3 of 307)	11.1% (3 of 27)
For use as a reading list (no mention of research)	0.3% (1 of 307)	7.4% (2 of 27)
Response was uncodable	2.0% (6 of 307)	3.7% (1 of 27)
Note: Unique responses mentioned by single individuals not included in Table 7. Formal comparisons between the two choice groups—regarding the likelihood of providing each type of response—were not conducted due to the clearly group-dependent nature of the responses.		

Particularly worthy of mention is that, in the ‘keep history’ group, a small but notable number of people indicated that they did fret about data privacy concerns in one way or another prior to making their choice. Thus, for many people the choice was straightforward, while for some it was indeed a relatively nuanced decision-making process.

Discussion

We explored how 588 patrons of a large, public university library thought and felt about li-

library data privacy, in the context of their decisions to have the library either maintain or delete their checkout history records. The situation in which library patrons were offered a chance to opt in or out of checkout history storage provided a unique opportunity to ask what types of thoughts, concerns, and emotions accompany such a decision.

When patrons were offered the chance to have a checkout history, versus not having a history, over 90% opted to have their circulation data retained. Members of this 'Keep' group, compared to the 'Delete' group, in general:

- were less concerned about privacy;
- expressed greater trust in the library versus the central university as stewards of their data;
- had used their histories more in the past (of those who had them); and
- were less likely to be in the campus role of librarians, archivists, curators.

These findings dovetail with other recent research showing that students, for example, are typically comfortable with their library data being retained and used for research, though not without a desire to have more transparency in data collection practices (Jones, 2023). Indeed, most of the students in the present study—and most in other campus roles as well—were comfortable enough with library data collection that they actively chose to have the library retain their checkout history data, despite the rather rare use of these histories on average.

People who wanted their checkout histories often reasoned that the data would come in handy in the future as a reading list record, or as an aide for future work, and most did not give deep thought to concerns like data privacy, although 8.5% did. Conversely, those who did not want their checkout histories described having data privacy concerns, the feeling that this was not proper practice for libraries, and the lack of their need for such data. Most of this 'don't keep' group did not describe wavering in their decisions, though a small percentage did consider the potential benefits of retaining a checkout history.

One of the more fascinating findings in the present study was that, although most people opted to retain a checkout history, the group of LACs (librarians, archivists, and curators) was significantly less likely to make that choice. This raises questions about the dynamics at play when a library employee feels more concern about a patron's data than does the patron themselves. Such a dynamic is ripe for future study. For example, does the small but significant difference between LACs and patrons—regarding views on data privacy—ever result in library data policies that are out of step with what many patrons want or need? Additionally, while we might be able to infer the role of age as an important variable when comparing faculty and students, future research is needed to formally measure and test this.

Future research could also explore the potential role that 'parentalism' (i.e., a non-gendered version of 'paternalism') plays in the data privacy decisions made by libraries. Parentalism involves restricting or controlling some aspect of others' experiences with the goal of serving their best interests (e.g., when states require all motorcyclists to wear helmets) (Carney et al., 2023). It could be the case that some libraries, or library employees, push for data policies that are viewed as best for patrons in the long term, even if it means denying a desire in the present (e.g., the desire for a checkout history). Studies of altruism have found that in some contexts, people do indeed engage in helping behaviors not only to satisfy the immediate needs/desires of others, but also to ensure the longer-term wellbeing of those others (Jacobsson et al., 2007). This can involve, for example, not engaging in helping behavior in the present when the helping may lead to longer-term negative consequences (Sibicky et al., 1995). Studies on parentalism in library policymaking could

add novel and important findings to both the literature on helping and altruism, and the literature on library data privacy.

Future research in this area could also consider various models of reasoning and decision making. Research has shown, for example, that when people are asked to make a morally- or emotionally charged judgment, intuition often drives the process, and explicit or rational reasoning about the judgment is often constructed after the fact (e.g., Haidt, 2001). Thus, it may be that the explanations provided by our participants regarding their past checkout history choices did not fully reflect the cognitive and emotional processes at play during the moment of opting in or out of data retention. Further, for some participants the decision may have been an emotionally charged one, while for others it may have been a simple and unemotional one; these potential differences in orientation toward data privacy decisions should ideally be considered in future studies of this topic and considered when creating and communicating about privacy issues.

Limitations

We acknowledge that this study was narrow and had shortcomings. First, we were very interested in the role that social identity might play in how people think about the privacy of their library data. We did ask a very broad question to ascertain whether respondents were members of marginalized and/or underrepresented groups (e.g., due to racism, gender discrimination, homophobia, etc.). We did not find much in the way of group differences regarding social identity, but this may have been because we did not ask nuanced questions about various aspects of respondents' identities. For example, it might be that people who LGBTQ+ are more concerned about the privacy of their library checkouts compared to heterosexual people; the approach of the present study did not allow for this level of exploration; future research in this area could benefit from stronger social identity questions.

In addition, we did not explore whether patron checkout history choices were related to other choices about data privacy, both inside and outside the library context. A study incorporating more than one data privacy decision will likely shed even richer light on how people think about their library data. Although we examined the association between trust in libraries and patron privacy choices, we did not explore causal relationships. It may be that libraries that offer patrons data privacy choices are more trusted than libraries that do not; such a question would be worth exploring, since the findings could have very practical implications.

Conclusion

What do the findings of the present study suggest about constructing or updating library data policies? First, the findings indicate that it may be beneficial for library policy makers to explicitly consider their own feelings and values related to data privacy alongside findings from recent studies that shed light on patron feelings and values. Noting where the orientations across library employees and patrons converge and diverge may enable policy makers to try new approaches to managing library data, and/or to communicate effectively and empathetically about the reasoning behind data policies. Second, the findings indicate that, when given some control over their library data, patrons typically have a very clear sense of their preferences. Library leaders should consider laying out the potential pros and cons of any data choices being offered to their patrons, so that patrons who might ordinarily make such a decision very quickly can slow down and consider future ramifications; this would

eliminate the need for libraries to engage in proactive parentalism while still allowing them to champion strong data privacy practices. Finally, our study indicated that many library patrons were unaware of their library's data policies and practices. Although campus communities are typically bombarded with messages from their colleges and universities, we strongly recommend that libraries find effective ways to alert their users to how their library data are managed, and what options are available to them (or not). This might involve one or more of the following: paying students to communicate about the library with their peers, liaison outreach to academic units, connecting with campus groups, using social media, placing messages on library websites, and creating welcome kits for those new to campus. Some of these communication efforts could also, of course, be paired with an explanation of library offerings and encouragement to use those offerings as well-informed, empowered patrons.

Acknowledgements

The authors are grateful to Monique Rio at the University of Michigan Library for providing the data used in this analysis. We also thank all those who took the time to participate in this study.

References

- American Library Association. (2021, July 21). *Professional Ethics*. <https://www.ala.org/tools/ethics>
- Asher, A., Briney, K. A., Jones, K. M. L., Regalado, M., Perry, M. R., Goben, A., Smale, M., & Salo, D. (2022). Questions of trust: A survey of student expectations and perspectives on library learning analytics. *Library Quarterly*, 92(2), 151–171. <https://doi.org/10.1086/718605>
- Carney, T., Bigby, C., Then, S.-N., Smith, E., Wiesel, I., & Douglas, J. (2023). Paternalism to empowerment: All in the eye of the beholder? *Disability & Society*, 38(3), 503–523. <https://doi.org/10.1080/09687599.2021.1941781>
- Charmaz, K. (2003). Grounded theory. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods* (pp. 81–110). Sage Publications, Inc.
- Gutierrez G., S. (2023, July 27). The Opt Out: Analog life is possible. But it'll cost you. *Popular Science*.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814–834. <https://doi.org/10.1037/0033-295X.108.4.814>
- Harper, L. M., & Oltmann, S. M. (2017). Big data's impact on privacy for librarians and information professionals. *Bulletin of the Association for Information Science & Technology*, 43(4), 19–23. <https://doi.org/10.1002/bul2.2017.1720430406>
- Ifenthaler, D., and Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics." *Educational Technology Research and Development* 64(5), 923–38.
- Ipsos MORI. (2021). Ipsos MORI Veracity Index 2021: Trust in professions survey.
- Jacobsson, F., Johannesson, M., & Borgquist, L. (2007). Is altruism paternalistic? *The Economic Journal* 117, 761–781. <https://doi.org/10.1111/j.1468-0297.2007.02049.x>
- Jones, K. M. L., Goben, A., Perry, M.R., Regalado, M., Salo, D., Asher, A.D., Smale, M.A., & Briney, K.A. (2023). Transparency and consent: Student perspectives on educational data analytics scenarios. *portal: Libraries and the Academy* 23(3), 485–515. <https://doi.org/10.1353/pla.2023.a901565>
- Lockwood, B. M., & Ritter, J. (2016). Maine state library: Trusted professionals survey 2016. Library Documents, Paper 101. http://digitalmaine.com/msl_docs/101
- Loveless, J. (2022, November 8). How does Apple's app tracking transparency framework affect advertisers? *Forbes*.
- MacDonald, S., & Klebe, S. (2018). GDPR: Understanding its impact and why library vendors rushed to comply. *Online Searcher*, 42(5), 24–44.
- Mathson, S., & Hancks, J. (2007). Privacy please? A comparison between self-checkout and book checkout desk circulation rates for LGBT and other books. *Journal of Access Services*, 4(3–4), 27–37. https://doi.org/10.1300/J204v04n03_02
- Matz, C. (2008). Libraries and the USA PATRIOT Act: Values in conflict. *Journal of Library Administration*, 47(3–4), 69–87. <https://doi.org/10.1080/01930820802186399>
- McClain, C., Faverio, M., Anderson, M., & Park, E. (2023). How Americans view data privacy, Pew Research Center. <https://www.pewresearch.org/internet/2023/10/18/how-americans-view-data-privacy/>

- Oakleaf, M. (2010). *Value of Academic Libraries: A Comprehensive Research Review and Report*. Association of College and Research Libraries. https://www.ala.org/acrl/sites/ala.org.acrl/files/content/issues/value/val_report.pdf
- Oakleaf, M., Varnum, K., Fransen, J., Nackerud, S., Brown, C., Mosbacker, B., & McCann, S. (2021). *Connecting libraries and learning analytics for student success*. <https://library.educause.edu/-/media/files/library/2020/12/classfinalwhitepaper.pdf>
- Robertshaw, M. B., & Asher, A. (2019). Unethical numbers? A meta-analysis of library impact studies. *Library Trends*, 68(1), 76–101. <https://doi.org/10.1353/lib.2019.0031>
- Sibicky, M. E., Schroeder, D. A., & Dovidio, J. F. (1995). Empathy and helping: Considering the consequences of intervention. *Basic and Applied Social Psychology*, 16(4), 435–453. https://doi.org/10.1207/s15324834basp1604_3
- Stevenson, J. F., (2020). Quasi-experimental designs, In P. Atkinson, S. Delamont, A. Cernat, J.W. Sakshaug, & R.A. Williams (Eds.), SAGE Research Methods Foundations. <https://doi.org/10.4135/9781526421036914289>
- Urquhart, C. (2013). *Grounded theory for qualitative research: A practical guide*. SAGE Publications.

Appendix A: Survey Instrument

[INTRO TEXT] Thank you for participating in our survey about your checkout history preferences. This survey will take about five minutes of your time. After data collection is complete, all identifiers will be removed from the data (e.g., your email address). When we report on the findings, we will be presenting aggregated data only. Your identity or name will not be connected to the data in any way. We will know only basic information about your campus role (e.g., faculty, student, staff).

As part of their review, the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board (IRB) has determined that this study presents no more than minimal risk, and is exempt from on-going IRB oversight (HUM00200146).

If you agree to participate, please click the next button below.

1. What is your primary role in relation to the University?
 - a. Undergraduate student
 - b. Graduate student
 - c. Faculty member (tenure, research, lecturer, or clinical tracks)
 - d. Faculty member (librarian, archivist, curator)
 - e. Staff member
 - f. Alumni
 - g. Community member
 - h. Researcher not affiliated with the University of Michigan
 - i. Other (please specify) _____
2. [FOR THOSE WHO ALREADY HAD A CHECKOUT HISTORY] Since mid-2016, the University of Michigan Library maintained a checkout history for you, which is a list of items you have borrowed from the Library. You were recently given a choice about your checkout history, and our records indicate that you recently chose to [INDIVIDUAL'S CHOICE PIPED IN HERE FROM BACK-END DATA SOURCE] Is this correct?
 - a. Yes
 - b. No
 - c. I don't recall making a choice about this
3. [FOR THOSE WHO WERE NEW TO LIBRARY] Since mid-2016, the University of Michigan Library has maintained checkout histories for some patrons, which is a list of items borrowed from the Library. You were recently given a choice about whether you want a personal checkout history, and our records indicate that you chose to [INDIVIDUAL'S CHOICE PIPED IN HERE FROM BACK-END DATA SOURCE] Is this correct?
 - a. Yes
 - b. No
 - c. I don't recall making a choice about this

[NOTE: ONLY THOSE WHOSE CHOICE WAS REMEMBERED AND CONSISTENT WITH OUR RECORDS CONTINUED WITH REMAINDER OF SURVEY]

4. What reason(s) led you to make that choice? _____
5. [ASKED IN DATA COLLECTION WAVES 3–5 ONLY] Did you give serious consideration to the opposite choice?
 - a. Yes
 - b. No
6. [IF YES TO PREVIOUS QUESTION] What were those other considerations? _____
7. [ASKED OF THOSE WITH EXISTING CHECKOUT HISTORIES] Prior to the recent opportunity you had to keep or delete your checkout history, were you aware that you could access your checkout history in the library's My Account tool?
 - a. Yes
 - b. No
 - c. Unsure
8. [IF YES TO PREVIOUS QUESTION] How often did you look at your Checkout History?
 - a. Never
 - b. No more than a few times a year
 - c. A few times a semester
 - d. About once a month
 - e. A few times a month
 - f. About once a week or more
9. [TEXT-ONLY ITEM] We're interested in your level of trust in the library, the University, and the internet more broadly with regard to responsible data management. Some types of data mismanagement include data leaks, data selling, and the sharing of identifiable information.
10. How concerned are you about data mismanagement, with regard to patron data stored by the University of Michigan Library?
 - a. Not at all concerned
 - b. Mildly concerned
 - c. Moderately concerned
 - d. Very concerned
11. How concerned are you about data mismanagement, with regard to student/employee data stored by the University of Michigan?
 - a. Not at all concerned
 - b. Mildly concerned
 - c. Moderately concerned
 - d. Very concerned
12. How concerned are you about data mismanagement, with regard to data stored about you by internet-based companies (e.g., Amazon, Apple, Netflix, etc.)?
 - a. Not at all concerned
 - b. Mildly concerned
 - c. Moderately concerned
 - d. Very concerned
13. We want to understand how groups that have been underrepresented or marginalized in academia think about data and privacy issues. Our goal is to serve everyone in our campus community sensitively and responsively. This question, like others

in this survey, is completely optional. Are you a member of one or more groups that have been underrepresented or marginalized on college and university campuses? (Examples include being underrepresented or marginalized on the basis of race, gender identity, disability status, sexual orientation, socioeconomic status, etc.)

- a. Yes
 - b. No
 - c. Prefer not to say
14. If you would like to share more information about your background or identity, please feel free to use your own words here: _____
15. If you have other comments you want to share about collection history and/or library data, please use the space below. (This is the final question on the survey.) _____

Appendix B: Coding Scheme

Chose to not have a checkout history. Coding categories for responses that were provided by (a) people in survey variants 1–5 who did not want a checkout history, and (b) people in survey variants 3–5 who thought carefully about both options.

- **Don't want or need it:** Statements about the checkout history not being wanted, needed, or useful
- **Library shouldn't keep it:** Statements about the library overstepping by retaining checkout histories
- **Privacy concerns:** Statements of worry about breaches to one's checkout history
- **Not aware of checkout history:** Statements that mentioned not having know about the history until being offered the choice
- **Not active in academic work:** Statements that mentioned no longer needing or using books from the library
- **Lack of trust in the university in particular:** Statements expressing a lack of trust that the larger university manages data safely
- **Uncodable:** Responses that did not make sense or were not categorizable

Chose to have a checkout history. Coding categories for responses that were provided by (a) people in survey variants 1–5 who wanted a checkout history, and (b) people in survey variants 3–5 who thought carefully about both options.

- **Want to have it:** A basic statement of want, with no explanation given
- **Want a reading list:** Statements about wanting to remember what one was reading in the past, preventing mistaken re-reading, facilitating intentional re-reading, tracking reading interests over time
- **Want to enable future research:** Statements about needing a citation list for dissertation or other research output, wanting to share things with other scholars
- **Other future use:** Statements about other future uses that is described in some detail (beyond a “just because” response)
- **No big privacy concerns:** Statements about a lack of privacy concerns, and about not having sensitive data in checkout histories
- **Expect this type of library service:** Statements about checkout histories being an expected or normal library service for patrons
- **Accessibility:** Statements linking accessibility support with the use of a checkout history
- **Uncodable:** Responses that did not make sense or were not categorizable

Other unique themes mentioned regarding the consideration of the opposite choice

- Privacy concerns related to specific materials being checked out
- Privacy concerns that were outweighed by expected benefits checkout history
- Privacy concerns that were outweighed by trust in the library
- Uncertainty about how to use the checkout history feature
- Concern about government access to records