Data for advocacy: A survey of data practices in African library systems

April 2021

Image credit: Miriam W. Mureithi, Principal Librarian, Kenya National Library Services Thika Branch, AfLIA Leadership Academy Cohort 1

African Library & Information Associations & Institutions

W
TECHNOLOGY & SOCIAL CHANGE GROUP
UNIVERSITY *of* WASHINGTON
Information School

TASCHA DESCRIPTION

The Technology & Social Change Group (TASCHA) at the University of Washington Information School explores the role of digital technologies in building more open, inclusive, and equitable societies. TASCHA researchers employ participatory approaches in partnership with organizations worldwide to co-design research projects and produce insights and practical resources that benefit communities.

AFLIA DESCRIPTION

AfLIA (African Library and Information Associations and Institutions (English) | Association Africaine des Bibliothèques et des Institutions d'Information (French) | Associação Africana de Bibliotecas e Instituições de Informação (Portuguese)) is an independent international not-for-profit organization which pursues the interests of library and information associations, library and information services, librarians and information workers and the communities they serve in Africa.

ABOUT THE AUTHORS

Renee Lynch is a Research Associate at the Technology & Social Change Group (TASCHA) and a Ph.D. Candidate in English at the University of Washington. Chris Jowaisas is a Senior Research Scientist at TASCHA, University of Washington Information School. Stanley Boakye-Achampong is the Research and Communications Coordinator at African Library & Information Associations & Institutions (AfLIA). Jason C. Young, Ph. D., is a Senior Research Scientist at TASCHA, University of Washington Information School. Joel Sam, Ph.D., is the Director of Research at AfLIA. Bree Norlander is a Research Scientist at TASCHA, University of Washington Information School.

ACKNOWLEDGEMENTS

COPYRIGHT

CONTACT

Technology & Social Change Group
University of Washington Information School
Box 354985
Seattle, WA 98195

Telephone: +1.206.616.9101
Email: tascha@uw.edu
Web: tascha.uw.edu

African Library & Information Associations & Institutions (AfLIA)
INSTI Building CSIR HQ
P. O. Box BC 38
Accra, Ghana

Telephone: +233 544 252 212
Email: info@aflia.net
Web: web.aflia.net

RECOMMENDED CITATION

ABSTRACT

African libraries are drivers of development in their countries. However, many of these libraries struggle to use data for self-advocacy in order to gain wider support from government bodies, development organizations, and other stakeholders. The University of Washington Technology & Social Change Group (TASCHA) and African Library & Information Associations & Institutions (AfLIA) conducted a survey of African library systems to examine their current data collection practices, data use cases, and data needs. This report summarizes findings from the survey including what data is collected by library organizations; how the data is used; how data is shared; how data is involved in decision making; and the experience of data collection. 29 responses from 19 African countries were analyzed to highlight major themes and suggest possible next steps for TASCHA and AfLIA to continue to support the data efforts of African libraries. The findings indicate that library systems value data and collect it purposefully for a variety of reasons, and most respondents collect operational and user data. Responses suggest that more training and support is needed in data analysis and dissemination as well as general ICT use. These interventions can foster library system agency with and through data.

140 CHARACTER SUMMARY

What are the current data practices and experiences of African libraries? What is needed to support their use of data for self-advocacy?

KEYWORDS

Africa, public libraries, data, data collection, international development, partnership

Table of Contents

1 Introduction: Data for Advocacy

African libraries are drivers of development in their countries. From Namibia (Isaaks, 2020) to Nigeria (Okojie and Okiy, 2019) to Kenya (Mureithi, 2020) and across the continent, libraries are creatively meeting the needs of their communities to positively impact those they serve. However, many of these libraries are in need of data in order to prove or demonstrate their value in development. Data is key for self-advocacy to gain wider support from government bodies, development organizations, and other stakeholders. For example, international development organizations desire output data which demonstrates that a potential partner is capable of contributing to a project (Lynch et al., 2020a). However, many libraries in African contexts experience challenges related to collecting and using data including data integrity in terms of completeness, accuracy, credibility, and relevancy; infrastructure; capacity; local investment in libraries; time; and participation of data collectors and respondents (Lynch et al., 2020b). These data challenges mean that libraries are often not perceived as drivers of development (Fellows et al., 2012), and their real value is often unrecognized and undersupported.

As a result, The University of Washington Technology & Social Change Group (TASCHA) and African Library & Information Associations & Institutions (AfLIA) conducted a survey of African library systems to learn more about their current data collection practices, data use cases, and data needs. This survey is part of an ongoing project called Advancing Library Visibility in Africa (ALVA), a joint research effort of TASCHA and AfLIA that aims to empower African libraries in development. The goal of the survey was to get a broad sense of current data practices and experiences as well as what is needed to support library use of data for self-advocacy. Additionally, we wanted to compare information about current data practices in African library systems with the data desires of international development organizations. We previously conducted interviews with development practitioners about how they choose partners and what is needed to shift perceptions of African libraries in this role (Lynch et al., 2020a). By knowing more about which types of data African libraries have access to and their current capacities for data use, we hope to connect African libraries and development organizations through potential partnerships made visible through data-based advocacy.

2 Survey Procedures

The survey was administered electronically using Qualtrics software. It had five sections asking about what data is collected by the organization; how the data is used by the organization; how data is shared; how data is involved in decision making; and the experience of data collection. The topics for these sections correspond to a life cycle of data from its collection, use and analysis, and adaptation based

upon experience in collection. The aim was to gain a broad understanding of the data life cycle within these organizations.

Participants were solicited using purposive sampling based on the professional networks and domain expertise of AfLIA. The original plan for conducting these surveys also involved an in-person follow-up session and was adapted to reflect the changed environment due to COVID-19. The 30 countries where AfLIA had established a contact with an in-country representative as part of the other ALVA activities were prioritized by the research team. This list of countries was then analyzed to attempt to gather responses from the various regions and methods of delivery of public library services, including centralized and decentralized public library systems. The research team then contacted representatives at national and/or federal/provincial/state libraries to confirm participation by those organizations and determine who would complete the survey on behalf of their library system. This means that each response to our survey could be taken to reflect the data practices of a network of libraries that the national or provincial/federal/state libraries directly managed. Participants were invited based on their roles within African library systems and/or involvement in the collection, use, analysis, or sharing of data related to public library services. 77 people were invited to participate in the survey, and we received 67 responses. These individuals, however, were permitted to delegate responses to the survey to other members of their staff, or to multiple staff members. This ensured that the individuals with appropriate institutional knowledge were responding to survey questions. Each complete survey was reviewed by the research team, and researchers followed up with respondents when responses to individual questions were unclear. Among these, 29 responses were complete. We classify a survey as "complete" if the respondent completed all required questions and then clicked the 'submit' button at the end of the survey. Note that even 'complete' surveys may not have responses to all questions, since not all questions were marked as required within the survey. 38 responses were incomplete, meaning that some questions had responses but the survey was not submitted by the respondent. For this report, we considered complete responses only, though this is a limitation that is discussed in "Limitations" below. Note that many of the incomplete responses were actually from participants that later submitted a complete survey.

Responses came from 19 African countries representing North, East, West, Central and Southern Africa.[1] Respondents chose to take the survey in Arabic, English, French, or Portuguese. For analysis, open-ended responses in Arabic, French, or Portuguese were translated to English using Google Translate.

---

[1] Classification was based on regional groupings as defined by the United Nations.

Overall, the goal of this survey was to gain greater insight into the data practices of African libraries. We hope that a better understanding of current data practices will allow library systems to collectively reflect on and improve those practices, so that data can be used even more powerfully to demonstrate the impact of libraries and to advocate for their needs. As the analysis will show:

- Library systems value data and collect it purposefully, for a variety of reasons
- Most respondents collect operational and user data
- Data collection and use is consistently tied to meeting the needs of library users
- More training and support is needed in the analysis and dissemination of data as well as general ICT use

2.1 Limitations

It is important to note some of the limitations to the responses we collected from this survey. Not all respondents answered all questions, even within surveys that were submitted and which we considered "complete." In this analysis, we have used "respondent" to refer to those who responded to a given survey question and "participant" to refer to anyone who submitted the survey. For participants, n=29. This distinction makes it clear when we are talking about all participants or about only a subset of them that answered a particular question. Additionally, some responses within one submission may have contradicted other responses within the same submission; for example, a respondent may have reported that their library system does not collect data on library users for one question but then selected types of user data that they collect in the follow up question. In these cases, we identified contradictory responses, followed-up with the respondent to obtain clarification, and changed the data set accordingly.

As noted above, we only analyzed responses that had answers to all of the required questions and were submitted by the respondent. This means that we only considered 29 responses out of a total of 67, and this limits the scope of our findings. Possible reasons for incomplete responses include inconsistent internet connectivity for respondents, confusion about the Qualtrics interface, unclear survey procedures, lack of time, the length of the survey, and general library system disruptions due to the global COVID-19 pandemic.

The results here are dominated by the public library perspective since this was our target participant group. Responses were primarily from National Libraries providing public library services (45%) and Public Library Services (34%). A smaller proportion came from State Libraries providing public library services (17%) and other library organizations (e.g. library association or community library; 3%). Therefore, while these results offer unique insight into the relationship between libraries and government, more specific outreach may be needed to apply these insights to community and other libraries.

Additionally, it should be noted that eight responses (28% of the total responses) came from one country with a decentralized library system, and this is a considerable proportion which may influence overall trends. Though the responses came from 19 different countries across the African continent, it is unlikely that they are representative of library systems across the entire continent. For example, East African libraries are overrepresented while Central and North African libraries are underrepresented. Therefore, any claims made in this analysis need to take this into account.

Originally, we intended to include an analysis of regional trends, for example, to compare data practices between East and West Africa, or for different libraries within the same national system. However, the data we collected with this survey does not support rigorous statistical generalizations in a way that would allow for responsible analysis. Further data collection would be necessary in order to compare data across regions, countries, or other types of trends.

3 Findings and Analysis

This analysis is focused on a few broad categories related to data culture in African libraries: data collection, use, and sharing practices. These categories were determined by the research team based on current knowledge and gaps pertaining to library system data practices, as an extension of prior research. We also consider connections between data practices and sustainable development work, based on our prior interviews with development organizations (see Lynch et al., 2020a).

The analysis here is a combination of quantitative and qualitative. Quantitative analysis consisted of creating descriptive statistics for closed-ended questions. We created charts of responses using Tableau software which aided in visualizing the responses for analysis. Qualitative analysis consisted of describing common themes and notable contrasts in the content of open-ended questions. For this process, all responses were translated, read, and then analysed for trends, using color coding to note themes which are mentioned in our analysis, with the number of responses within that theme noted where possible. Examples are provided to illustrate typical and atypical responses. Where possible, the analysis mentions the specific question in the survey that the analysis is drawn from as (Qsection number.question number), e.g. (Q2.20). Please see the appendix for the survey instrument.

3.1 Data Collection

3.1.1 How is data collected and verified?

Who Collects the Data
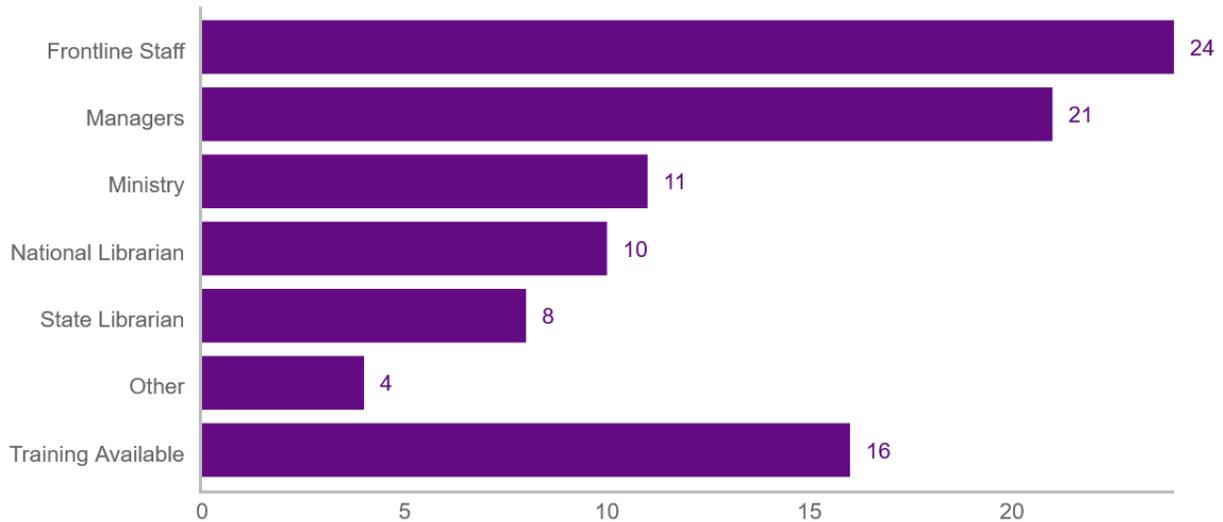
Figure 1. *Displays the sum of how many library systems involve each type of staff role in the collection of data. Participants could select all answers that apply to their library system. All participants (n=29) responded to this question.*

Frontline staff and managers are involved in data collection across the most number of library systems (Figure 1). This may suggest that the bulk of library data collection happens in the library itself and is carried out by local staff. Approximately 40% of the participants report that the ministry responsible for libraries collects data, as well as national and state librarians though to a lesser extent. However, only 55% of the participants report that training is provided to data collectors/reporters prior to collection/reporting. This echoes the need for staff training related to data which was frequently mentioned in open-ended responses.

All 29 participants provided open-ended responses to Q2.20, which asked respondents to describe the process for the entire data cycle from collection to analysis, but not to the same degree of detail. Some respondents provided many details while others provided minimal descriptions, so that must be taken into account and balanced with other responses in this survey. Their descriptions of the data cycle reveal a fair amount of variety among library systems. For example, in terms of frequency of data collection, some mention daily (6 respondents), weekly (2), monthly (5), or quarterly (4) schedule while two respondents mention annual data collection only in their response to this question.

Overall, three organizational patterns seem to be present in data cycles: hierarchical, issue-based, and independent. Most participants (17, or 59%) describe a *hierarchical* system whereby data is collected by

library branches on a regular schedule and passed up to a higher authority who compiles, analyzes, and submits the data to higher levels, making this the most common pattern for data cycles. There were two exceptions to this pattern. First, some organizations adopted an *issue-based* approach to data whereby involved persons "identify issues and have conversations around the issues" and then decide the collection method, indicators, and necessary training (2 respondents, or 7%). Others described an *independent* approach whereby libraries collect their own data and may or may not share it for compilation or analysis at higher levels (4 respondents, or 14%). This happens when the national library does not have a mandate to collect data from branches (1 respondent), or the national library body is a "semi-autonomous organization" that compiles data and publishes reports "by [them]selves" to share with the government as a courtesy (1 respondent). It seems that the independent approach may incur challenges for libraries; as one respondent notes, "There is an absence of a central agency/ministry to collect statistical data on library services at national level. If statistics are required, librarians rely on a very informal way on their network of contacts." For six responses, the nature of the data cycle is not clear, with references to data that is "passively compiled" (1 respondent), simply not collected (1 respondent), or unknown due to lack of details provided.

Other commonalities include issuing an annual report based on compiled data (7 respondents) and utilizing a mixture of paper-based and electronic methods. In addition, three respondents mention the use of standardized forms for data collection, and seven mention data verification of some type in this open-ended response, though more respondents report data verification processes in the close-ended question on this topic.

Do Libraries Use Verification Processes
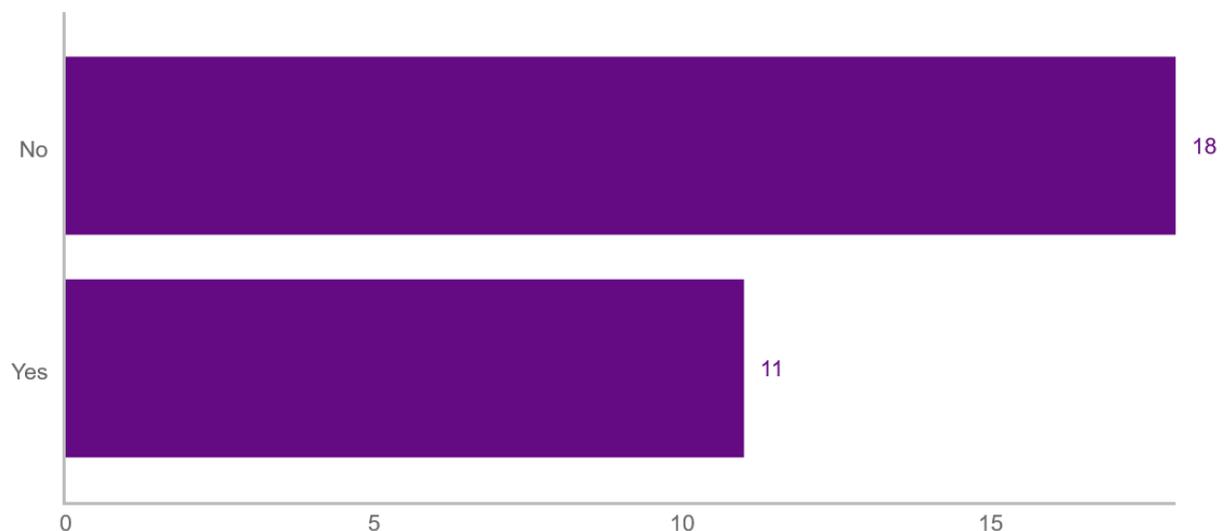
Figure 2. *Displays the sum of how many library systems do or do not have a verification process that they use to ensure accuracy of collected data. All participants (n=29) responded to this question.*

In Q2.26, responses indicate that most of the data collected by libraries does not undergo a verification process (Figure 2). Only 38% of respondents report that there is a verification process for accuracy of

collected data, while 62% say there is no such process. When asked to describe the process of data verification (Q2.27), a common theme was passing the data up the levels of hierarchy, for example, by having data checked by a manager before sending to the library headquarters or ministry. Descriptions of the exact activities of verification were limited; respondents mentioned "proofreading" or "cross-checking" data between sources, such as the physical record and its digital entry, or utilizing trend analysis to target anomalies with previous data collected on similar topics. In two instances, responses mentioned the involvement of a third party, such a national office of statistics, for external validation. For most, however, the process of verification was expressed as gaining the approval of a higher authority. This echoes the predominance of hierarchical models in the data cycle, as described in the open-ended response.

3.2 Types of data

3.2.1 What types of data are most commonly collected?

Figure 3. Describes the general types of data collected by library systems. Participants could select all answers that apply to their library system, and options included operational data, user data, demographic data, Sustainable Development Goal (SDG) and National Development Plan (NDP) data, data for NGO-led or private projects, and other. To ensure confidentiality, library names were replaced with a pseudonym that consists of the type of library (e.g., 'national library') and a unique number.

Purple represents data types collected by the library system, and white represents data types not collected by the system.

Data Types Collected - General Types
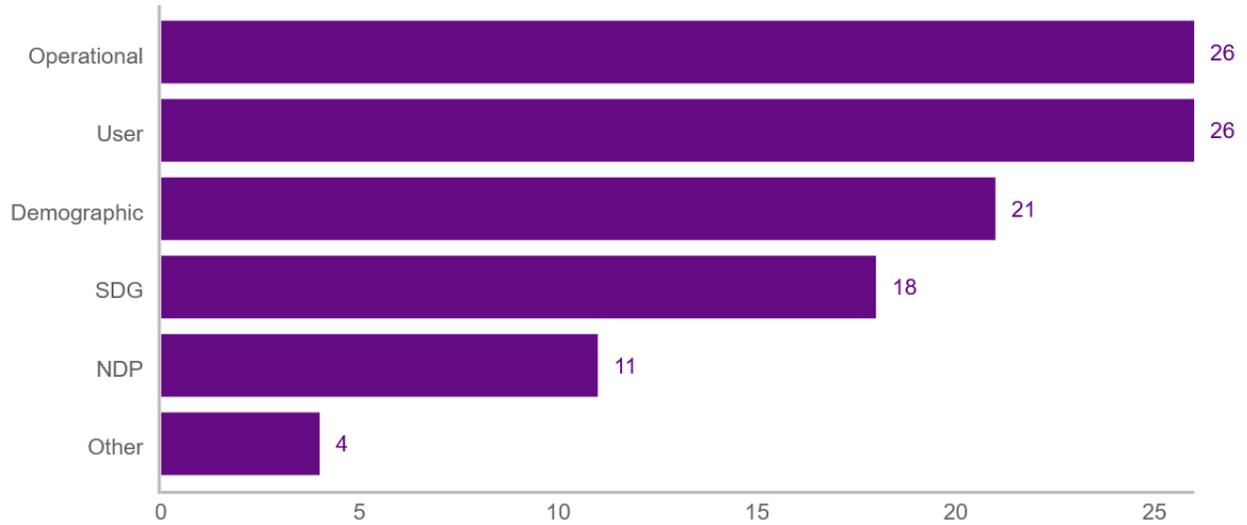
Figure 4. Displays the sum of how many library systems collect each general type of data. Participants could select all answers that apply to their library system, and options included operational data, user data, demographic data, Sustainable Development Goal (SDG) and National Development Plan (NDP) data, data for NGO-led or private projects, and other. All participants (n=29) responded to this question.

Data Types Collected - Specific Types

Data Type	Count
Number of Patrons	20
Print Materials for Circulation	16
Public Internet	16
Number of Staff Computers	16
Public Service Hours	16
Seating Capacity	14
Types of Info Sought by Users	13
Tech Training	13
Staff Training	13
Staff Demographics	13
Number of Other Staff	13
Capital Expenditures	13
Number of Books	13
Visitor Demographics	12
Children Programs	12
Teen Programs	12
Public WiFi	12
Number of Public Computers	12
Number of Librarians	12
Other Materials	12
Number of Holdings	12
Skills Training	11
Internet Type	11
Other Digital Equipment	11
Staff Skills	11
External Use of Library	10
WiFi	10
Total Paid Staff	10
Program Attendance	9
Youth programs	9
Other Infrastr. for Patrons	9
Number of Volunteer Staff	9
Revenue by Regional Source	9
Outreach Measures	8
Electricity	8
Operating Expend - Other	8
Number of Electr. Materials	8
Electr. Materials Expend.	7
Print Materials Expend.	7
Operating Expend. - Salaries	7
No. of Users for Programs	6
No. of Users for Training	6
Generator	6
Automation/ILS	6
Website Presence	6
Building Size	6
No. of Volunteers	6
Operating Expend. - Benefits	6
Revenue by National Source	6
Revenue by Local Source	6
No. of Audiovisual Titles	6
Other Material Expenditures	5
No. of Mobile Libraries	4
Service Area Population	4
Revenue by Other Source	3
Other Staff	2
Other Library Use	1
Other Financial	1
Other Services	0
Other Infrastructure	0

Figure 5. *Displays the sum of how many library systems collect each specific type of data. Participants could select all answers that apply to their library system. Participants only provided information about specific data types if they confirmed that they did collect the general data type into which each specific type falls. For example, 'number of holdings' is a form of operational data. Therefore, only those participants that indicated that they collect operational data were asked whether they collect data on number of holdings.*

Responses suggest that most libraries collect multiple types of data. Overall, the most commonly collected types are operational data and data on library users (Figure 3, Figure 4).

The survey defined operational data as data about "collections and materials, financial, infrastructure, human resources & administration, use of the library, and services & programs." 90% of respondents report collecting data of this type. The top five most commonly collected types of operational data include: number of patrons, print materials circulation, public internet, number of staff computers, and public service hours (Figure 5, as compiled from Q2.3#1, Q2.4#1, Q2.5#1, Q2.6#1, Q2.7#1, Q2.8#1).

Regarding user data, 90% of respondents report collecting some type of data about library users (Figure 4). 81% report collecting demographic information about user age, gender, occupation, and other aspects, with education level and purpose for attending the library given as examples.

This suggests that these data types (operational and user data) may be the most universally available across library systems. Thus, if a data platform is made or adjusted to make library data visible to potential partners, the platform could include space for those categories and the most commonly collected types of operational data mentioned above. Operational and user data were also desired by development practitioners in interviews, which furthers the case for making these types visible.

Another question asked whether libraries collect data specifically for privately or NGO-funded projects (Q2.16). Of the 12 participants that responded to this question, 11 indicated that they did collect this data. This represents 92% of those respondents, or 40% of the total participants. Examples of privately-funded or NGO data provided were generally program-focused, such as attendance at programs, gender of participants, feedback on/impact of specific programs, and general community-focused such as services available for the blind, data on political violence, and school locations and enrollment. While it is not clear if libraries were collecting this data independently before the privately or NGO-funded project, the large proportion of library systems collecting data to support NGO projects suggests that these types of data are related to NGO priorities. It would be interesting to do further research with a library that collects data for such purposes to learn more about their data-related challenges and successes in interacting with private or NGO funders.

3.2.2 How are these types of data chosen to be collected?

What Data Are Legally Required
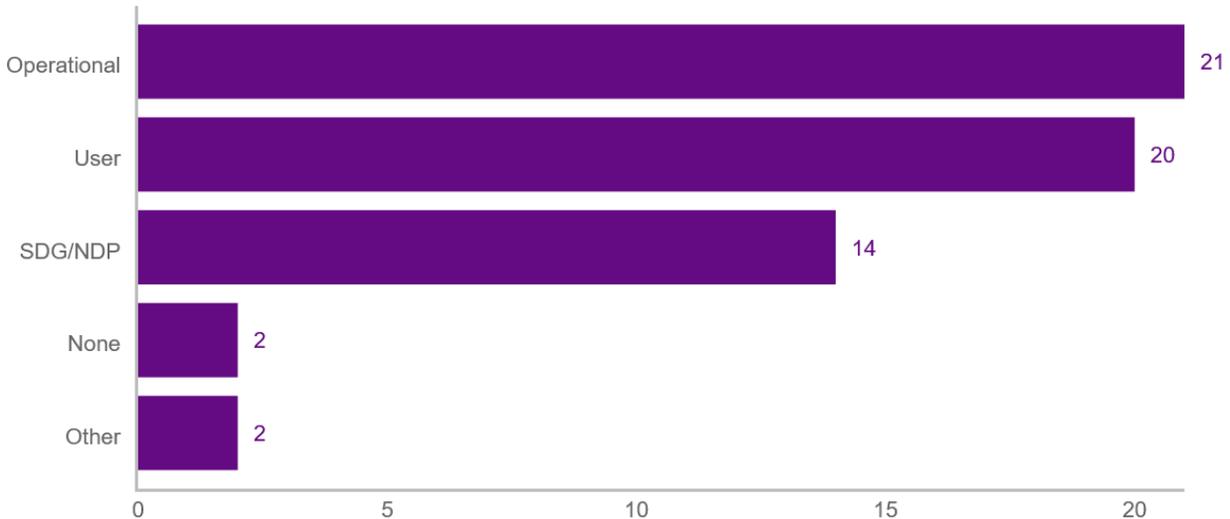
Figure 6. *Displays the sum of how many library systems are mandated or legally required, by their ministry or another government authority, to collect each type of general type of data. 25 participants responded to this question.*

The most commonly collected types of data, operation and user data, are also the most commonly mandated among respondents (Figure 6). This suggests that the main determiner of what data is collected is the mandate of a ministry or government authority. To a lesser extent, development goals drive which data is collected, and nearly half of the participants (48%) report that they are required to collect data related to SDGs and NDPs. Additionally, two participants noted that other types of data are required by their ministry or government authority, including "text" and "staff data." It is clear that most library systems in this survey have some sort of mandate about which data to collect, with only two participants (7%) reporting that they have no specific directives about which types to collect.

When asked to describe the process for choosing the data that is currently collected and why it was chosen (Q5.8), the most common theme is choosing data to understand library usage. 12 participants (41%) mentioned collecting data that is tied to library users, either to assess their needs or measure library success in meeting these needs and identifying "gaps." Another common theme is demonstrating accountability via data, for example, when advocating for government funding as well as meeting goals expressed in a strategic plan. As one respondent reports, "The data collected was chosen because it acts as evidence during budget allocations, Human resource acquisition." Another adds that data is chosen based on "the relevance of these data and the ease of making advocacy with government units and other partners." Only one respondent reported that "there is no process," which indicates that for the majority of library systems represented here, the types of data are chosen purposefully. Overall, these responses suggest that data to be collected is chosen strategically relative to library goals.

3.3 Uses of Data

## 3.3.1 How is data used by organizations - for example, to change internal policy or user-facing policy?

How Data Are Used
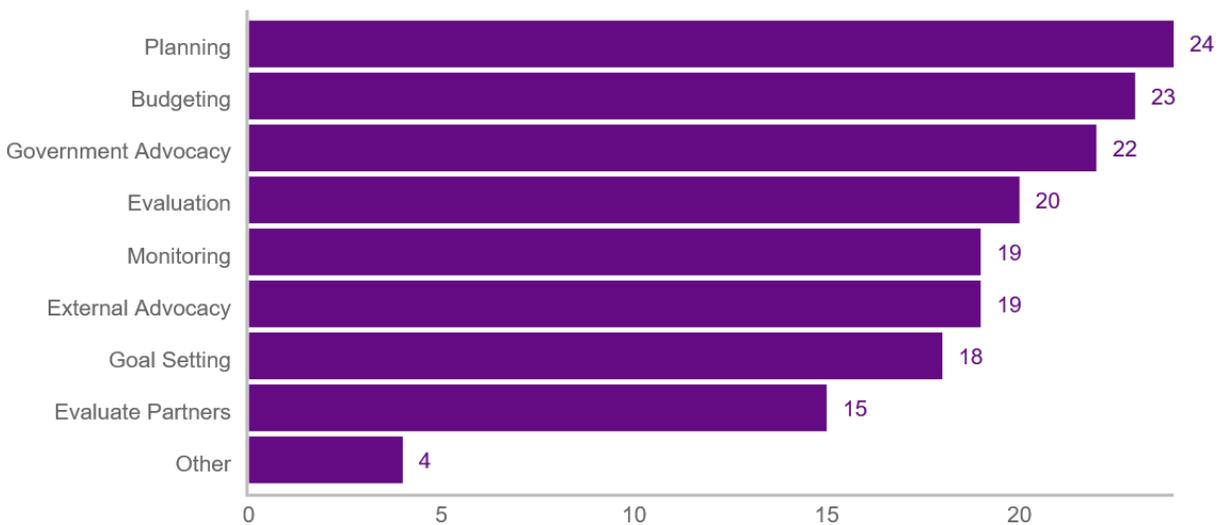
Figure 7. *Displays the sum of how many library systems use the data they collect in order to perform the following tasks: planning of programs and services, budgeting of programs and services, goal-setting, monitoring, assessment/evaluation, advocacy with other government entities, advocacy with external entities, evaluation of potential partnerships, and other. Participants could select all answers that apply to their library system. 26 participants responded to this question.*

Responses indicate that library data is used for a wide variety of purposes. Most participants (approximately 80%) report using data for operational purposes related to planning and budgeting (Figure 7). However, a substantial proportion (76%) also report using data for advocacy with other government entities. Approximately two-thirds of libraries use data for monitoring and evaluation as well as advocacy with external entities. 62% of participants say data is used for goal setting, and just over half report using data to evaluate potential partnerships. "Other" uses for collected data include policy-making, assessing the number of staff needed, and preserving memory in general. Taken together, the varied uses of data reported here support the finding that most libraries are collecting data strategically and making data-informed decisions.

Value of Data
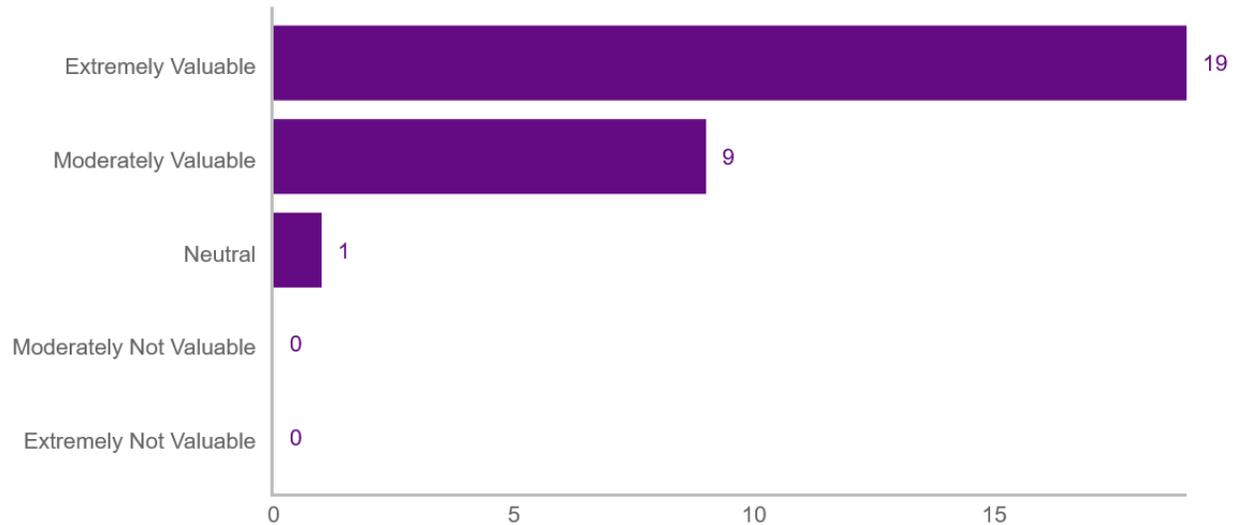
Figure 8. *Displays the sum of how many library systems rated the value of data within each category. Participants were asked to consider how valuable data are to them relative to the amount of time spent collecting that data. Participants could select only one category. All participants (n=29) responded to this question*

It is no surprise, then, that a vast majority of participants (97%) find some value in collecting data, relative to the amount of time spent collecting it (Figure 8). Most of these (66%) say that data is "extremely valuable" while only one respondent holds the view that the value is neutral, "neither valuable nor not valuable." Overwhelmingly, those who responded to this survey think that collecting data is a worthwhile activity. It is important to note that a majority of the participants to this survey were in leadership positions or other positions of authority when interpreting these results. On the one hand, this is an important finding, in that these library leaders are the ones most likely to use data to carry out activities such as advocacy, budgeting, staffing, re-stocking, managing library investments, and more. On the other hand, it would be interesting to see if this perception holds for other library staff such as frontline librarians and managers who may not have had the chance to respond to this survey. Given that these frontline staff are the ones most likely to collect data but perhaps less likely see the impact of the data they collect, it is important that they see the value in the activities that they are carrying out. Future research efforts would be valuable to explore whether there is a disconnect between those who see data as valuable and those who collect it. There is some evidence to support the presence of a disconnect in the "Challenges" section below, in which respondents note "lack of staff interest" among other personnel-related issues. This may indicate that it would be valuable to help frontline staff better understand and benefit from the value of data, so that they are more invested in collection practices.

Do Data Impact Internal Policy
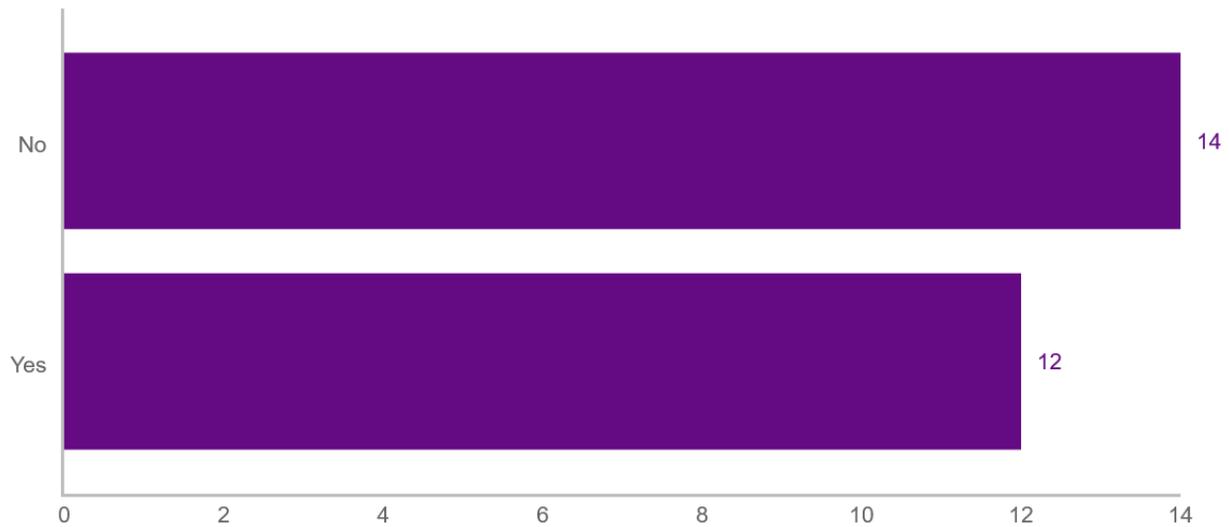
Figure 9. Displays the sum of library systems that have changed an internal policy in the last two years based on data they have collected. 26 participants responded to this question.

Do Data Impact External Policy
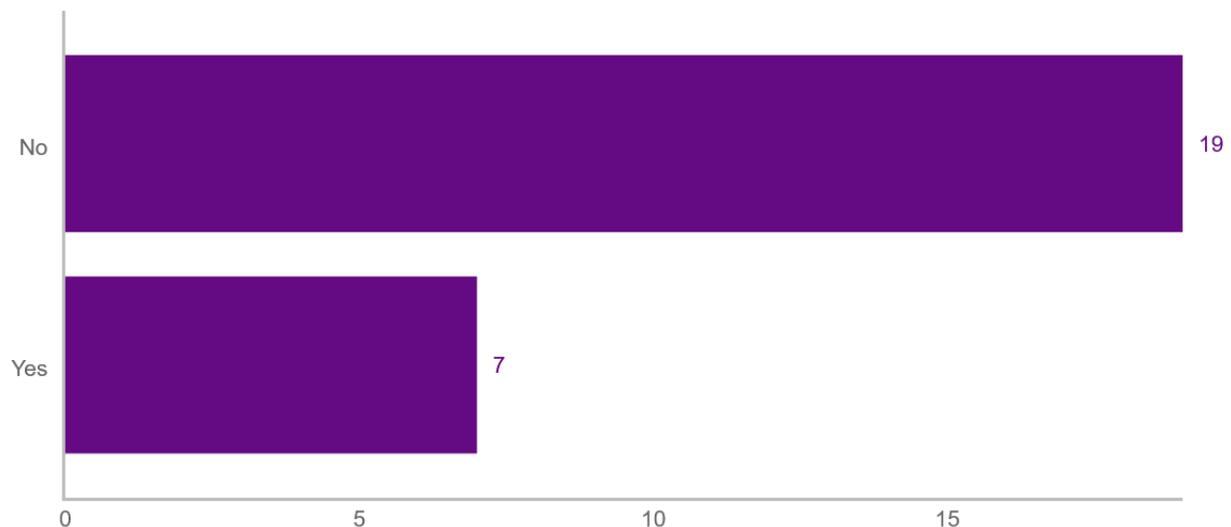
Figure 10. Displays the sum of library systems that have changed a user- or patron-facing policy in the last two years based on data they have collected. 26 participants responded to this question.

While most respondents consider data very important, a majority of libraries have not used the data to make internal or external policy changes. For those who have changed policy as a result of data, they seem to influence policy more internally rather than externally (Figure 9). 41% of participants report an internal policy change within the last two years due to data collection; most of these changes were

operational such as adjustments in library hours, staffing allocation, budgeting, and their own data collection practices. In an interesting example, one library system augmented a new fee system after discovering that visits to rural branches were down after the introduction of fees, and data encouraged them to lower fees for rural areas and admit children for free. In an interesting example, one library system discovered that visits to rural branches declined after augmenting a new fee system. This data encouraged them to lower fees for rural areas and admit children for free. For those that reported external (user- or patron-facing) policy changes within the past two years (24%) (Figure 10), a common theme was expanding services for users such as more programming for rural areas, increased hours, and renovated spaces. However, two respondents reported limiting library lending services as a result of data on loss of materials. Nonetheless, as seen with uses of data, libraries most commonly report using data for policy changes to meet user needs.

3.3.2 How is data used to make partnership decisions? What type have been persuasive?

Importance for Choosing Partners
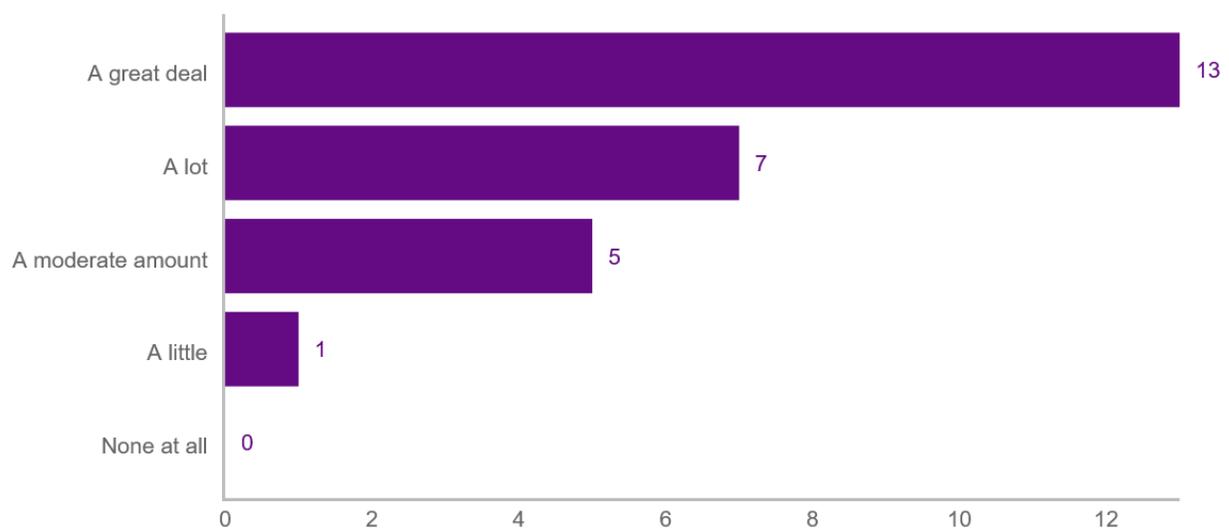
Figure 11. *Displays the sum of how many library systems rated the importance of data within each category. Participants were asked to consider how much of a factor data play in the decision to partner with another entity (i.e., governmental or NGO). 26 participants responded to this question.*

Importance for Soliciting Support
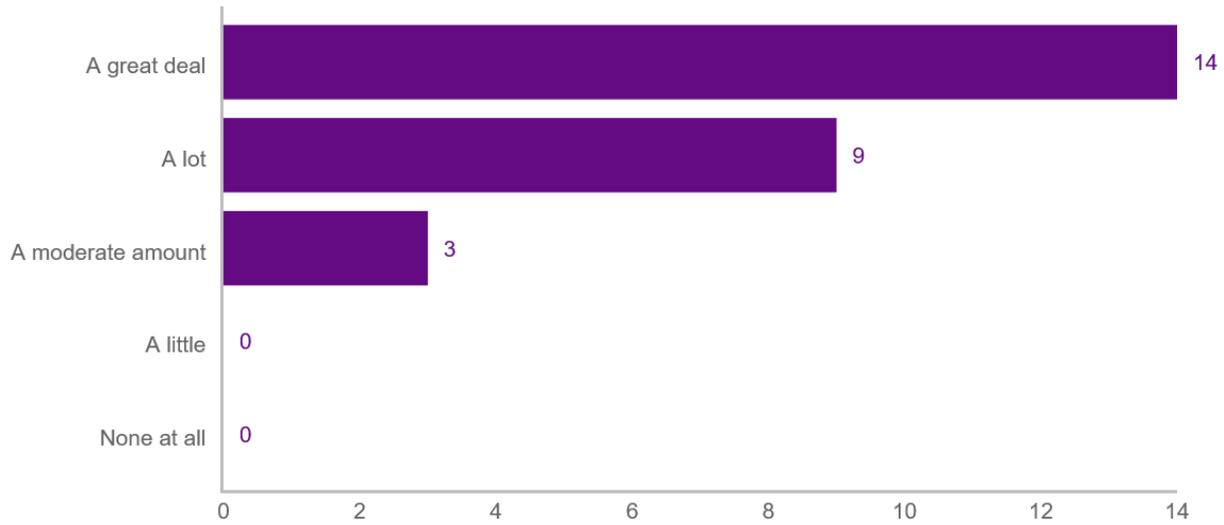
Figure 12. *Displays the sum of how many library systems rated the importance of data within each category. Participants were asked to consider how much of a role data plays when soliciting support from another entity (i.e., governmental or NGO). 26 participants responded to this question.*

Respondents see data as important in both deciding and soliciting partnerships with governmental, NGO, and other entities. Most participants see data as important in the decision to partner, with two-thirds reporting that data has "a great deal" or "a lot" of importance (figure 11). Similarly, over 50% of participants report using data to "evaluate" potential partnerships (Figure 7). A greater proportion (80%) of participants think that that data is important for soliciting external support (governmental or NGO) (Figure 12). As a concrete example, one participant shared that the evaluation data from a school library project led to increased funding for four more school libraries (response to Q3.4). This illustrates lived experience of the power of data for library systems. Taken together, these findings demonstrate that libraries have bought into the value of data for partnerships.

When asked what types of data have been the most persuasive for advocacy with other government entities (Q3.12), there is a notable emphasis on numerical or quantitative data. For example, one respondent lists "Reading room attendance data; Number of reading room subscribers, Number of works available to users; Number of subscribers per year; Number of loans" which illustrates the predominance of numbers as data. Overall, most of the respondents mentioned data points related to library usage or facilities, such as memberships and computers available. Three respondents mentioned data that illustrates library impact, such advocacy and innovative outreach efforts, as persuasive for government entities, which may be lower than expected.

For advocacy with external entities such as NGOs, private companies, and other donors (Q3.13), the types of data reported as persuasive are largely the same: operational, usage, and impact data. For external entities, there is an additional theme of data to highlight challenges in the community, such as

rates of illiteracy, disabilities, and lack of materials. In the words of one respondent, data on the limited resources available "have allowed donors to provide some support to meet some needs… recognizing that this support in [country] is very scarce." Two respondents also mentioned data that highlights the "expressed needs" of staff to build professionals skills. Though more context would be needed, this difference suggests that libraries tend to present their successes internally and their needs externally. Given connections to our development interviews below, libraries could be encouraged to shift their strategy to emphasize successes externally in order to be perceived as capable, valuable development partners

3.3.3 What skills are needed for staff to utilize data more effectively?

As discussed above, over half of the participants indicated that their library systems provided training on data collection (see Figure 1). Nevertheless, many participants discussed additional training needs related to utilizing data more fully. In open-ended responses asking what skills would be needed for library staff (Q3.14), respondents mention a wide range of skills that they perceive as needs, from "computer appreciation" and "[t]he fundamentals of data literacy" to "web creation and design" and "data protection." This suggests that needs may vary drastically depending on context. Most responses tend to pool around skills in analysis, as in "basic training in statistics," "data interpretation skill," or simply "data analysis skills." Many of these emphasize the use of technological tools for analysis such as "use of excel [sic] in data manipulation" or "spreadsheets statistics." In addition, eight respondents mention a need for skills in data collection, specifically in technical or online methods. Seven respondents mention a need for data presentation skills related to translating, reporting, and disseminating data. Overall, these responses highlight the intersection between general ICT skills and the ability to work with data, emphasizing that in most cases, there is a need for both.

More generally, two responses mention "teamwork" or "collaboration skills." Importantly, one respondent notes, "The training must include the construction of these by means of computers as well as harmonization with the models universally applied in this area without neglecting the need to include specific data in our country." This comment emphasizes that any training should be able to apply to types of data and data collection methods already in place, recognizing that libraries may not have much agency to implement drastic change.

3.4 Data Sharing Practices

3.4.1 To what extent is the data open? With whom are data shared?

Figure 13. Describes with whom library systems share their data. To ensure confidentiality, library names were replaced with a pseudonym that consists of the type of library (e.g., 'national library') and a unique number. Purple represents audiences with whom the library system shares data, and white represents audiences with whom they do not share data. 27 participants responded to this question.

Data Sharing Practices
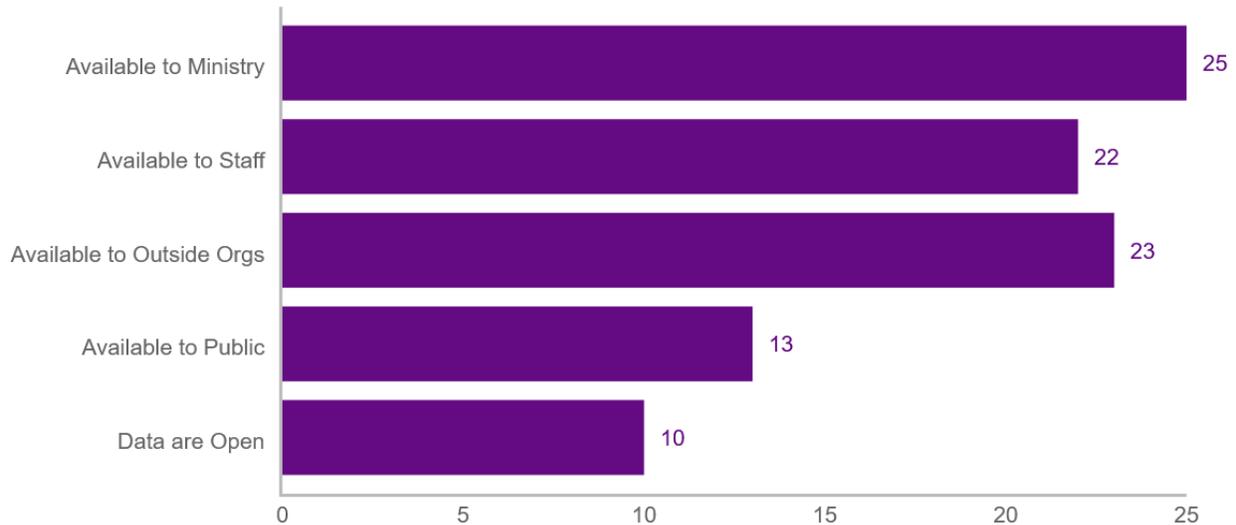
Figure 14. *Displays the sum of how many library systems share data with each category of audience. Participants could select all answers that apply to their library system. 27 participants responded to this question.*

Survey responses suggest that a majority of library systems do not provide data openly. Instead, data is more available to official entities such as ministries and library staff rather than the general public (Figure 13). Only 34% of participants indicate that data is uploaded to open portals or websites, and only 45% make data available to the public (Figure 14). Data is most commonly accessible to ministries (86%). Several library systems indicated that they are required to report data to their national statistics agency or ministry and described varying reporting requirements (Figure 15). Respondents most often reported user (60%) or operational (56%) data, followed by SDG or NDP data (44%). Libraries also shared data with "entities outside [their] organization" (79%) and library staff to a slightly lesser extent (76%). Interestingly, two participants indicated that data is provided openly but not to the public. This raises the question of a disconnect between our understanding of "open data" and participant understanding.

What Data Are Reported to the Ministry
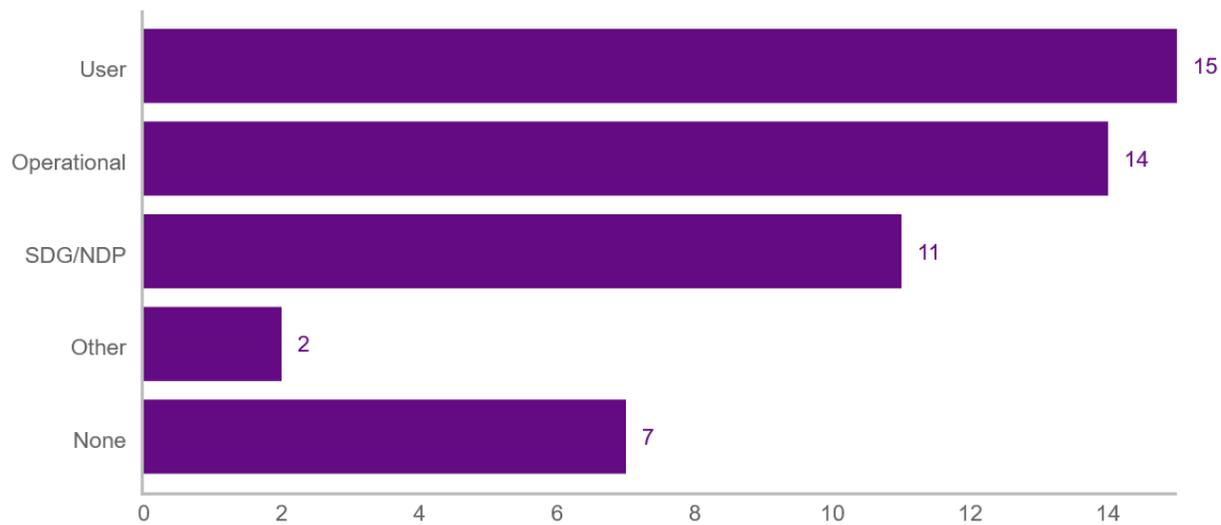
Figure 15. Displays the sum of library systems that report each type of data to the national statistics agency or ministry. Participants could select all answers that apply to their library system. 25 participants responded to this question.

3.4.2 How is it shared?

48% of respondents say that data collected from the libraries is available to the public through electronic/digital access, for example, a website or Facebook page (Figure 16). Of the links provided by respondents (Q4.3), seven are library websites, four are Facebook pages, two are ministry websites, and one is a Google Drive. 37% say that data is uploaded to any open data portals or websites, and examples include three publicly accessible annual reports, two NGO-hosted sites, two government-hosted sites, and one Facebook group. This reflects a variety of ways of sharing data.

Asking about the relative risks and benefits their organization would receive from sharing data it collects more widely (Q4.9, 4.10) revealed a rich discussion around attitudes towards open data which could be explored further. Perceived benefits include those for libraries, with "increased visibility" and partnerships as main themes, as well as benefits for users such as increased awareness of services and increased advocacy to expand services for users. Among reported risks, a strong theme is fear of data misuse, if they should fall into "the wrong hands and use for the wrong purpose" such as hacking, misinterpretation, or discrediting libraries. In one example, a respondent states, "Resources are very scarce. If we are able to attract help from an organization that is politically affiliated to an opposition party, the ruling party may blacklist us." Another fears "interpretation of the data out of context to harm the [library] and create confusion in the minds of the public in general, and more particularly our [stakeholders.]" These and other comments highlight the real consequences and complexities of what constitutes "fair use" of data for these library systems. Additionally, some are concerned about data privacy and crediting the data source properly. Four respondents perceive no risks associated with sharing data more widely. Overall, this could be an area for deeper research and analysis.

Data Made Available to Public
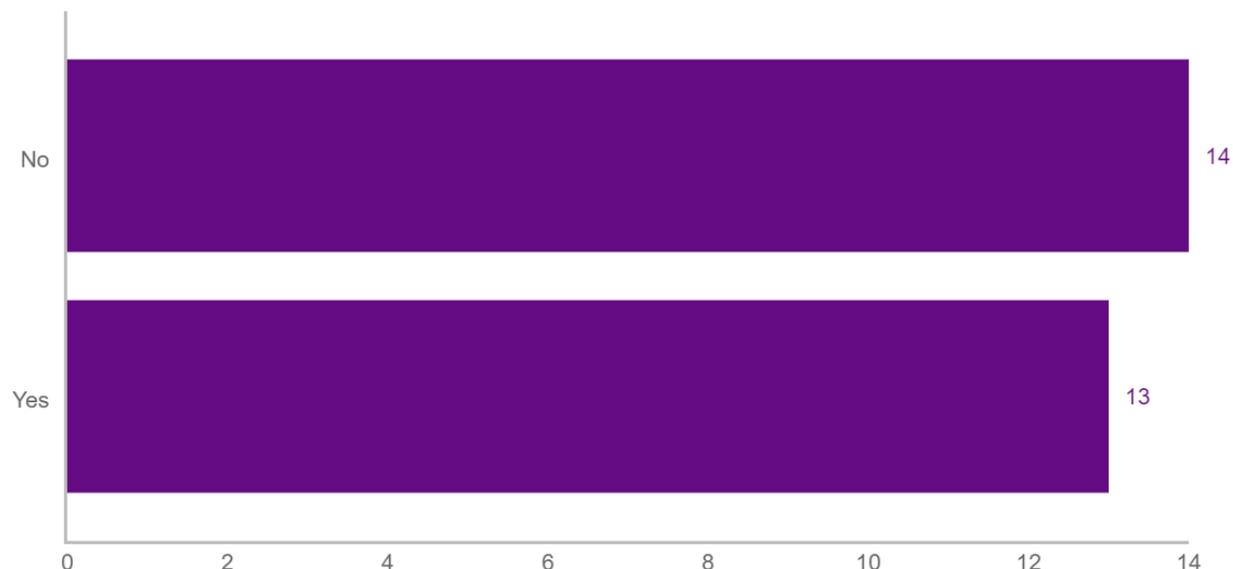
Figure 16. *Displays the number of library systems that make their data available to the public through electronic or digital access. 27 participants responded to this question.*

Combined with the abovementioned desire for staff training around data presentation, reporting, and dissemination, these responses represent an opportunity to empower library systems to have agency over their own data sharing. Some express an explicit desire for capacity building in this area. For example, one respondent said, "My organization needs to have a way of disseminating the data to the public and other organizations. Data should be readily available if needed by external entities." If libraries feel confident using and sharing data responsibly on their own terms, that could be a possible intervention in enabling more equitable data practices.

3.5 Challenges

3.5.1 What are common challenges related to data?

Are Data Collected in a Complete and Timely Manner
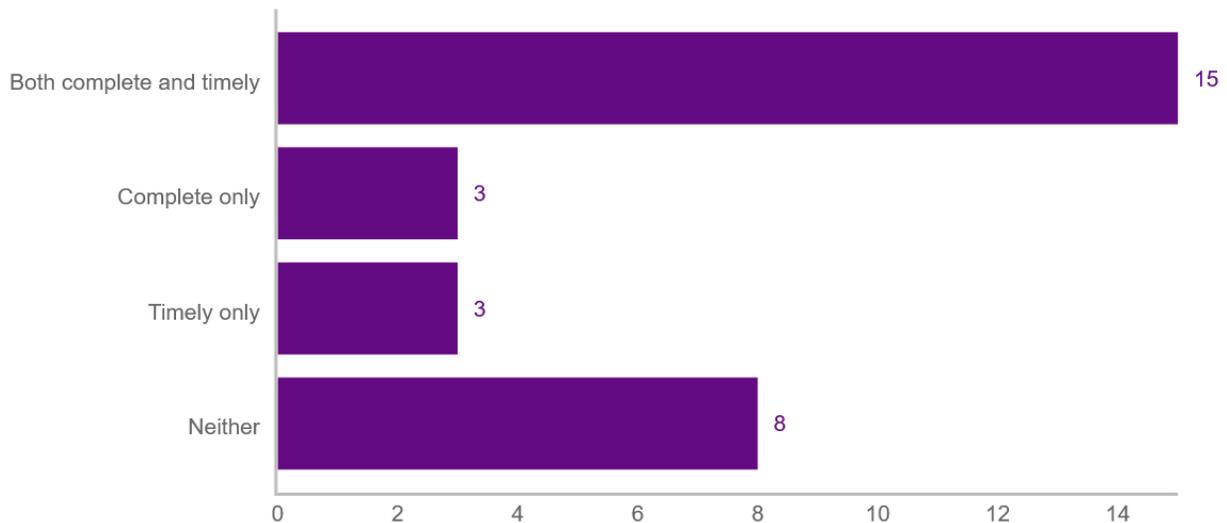
Figure 17. *Displays the sum of library systems that described data collection within each category. Participants were asked to reflect on whether libraries regularly return data that are complete and timely. All participants (n=29) responded to this question.*

In the words of one respondent, "The major challenges are numerous." These challenges can be characterized in two main categories: logistical/infrastructural and staff-related. 12 respondents mentioned logistical issues. A commonly cited issue is a lack of internet and/or electricity, especially in rural areas, which makes data collection and reporting by digital means difficult. Also common was inadequate funding or budgeting as well as the lack of tools, such as computers or "proper internet structure [or] Integrated library system," to enable data collection and sharing.

Eleven respondents mention challenges related to library staff, including insufficient numbers of staff, limited technical and data management skills, or delayed responses. As indicated in Figure 17, only half of respondents say that libraries regularly return data that is complete and timely, with 28% reporting that data received from libraries is neither complete nor timely. Lack of data and technical skills among staff is a notable theme, though some highlight lack of data analysis skills as an issue larger than just staff, noting "lack of understanding of the importance of data collection" and "misinterpretation of data by stakeholders" as additional issues. The challenges of the staff may be related to a lack of policy or directive around data, which may in turn impact their motivations; as one respondent explains, "The other challenge is the measurement of the time spent in data collection and the benefits that are expressed through the outcomes of the information outputs. It is not easy to determine the cost-benefit of the data collected and processed." This combined with "[n]o legal or statutory obligation to collect data at national level," "government protocols," "bureaucracy," and "changes to definitions and policies

and maintaining data comparability" may contribute to the reported "lack of staff interest." It's likely that these socially-based challenges may be further compounded by the logistical issues such as inconsistent electricity and internet access. It would be interesting to consult with more frontline staff in order to gain insight into the confluence of these challenges.

3.5.2 What support is needed?

The reported challenges reinforce the needs for additional capacity building around the collection, analysis, and use of data. Given the logistical challenges, increased funding for staff and infrastructure such as the internet and computers could also assist with this capacity building. In addition, it seems there may be a disconnect between higher authorities and frontline staff with regards to the value of data given the labor involved. It's possible that increasing staff ownership over data, perhaps through offering opportunities for local analysis and actions as a result, may increase frontline buy-in on data collection. However, this issue is complex and should also be considered in light of hyperlocal power dynamics.

3.6 Data and SDGs and NDPs

3.6.1 To what extent are the types of data collected aligned with SDGs and NDPs?

A majority of participants (62%) report collecting some data to support development goals such as Sustainable Development Goals (SDGs) and/or a National Development Plan (NDP), as shown in Figure 18. Almost half of participants (48%) collect data to support both SDGs and NDPs, and this is more common than collecting data to support SDGs (7%) or NDPs (7%) only. A notable portion (38%) do not collect any data aligned with development goals.

Collect Data to Support NDPs or SDGs
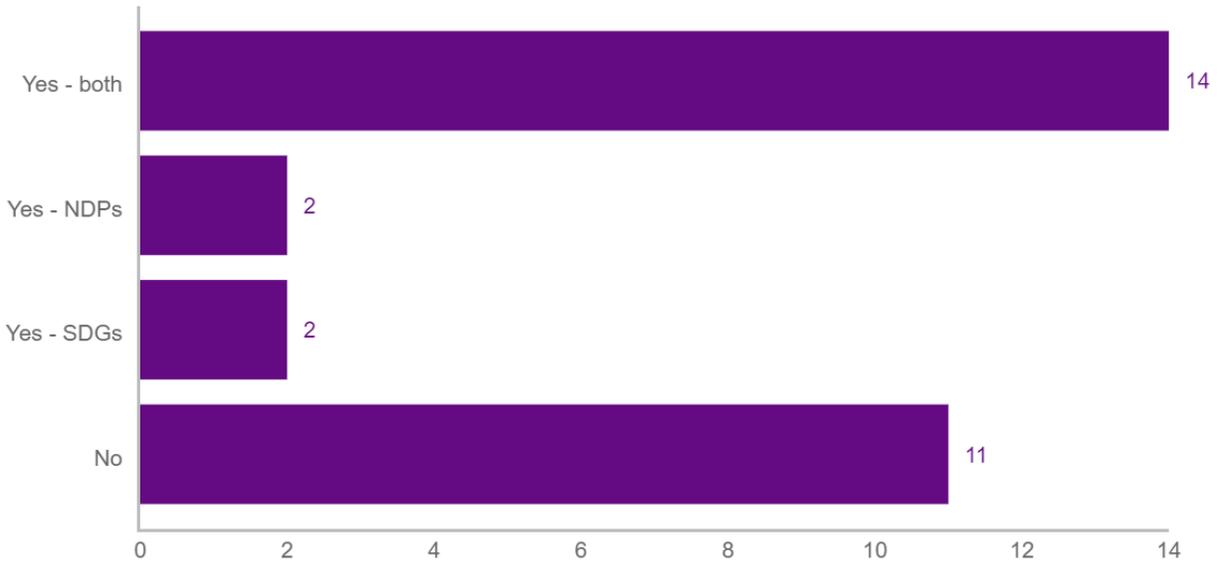
Figure 18. *Displays the number of library systems that collect data to support either National Development Plans (NDPs) or Sustainable Development Goals (SDGs). All 29 participants responded to this question.*

When asked to list or describe the types of development data they collect (Q2.14), a common theme was collection of data to support development goals in education. At least eight respondents mentioned education-related data points such as the number of schools in a given area, number of books given to schools, and programs to support reading and digital literacy. Another common theme was collecting development data focused on their local communities such as the access needs of marginalized populations, the number of local health centers, and the availability of electricity and other utilities. These themes support the role of libraries as development partners in education and community-embedded institutions.

Have Data Practices Changed to Support Development Goals
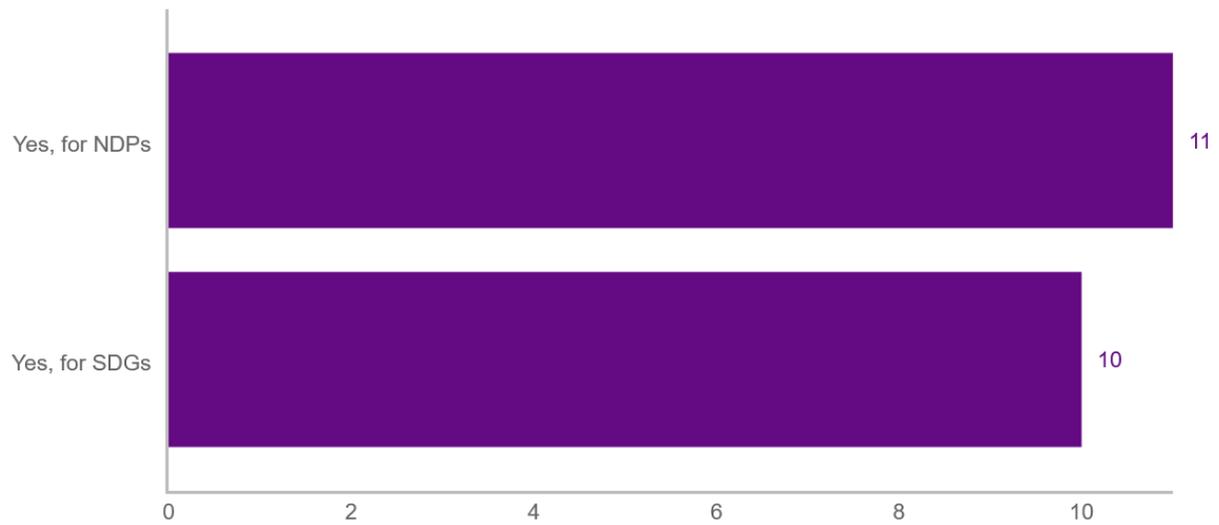
Figure 19. *Displays the sum of library systems that have changed their data collection practices so that they better align with or support either the Sustainable Development Goals (SDGs) or a National Development Plan (NDP). 25 participants responded to this question.*

Notable proportions of participants report that data collected has been changed so that they better align with or support development goals: 38% for NDPs and 34% for SDGs (Figure 19). This may further support the potential of libraries as partners in development: many have already shifted their data collection to be driven by development policies, and so these library systems are well-positioned to contribute desired development data to agencies.

3.7 Connections to Development Organization Interviews

3.7.1 What are some of the connections with interviews?

These responses suggest that most libraries already collect the types of data we recommended based on our development organization interviews. We previously recommended that African libraries collect quantitative data on their output (e.g. visits, services) with a special focus on usage and demographics of their users (Lynch et al., 2020a). Responses to this survey indicate many libraries collect data on number of library visits, the number of users attending programs, and user demographics. Furthermore, many respondents have found that quantitative data on users, programs, and facilities has been persuasive for internal and external stakeholders, which also supports our conclusions from development organization interviews that quantitative data is useful to demonstrate success. Overall, these connections suggest that libraries already have data that is beneficial for attracting partnerships with development organizations.

Two areas of potential disconnect between African libraries and development organizations may lie in (1) the use of impact data and (2) how libraries may currently be positioning themselves. Based on our interviews, development organizations expressed strong desire for impact data but a willingness to accept output data as a proxy. Throughout the survey, however, there were multiple mentions of library systems collecting data on library impact, such as assessment of programs or user satisfaction surveys (Q2.17, 2.25, 3.12, 3.13, 5.8). Given that only three respondents mentioned that impact data was persuasive for advocacy with external entities, library systems may not be aware of the full value of impact data in attracting partnerships. Furthermore, libraries may not feel capable of assessing impact; when reporting staff training needs, two respondents mentioned impact assessment, with one stating the need as "data analysis skills for librarians to determine the impact of library serve [sic] on the communities they serve." This suggests possible benefits of training related to measuring impact of not only improving library services but also, as one respondent affirms, "Organizations are likely to fund projects where they could see impact data."

Development organization interviews also disclosed a desire for partners that are capable and functional, with local expertise and insight into communities. In contrast, respondents tended to emphasize their experience that data about their needs is the most persuasive for partnerships with external entities. This may represent a gap in expectations between development organizations and libraries, in which libraries are overemphasizing needs instead of capability. This may have the unintended effect of amplifying perception of the weaknesses of libraries, rather than of their suitability as development partners. Though respondents report that they have found expressing their needs to be persuasive, they may solicit more partnership opportunities by marketing their successes, such as innovative programming, community impact, growth in visitor ship, or similar data that promotes them as capable partners in development.

3.7.2 What are next steps towards bridging the gap between libraries and development orgs?

If a majority of libraries already collect the types of data that match the desires of development organizations, a next step would be making that data visible to potential partners. A data platform, or an adjustment to an existing platform, could include the following variables:
- Average visitors per day/week/month/year (unit chosen by library)
- Average visitors age
- Percentage of visitors by gender
- Number of visitors attending most recent/specified program (chosen by library)
- Operational data including number of patrons, print materials circulation, public internet, number of staff computers, and public service hours or other commonly collected types
- Impact data including links to most the library's recent annual report, and/or impact assessment if available
- Others, illuminated through further analysis

In conjunction with these efforts, another step could be capacity building around the sharing of data in ways that empower libraries to analyze, report, and disseminate their own data on their terms. Though there is some hesitation around data sharing and its risks, multiple respondents expressed a need for these additional skills. As one respondent states, "My organization needs to have a way of **disseminating the data to the public and other organizations**. Data should be readily available if needed by external entities. [Country's national library service] needs to align data collection with development strategies in order to provide relevant services." This training could include specific modules on assessing impact, understanding that output data can be used to illustrate impact, in order to better tell the story of libraries' role in African development. Training could also make use of existing resources, such as the IFLA Storytelling Manual (IFLA, 2018) which library systems can use to collect, frame, and compose stories about how they are making differences in the lives of their users as proof of how they are impacting development in their communities.

4 Conclusion: Data as Agency

Overall, responses reveal considerable variety in data practices in library systems across the African continent. This variety is to be expected, but consistent themes have emerged in responses to our survey: data of multiple types is collected purposefully, usually with regards to library users and their needs or operational functioning of the libraries. Library systems see the value of data and have had some success using it instrumentally with various stakeholders, but there are still many challenges around staff capacity for analyzing, sharing, and using data effectively. In spite of challenges, many library systems are well-positioned to partner with development organizations and contribute desired data. To continue to foster these connections, library advocacy organizations can provide increased training and support in the analysis and dissemination of data with attention to library impact. However, these supports must also cultivate local ownership and agency with data in order to sustain self-advocacy efforts and continue to benefit the communities that African library systems serve.

References

Fellows M, Coward C and Sears R (2012) Beyond access: perceptions of libraries as development partners. Seattle: Technology and Social Change Group, University of Washington Information School. Available at: https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/20916/BeyondAccess _Perceptions.pdf?sequence=1.

IFLA (2018) Libraries and the sustainable development goals: a storytelling manual. Retrieved from https://www.ifla.org/files/assets/hq/topics/libraries-development/documents/sdg-storytelling-manual.pdf.

Isaacks V (2020) Public libraries and dissemination of information at correctional services. 24 January 2020. Accessed 29 July 2020. https://web.aflia.net/public-libraries-and-dissemination-of-information-at-correctional-services-victoria-isaacks/

Lynch R, Young JC, Jowaisas C, Boakye-Achampong S. & Sam J (2020a) African libraries in development: Perceptions and possibilities. *International Information & Library Review*, DOI: 10.1080/10572317.2020.1840002

Lynch R, Young JC, Jowaisas C, Rothschild C, Garrido M, Sam J, and Boakye-Achampong S (2020b) Data challenges for public libraries: African perspectives and the social context of knowledge. *Information Development*. https://doi.org/10.1177/0266666920907118

Mureithi M (2020) Libraries as youth empowerment centres. 24 January 2020. Accessed 8 March 2021. https://web.aflia.net/libraries-as-youth-empowerment-centres-miriam-w-mureithi/

Okojie V and Okiy R (2019) Public libraries and the development agenda In Nigeria. Paper presented at: IFLA WLIC 2019 - Athens, Greece - Libraries: dialogue for change in Session 233 - Africa Section. From http://library.ifla.org/2496/1/233-okojie-en.pdf

Appendix: Survey Instrument

Due to its length, the complete survey is not included here. Please find the survey instrument here.
Survey instrument:
https://digital.lib.washington.edu/researchworks/bitstream/handle/1773/46926/Advancing_Library_visibility_in_Africa_-_Data_Survey.pdf?sequence=1&isAllowed=y
ALVA project website:
https://tascha.uw.edu/projects/advancing-library-visibility-in-africa/