What Can We Learn From CAAs' COVID-19 Experiences?

DATA INFORMED COMMUNITY ACTION AGENCIES

June 2022 Report

By Paige Teegarden, MPP, CEO CSST Software, LLC

Helping People. Changing Lives.
community Action PARTNERSHIP
AMERICA'S POVERTY FIGHTING NETWORK

TABLE OF CONTENTS

Introduction 3

Background 6

Impact of the COVID-19 Pandemic 9

Data-Informed 13

Data Capacity 13

Leadership 15

Motivations for Strategic Development of Data Capacity 16

Resources Committed 17

The “Data Professional” Position at CAAs 18

Org. Structures Supporting Data Professionals’ Success 25

The Data Management Pipeline and Data Governance 28

Technical Approaches to Accessing Data 28

Conclusions 28

Helping People. Changing Lives.
community Action PARTNERSHIP
AMERICA'S POVERTY FIGHTING NETWORK

INTRODUCTION

This report summarizes the highlights of interviews with staff that have data management responsibilities at Community Action Agencies (CAAs) that were recommended as exceptional in their use of and/or aspirations for using data. This report focuses on the collection, management and use of customer-related data and information to inform key decisions about programs, partnerships, operations and more. This work was done in the context of the COVID-19 pandemic and CAAs' use of CARES Act funding. We focused our questions on the following, with particular emphasis on how the COVID-19 pandemic motivated changes in data collection and analysis or presented new barriers to agencies' use of data:

- What is required to shift the mindset around data to 'data as an asset' and what does it mean to be a data-informed Community Action Agency?
- What are the skills and knowledge of the critical players in helping improve CAA data capacity?
- What operational processes and investments appear to be advantageous for increasing the use of data?

BACKGROUND

Despite the inherent difficulty in managing multiple data streams, which is the norm for most Community Action Agencies (CAAs), there has been a clear call to action to use the data CAAs collect. COVID-induced shutdowns and the subsequent shift to a contactless provision of services forced CAAs to examine their existing technology infrastructure. At the same time, CARES Act funding provided a means of improving the use of technology to help with data collection and data management.

The National Community Action Partnership seeks to learn from CAAs' experiences shifting to contactless work, and to push the network's data capacity forward to prepare for the next crisis. In March 2021, NCAP organized the Data Ambassadors Cohort as part of the Network's response to the pandemic. The cohort consists of CAA data professionals focused on upskilling organizational data management abilities throughout the Network. These professionals seek to provide high-level guidance to the agencies in our Network to increase capacities for problem-solving, innovation, and continuous improvement in their practices, a need that has been exacerbated due to the pandemic. The work of the Data Ambassador cohort, of which this set of interviews and report are a part, was born in this context and seeks to:

- Expand our knowledge of what it takes for

CAAs to build data capacity

- Elevate promising practices
- Provide access to tools and resources for agencies seeking to build data capacity

IMPACT OF THE COVID-19 PANDEMIC

In many of the agencies we interviewed, as with most CAAs across the country, the COVID-19 pandemic brought with it intense operational challenges. In almost all cases, our interviewees turned to technology to help continue serving their communities. In a few cases, the operational stresses wrought by COVID even opened up opportunities for leaps forward in data management and analysis at agencies. For example, although Wayne Metro Community Action Agency (or Wayne Metro) of Detroit planned to centralize and digitize their intake process, this effort was pushed forward by as much as a year as a result of the pandemic.

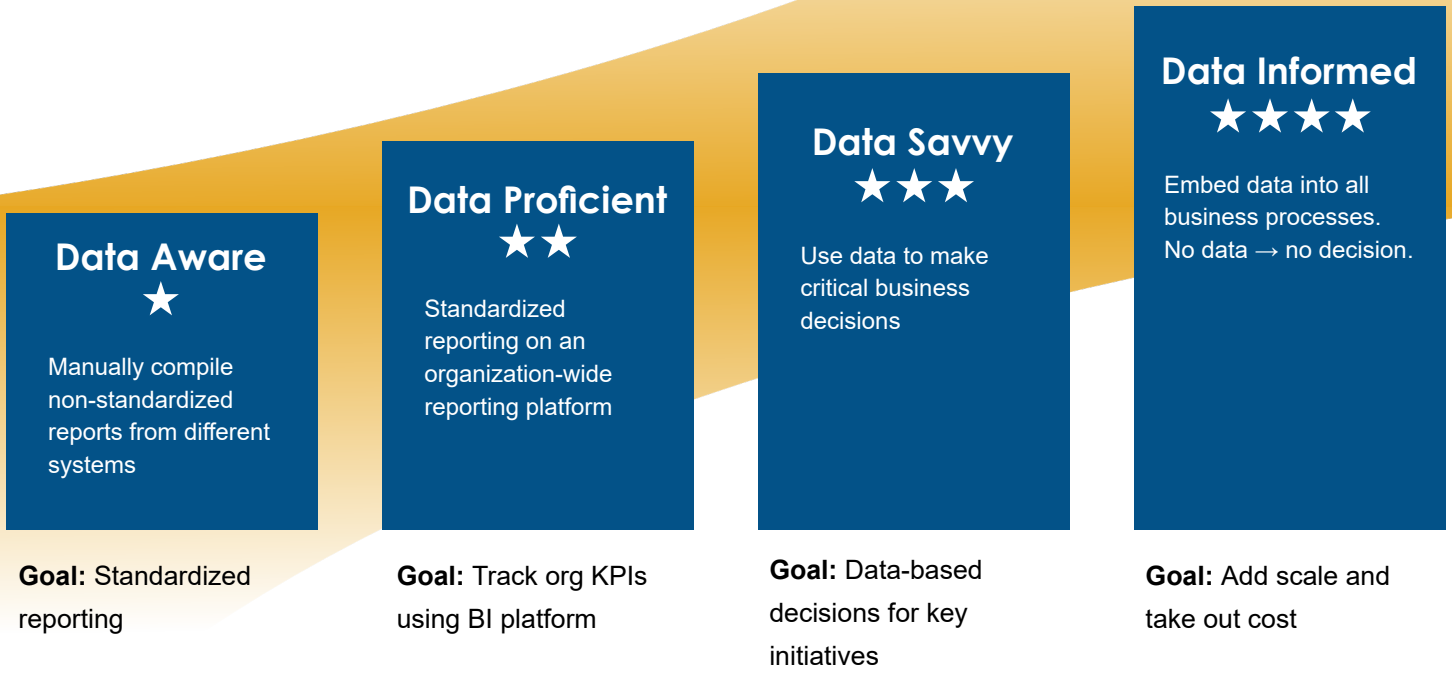
The impacts of the COVID-19 pandemic yielded advances in data use from several directions:

- Digitizing and centralizing data made it easier to use
- Customers and staff simply got better at electronic interaction, whether it was for groceries or services
- Agencies were more likely to have better customer contact information, which can be leveraged to gather more information
- The lack of alternatives to technological platforms for service delivery (no physical contact during large periods of time during COVID) forced all parties to fully embrace electronic interaction, overcoming the typical hesitation that comes with new technologies

DATA-INFORMED

Being a data-informed organization means using data (as opposed to intuition, personal experience or opinion) to drive the actions and decisions of the agency. It involves using data to inform strategic decisions, and to continually improve operational and program processes. It is a journey that organizations undertake, and not a result that happens overnight (see image above). The journey allows development of not only the tools, but also the culture and mindset shift that needs to happen. Becoming more data-informed yields a stronger, more resilient organization able to capitalize on opportunities and mitigate challenges. The COVID-19 pandemic demonstrated how data-informed CAAs were ready to quickly adapt practices, using timely data to inform service changes that met evolving community needs.

In the context of Community Action, being data-informed includes doing the following:



1. Using data to inform your local theory of change

2. Analyzing data about services, outcomes and customer demographics (for example what demographic profile receives what services and outcomes) to make shifts in program services and operations and to ensure equity in addressing the needs of all community members, especially at the start of the COVID-19 pandemic

3. Using data to continuously monitor operations in order to make improvements in service and efficiency (which was happening at a rapid pace at the beginning of the COVID-19 pandemic when CAAs were pivoting to meet the emergency needs of customers and communities) and meet performance targets

4. Using data to understand needs and opportunities in your community

The Data Ambassadors provide valuable front-line perspectives regarding the important role of data in Community Action. During the COVID-19 pandemic, data regarding customer needs, the delivery of services and the achievement of appropriate customer outcomes was critical to driving effective COVID response strategies.

A Data Analyst and member of the NCAP Data Ambassador Cohort from Action for a Better Community in New York put it this way: “How do we know we are being effective? What is the

data telling us and why?" Digging a little deeper into "what it means to use data," the Integrated Services Coordinator at YCCAC and another NCAP Data Ambassador, says: "It's about asking questions and making decisions based on data. You have to ask questions about success and where we are going, and then you have to dig into the data to inform operations."

The Vice President of Planning and Evaluation at People Inc., a CAA in Virginia, stresses the importance of investing in quality data so the agency can make informed decisions and use data to continuously improve its services. It's not something that will just happen on its own.

One example of how to begin the work of becoming data-informed comes from the Executive Director for Information Strategy at Wayne Metro. "We have principles (noted below) as part of our data strategy that really get at what being data-informed means to us."

- Align true need with resources to drive outcomes
- Streamline data usage to enable forward looking services and practices
- Employ adaptive program design and funding mechanisms

Finally, the Data and Policy Manager at Utah Community Action highlighted the importance of presenting data in ways that make sense to others so that they are able to take data into account when making decisions.

How Do You Know If Your Agency is Data-Informed?

- **Has your agency eliminated a program that was showing poor outcomes or usage?**
- **Has the agency made an adjustment to program operations based on service and outcome information?**
- **Do staff at multiple levels ask for and receive data to help them support their work? Do they approach this data with curiosity?**
- **Do all managers and leaders communicate expectations and follow up on these expectations by holding staff accountable to them?**

DATA CAPACITY

In order to be data-informed, an agency must have data capacity: the ability to get the right data to the right people at the right time - and in the right format for people to use that data. During the COVID-19 pandemic, when community conditions and customer needs changed frequently (sometimes daily), it was especially important for CAAs to preserve or increase their data capacity.

Right Data

The "Right Data" fundamentally means accurate, consistent data that measures what you think it measures. It requires using valid data collection tools and ensuring that the data points being collected are understood in the same way by those using the data.

While this seems straightforward, it usually entails grappling with which data definitions to use for terms like "Income," who to include as household members, or even who is receiving the service within a household (i.e., does "Income" refer to gross or net income, and how do staff know this?).

Another example might be housing status and making sure that what is meant by 'homeless' is understood in the same way by anyone collecting, analyzing and acting on the data (i.e., is couch surfing homelessness?).

Data capacity means the ability to get the right data to the right people at the right time.

The Right Data also implies that data is collected and stored at sufficiently specific levels (lowest reasonable level of specificity) to be useful for analysis. For example, it is more effective to store date of birth rather than age, because the former will remain unchanged over time and not requiring updating. The Right Data needs to be defined consistently across the agency so that conflicting definitions can be surfaced and resolved.

Right People

The "Right People" includes people at all levels of the organization, including customers. It should include every stakeholder that will act on the data (staff), whose interests are represented by the data (customers), or who will use data for organizational accountability (leadership/Board members). For example:

- Providing **customers** with information about programs and requirements can help them make decisions about engaging with the CAA

- Providing **frontline workers** with data about the actions of their recent customers and when to follow up can help them be more efficient and effective

- Providing **managers** with data about which staff are completing which types of services and achieving which outcomes can yield more effective workflows

- Providing **leadership** with summaries of progress and results across the agency, sometimes in comparison with targets or agency goals, promotes accountability

- Providing **all stakeholders** with information about both agency activities and results in an aggregated, easily comprehensible way promotes CAA effectiveness and the establishment of more effective community partnerships

Right Format

The roles that people play inform which data is needed and in what format it needs to be for them to use it effectively. In general, the more strategic the role, the more aggregated the information. Note, however, that aggregated information does not mean lumping all customers together into a single number. Some important agency strategic priorities – such as the advancing of equity – will require leadership to consider data analyzed in terms of different groups. These cases will permit leadership to adopt strategies that advance "targeted universalism:" the idea that different groups in the community may require different approaches in order for everyone to achieve equitable community standing.

In general, however, staff considering agency strategy and change will usually look at data about groups of people – and comparisons between groups - rather than individual-level data, whereas staff providing more direct or hands-on assistance are more likely to need to view data about individuals and families in real time.

Consider the following example: A case manager needs to see the data for the particular household or person with whom they are working to determine effective service delivery. By contrast, an executive director or board member generally needs to see counts and percentages of particular data (aggregated information) to understand program/agency performance trends. Again, sometimes leaders will need disaggregated data to ensure CAA services and outcomes advance equity in the community. Finally, in some cases, funders may require tabular counts of services or demographic profiles, whereas community-facing reports will usually present information in graphs, images or infographics more easily accessible by the broader community.

Right Time

This means the data should be available to people when they need it to make a decision. It doesn't matter if you can produce a beautiful graphical and narrative analysis if it is not available to review and make decisions until after the meeting. In another example, if the data to successfully offer an intervention is not entered in time for the staff that will actually engage with the customer and offer an appropriate intervention, then collecting that information loses significant value. For example, if a coordinator or doctor or navigator who is responsible for offering assistance with issues like homelessness does not know the person is homeless when they meet with them because that data has not yet been entered or is not visible to them, then they either have to waste time re-asking questions or they miss the opportunity to offer timely assistance.

LEADERSHIP

COVID-19 has exemplified how important it is that leaders understand the value data brings to the decision-making process. Perhaps one of the most important prerequisites for moving an organization toward greater use of data, including using the data the organization collects, is for the organization's leadership - both the staff leadership team and the Board of Directors - to understand the value of data and how that value can benefit an agency. Then that leadership team must prioritize the systems and resources necessary to appropriately collect, manage and use customer data. This has been noted in literature on the topic (eg. Urban Institute's Continuous Quality Improvement Study of Head Start Programs), as well as by the agencies interviewed for this paper and in stories of agencies' failures along the path to using data.

The following are core leadership actions in data-informed nonprofit organizations:

- Serving as a role model(s) for data use
- Distributing data responsibilities to staff at all levels
- Motivating staff to use data
- Communicating expectations around data use clearly and holding staff accountable to those expectations

In addition to the core actions above, leadership staff of Community Action Agencies can advance their organizational data culture by reinforcing the following:

- Demonstrating a consistent interest and ability to articulate a strategic, agency-wide perspective on data collection and analysis
- Connecting data collection and analysis to a

clear strategic vision

- Promoting integrated agency services (breaking down programmatic silos) by developing a willingness and ability to weave together a shared vision and motivation regarding data use

Examples of leadership driving data use

At high-performing data-informed agencies, top leadership supports, and usually explicitly engages questions of agency data and strategy. Leadership support was and remains essential during the COVID-19 pandemic. The Data Ambassadors provide several concrete examples of how leadership can effectively drive increased CAA data capacity:

- At Action for a Better Community in New York, the CEO asked for a monthly outcome report, and is shifting the conversation to include a focus on results across programs. The questions being asked, and the data requested, are helping drive greater interest in collecting accurate data and effectively using the data that is collected.

- At YCCAC, the leadership team decided they wanted to focus on cross-agency referrals and warm hand-offs. Over the course of months of training and meeting with staff, they were still seeing relatively low referral numbers. Once it was clear that there was no underlying training issue, the leadership team engaged in an honest conversation about accountability and operational management. This is an example of the interrelated role that leadership plays in encouraging the use of data: a leadership decision leads to an operational implementation about recording data, which leads to challenges around that data being recorded, which, in turn, leads back to management and leadership's response.

- At Hawkeye Area Community Action Agency (HACAP) of Hiawatha, IA, a new executive director sought to move their agency toward a philosophy of "We are one agency, we are HACAP." Originally, program directors and other staff used terms like "my families" to refer to their customers, and acted on that sentiment by protecting their program, operational processes and data. After data showed families were getting an average of three services across the agency's programs, this mentality began to shift. The concrete use of simple service data helped generate a shift in underlying organizational beliefs and attitudes.

- At People Inc., there is a board evaluation committee that reviews customer trend data, prompting an agency-wide examination of services and outcomes data on a regular basis. Engagement from the board further highlights the importance of the data being

collected and encourages new analysis of data.

MOTIVATIONS FOR STRATEGIC DEVELOPMENT OF DATA CAPACITY

The need for CAAs to holistically understand who they serve, what they do and what their impact has been embedded in the ROMA approach since its inception. (ROMA, or Results Oriented Management and Accountability, is a cyclical, data-informed decision-making framework that leads to continuous growth and improvement.) Despite this, many agencies struggle to effectively develop the data capacity inherent in this need. Looking at some of our top performing CAAs, ***what was the impetus for them to commit resources to building their internal data capacity?***

Prioritizing resources for any project or need involves both strategic and operational considerations. For the decision to build data capacity, an example of a strategic decision is needing to understand who you serve in order to better serve the community. An operational consideration may be the need to more efficiently produce reports that are useful for management and funders. Within the agencies interviewed, we find clear examples of strategic and operational considerations like those highlighted above.

There was also a specific driving impetus within each of the Data Ambassadors' agencies. The COVID-19 pandemic heightened the need for even more significant increases in organizational data capacity. The following are some of the drivers of change raised by our interviewees:

- Newer executive directors who wanted to get a clear "picture" of the agency
- Issues or questions about measuring success and the agencies' data raised in Strategic Plan or Audit/Assessment
- A Strategic Plan that highlighted a need for holistic services, and/or an explicit decision to increase cross-communication and coordination among agency programs in order to serve the community more effectively and efficiently
- Wanting/needing compelling numbers and stories for stakeholders (a greater ability to tell the agency's story) and thereby increase funding and other support
- Accurate CSBG Reporting
- The need to use data in new ways as a result of the pandemic and/or natural disasters

Several of our interviewees talked about new

leaders (executive directors, deputy directors or critical program directors) and their priorities as "the reason" for the shift toward using data and the subsequent investment of resources in these efforts. Often this new leadership, combined with a current strategic planning process or assessment/data audit, created a case for beginning to seriously consider the agency's use of data.

In three interviewees' agencies, incoming executive directors focused on their agencies' data capacity as a key strategy in promoting collaboration and a coordinated implementation of the agency's strategic goals.

- At Tioga Community Action, New York, a newer executive director came into the position with a particular interest in driving greater cross-departmental communication and collaboration and promoting an understanding of how an agency-wide data approach could facilitate those goals

- At Utah Community Action, a new executive director arrived with a history of work in social services and wanted to facilitate a collaborative, cross- functional approach in the agency. She teamed up with an agency data expert (a Data Ambassador) to utilize data in service of this approach

- At HACAP, Iowa, a new executive director formed a strategic leadership team to help focus on the agency-wide mission and strategies

Leaders other than executive directors also play a key role in supporting the strategic development of data capacity. For example, at York County Community Action Corporation (YCCAC) in Maine, a new deputy director with experience in Results Based Accountability (RBA) and a program director responsible for CSBG with expertise in organizational change reporting came together during their agency's strategic planning process to focus on how best to articulate and measuring success. Over time, this collaboration led to a focus on de-siloing their data and spreading data capacity and use throughout the agency.

Agencywide audits and/or risk assessments can also drive the development of greater data capacity. At Wayne Metro, an IT Security Audit led to an acknowledgement that the agency's data was dispersed and there was no clear 'ownership' (or governance) over the data. This in turn led to the creation of a Director of Data Strategy position, which, in turn, facilitated the writing of a data strategy and plan.

In several of the interviewees' CAAs, the desire for accurate and timely reporting of CSBG service indicators and outcomes (enhanced by the need to track CARES-funded services during the COVID-19 pandemic) played a role in helping motivate a focus on agency-wide data

rather than program-specific information. At ABC in New York, Gulf Coast Community Services Association (GCCSA) in Texas, and Community Action Partnership of Lancaster and Saunders Counties (CAPSLC) in Nebraska, there was an explicit desire to better understand the data being reported for CSBG and use it to inform decision-making internally.

Our interviewees cited the COVID-19 pandemic as pushing a new reliance on technology and the data consolidation that it can help achieve. In addition, natural disasters in the Midwest and Hurricane Harvey in the South created an impetus to shift processes and use data collected from customers and from the community to make decisions.

RESOURCES COMMITTED

Whatever the impetus or motivation to focus on data use and data capacity, all agencies that have made significant progress toward becoming data-informed committed explicit resources to that work.

Across the agencies interviewed for this paper, the percentage of the agency's overall annual operating budget spent on data systems, data management and data capacity ranged from 0.5% to 3.9% of the agency's annual budget, with an average of 1.3%. In absolute dollars, the investment ranged from about $50,000 to $500,000 annually, with the average being $234,000 annually. These resources directly funded data-related staff positions focusing on agency-wide data and information, paid for enterprise customer tracking systems and technology customization projects, and funded training and other capacity-building efforts for staff.

While not every CAA has the discretionary resources available to invest at these levels, there are ways to support both individual agency data capacity building as well as supporting the efforts of the entire network. Examples of strategies beyond individual agencies spending significant funds include:

- Supporting associations in developing their data capacity, which can then be contracted out to agencies or used to provide local training and other support
- Educating State CSBG Offices about how they can support local agencies' data capacity through data collection guidance and training, as well as through their administrative or discretionary spending
- Increasing the focus on data capacity within ROMA Implementer and Trainer curriculums
- Gathering a cadre of consultants who have worked with local CAAs on data capacity and encouraging them to provide services to additional agencies in the network

- Provide funding explicitly for data capacity

THE "DATA PROFESSIONAL" POSITION AT CAAs

The existence of a staff person who can lead the charge on data collection, management and use is a critical resource in high-performing, data-informed CAAs. The interviewees for this project are in positions that have significant responsibilities for using and understanding the data that the agency collects as a whole. In most, but not all cases, it is a dedicated position that has direct access to agency leadership. Specifically, seven out of the ten agencies interviewed had at least one position fully dedicated to managing the agency's data. In about half of those cases, the data position was a member of the organization's Senior Leadership Team; the other half had direct and easy access to the top leadership in the agency as a direct report.

There were several common characteristics of the people we interviewed, which are summarized below. We delve deeper into these and how to potentially recruit and hire for these positions in a separate report about Data Expert Competencies (to be released in Spring 2022).

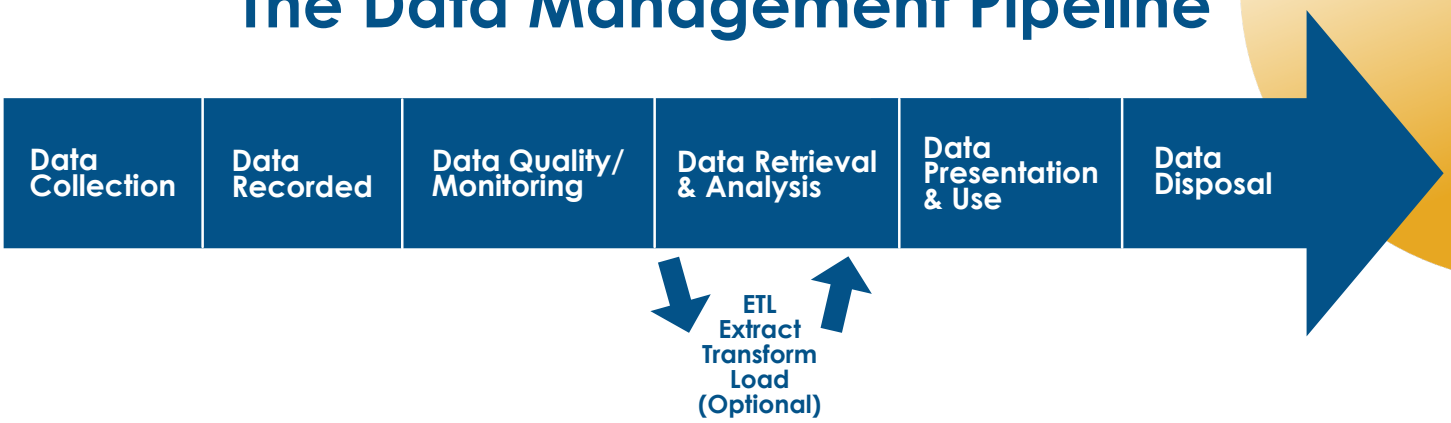
Characteristics of Successful Data Experts:

- Work experience in education, training and/or consulting
- A leadership style that is open, engaging, and collaborative
- Trusted and friendly relationships with people across the agency despite sometimes having to highlight opportunities for improvement
- Predisposition to getting things done and learning along the way
- Determined, unlikely to let barriers stop them from achieving their goals
- Genuinely curious about data and what it could mean
- Act as "translator" of program concerns to data points and collection methods to data/technology and back to program operations Able to connect big-picture questions to the details of operations; able to break the big picture into smaller understandable pieces

In some cases, where the positions were more focused on retrieving and analyzing the data, the staff possess a background in database administration, data analysis, data science or IT. In these cases, understanding tools like SQL and/or Microsoft Power BI is critical.

In addition to data positions that operate at a high level in the agency, there can be program specific data staff who focus on pulling,

The Data Management Pipeline



analyzing and reporting data for specific programs. It is important to note that the type of cross-program, agencywide management and analysis of data advocated for in this paper requires a position that is operating at a higher, agency-wide level.

ORGANIZATIONAL STRUCTURES SUPPORTING DATA PROFESSIONALS' SUCCESS

Even with supportive leadership and the right person in a data expert role, it is critical to have supporting infrastructure within the organization. Think of this as the mechanisms through which expectations about data are communicated and analyzed, and decisions using data happen.

First, agencies had a well-defined leadership team that team met regularly, most often monthly, to consider strategic questions for the agency. These leadership teams typically included the executive director/CEO, CFO (or other financial leadership), operational and/or compliance leadership (if applicable), and directors of the major departments or divisions. In most cases, these leadership teams had a clear 'charge' or purpose that included the strategic direction and implementation of new initiatives, financial oversight and risk assessment. These teams often spent time explicitly reviewing service, outcome and operational quality data related to strategic decision-making. In many cases, leadership teams met even more often to deliberate on operational questions caused by the COVID-19 pandemic. The leadership teams were unlikely to get into program-specific data except as it related to strategic issues and questions.

In several cases, additional defined teams met and focused on specific agency-wide and programmatic data. These teams might meet every week or every other week and get into the

nuts and bolts of making meaning out of the data. Examples of these second teams include an integrated services team at YCCAC, which grew out of their strategic plan and meets bi-weekly; a weekly general leadership meeting and weekly/biweekly departmental and inter-departmental meetings at Wayne Metro; a management team meeting at Utah Community Action; and a twice-monthly strategic leadership team meeting at HACAP, which includes directors and (at times) program managers and focuses on building up the agency's management and leadership capacities. In all of these examples, integrating data management plans, strategic analysis, and a simple review of numbers is a critical and consistent part of the agenda. One important consideration is who controls the agenda for these meetings and how often customer data plays a part in them.

Even in agencies where there were not formal secondary structures that had a data component, data staff met regularly with different programs and departments. They often would focus on a particular program's data, asking questions about it and ensuring the department had the information it needed and understood said information.

One example of this is in ABC, NY, where staff meet with departments to go over outcomes reports that go to executive staff. In other cases (like at GCCSA in Texas), the submission of monthly and quarterly CSBG data is a natural time to think about services and outcomes across the agency and discuss possible operational shifts that might be necessary.

THE DATA MANAGEMENT PIPELINE AND DATA GOVERNANCE

The data management pipeline or lifecycle refers to the stages through which data passes in a typical CAA. These are the stages people within the organization must understand. As the name implies, agency staff need to understand and manage the data through these processes, including identifying who is responsible at different stages and potentially considering the question of data 'ownership' (who is responsible for the integrity and accuracy of specific data). Agencies need to ensure that their systems - both technology systems and human systems - which manage these stages for different types of data are secure, consistent, and consistently understood by stakeholders. Often, these conversations are handled in Data Governance discussions and documentation.

The corporate sector is home to a variety of complex models and frameworks for data management, data management job positions/roles and associations. While many of those tools and ideas are relevant to Community Action, they are generally positioned for people with significantly greater data and technical skill

than is reasonable to expect for most CAAs. Further, the concept of data management, particularly within the context of organization management and the different stages of data processing, is something that has not been well defined or attended to by most nonprofit agencies. The model presented on page 14 is an adaptation of these corporate concepts for CAAs, and is being presented to focus attention on the need to manage data, which includes understanding the relationship of data processes and standard program operations.

In our interviews, there are examples of attention being paid to a variety of points on the Data Management Pipeline and, in most cases, there were indications that staff possessed an intuitive understanding of the pipeline. An interviewee at YCCAC said, “I’m involved in the entire lifecycle of data at the agency from intake through to analysis.” Or, as a Wayne Metro interviewee highlighted, “We are focused on process improvements and efficiencies, getting data to a variety of stakeholders with efficient methods. We work on providing information that is useful to the programs.” By contrast, another interviewee pointed to the challenges of data management for CAAs, stating that “we do a lot of data entry, but because we don’t control all the systems, we have difficulty getting data out, making [data management] more difficult.”

Other interviewees commented on the interrelated aspects of data collection and programmatic operations. One stressed the importance of buy-in across the agency and the need to integrate data collection into program processes. For example, Head Start programs have a clear operational model, success is well-defined, and data/metrics collected are well-understood. Consequently, data collection is well-integrated into operational processes. In other cases where those elements are not in place, it is much more difficult to achieve buy-in from staff to support the collection and use of operational data.

Interviewees at Utah Community Action, ABC in New York, and YCCAC in Maine all mentioned meeting with each program department to understand what metrics they want to look at regularly and then define what data must be collected to develop those metrics. There are several considerations during these discussions, including understanding why they want to look at specific information, considering if existing data collected can answer that ‘why’ question (or can be converted into a metric to understand the same topic), and if not, then what effort will be required to collect the information (and whether that effort worth it). Finally, if the data is being collected, then can the agency get access to that information in the format they need to answer the question? These conversations tie together most elements of the data management pipeline.

Whether it is called a workflow or a data entry plan, it's helpful to maintain a written document about:

- **When staff gather specific pieces of information**
- **Where that information is entered**
- **And at what time intervals that information is entered.**

Paying Attention to Data Collection

Data collection is fundamental to every part of the data management pipeline. If the information collected isn't accurate or isn't consistently applied across both the agency and its customers, it becomes impossible to draw accurate conclusions about what the data means.

YCCAC staff talk about written 'workflow' guides or instructions on exactly what information is requested at what point in interactions with customers, what the options for answers mean, how to verify that information and where to record the information in their data system. At People Inc, staff have developed data entry plans serving a similar purpose. These guides are part of ongoing training and support for staff who are responsible for gathering information from customers and/or verifying information that customers provide.

The YCCAC interviewee notes, "A business process or flow will happen: Staff will interact and gather information. If you don't define that process, then it will just vary from staff person to staff person, and it will be impossible to know what that data actually means."

Recording Data

Agencies must understand where the data they collect is being recorded, at what point in the service model it is being recorded, and by whom it is recorded. When data is in an unknown location or in a format that isn't accessible to those who need it, you can't guarantee the integrity of the data, and realistically, you can't efficiently use that data. For example, if you have information, but it is written on paper forms stacked on someone's desk or in an Excel file on a staff member's computer, you don't actually have control over that data and you can't easily use that information.

Recording data also means being clear about the definitive source for different types of information. This answers the question, "Where do you go for accurate information about demographics, about program services, about

specific outcomes?" Wherever that definitive source is located, it should be well-documented and 'enforced' (i.e. if you say it is XYZ software and agency leadership accepts data from another source, then XYZ is not the definitive source for that information).

Examples of Recording Data From Interviews

The System Administrator at CAPLSC in Nebraska discussed the clear expectations that are set when specific data is entered into their agency-wide system. Staff follows up and monitors program information based on these deadlines, which adds consequences for not meeting the expectation.

The Community Services Coordinator for Douglas County, Colorado discussed the time spent meeting with partner agencies up front as part of the rollout of their joint data system. In this case, staff is providing the data platform for the partners' use, and this initial setup time is focused on ensuring that staff are able to configure the system to capture information needed for funders. It also ensures that the agency, as a facilitator of collaboration, is able to capture the information that will be useful to the partnership. This upfront time results in a closer match between the processes of data collection, recording data and reporting data.

At Wayne Metro, Detroit, Michigan, staff decided to establish the definitive source of information by holding intake information, which includes demographic information and program/service requests, in the central system. After that point, more program-specific information may be held in program-required data systems.

Data Quality/Monitoring

Even when data collection and recording are well-defined, it is important to monitor and use data regularly. The best methods of data quality and monitoring are integrated to the analysis and use of the data, since having accurate, good-quality data that is never used is a waste of resources. Agencies should define who is overseeing data and at what frequency, and make sure these responsibilities are well-known by agency stakeholders.

In one example, staff at CAPLSC in Nebraska monitors data every month and analyzes information against service and outcome targets and looks at trends over time. If any of these show a cause for concern or raise questions, staff has discussions with the program manager; If there are still questions about data, then the executive director will get involved in helping to figure out what's happening. This helps to ensure the data is entered accurately and in a timely manner. CAPLSC staff also highlight that it is more important that information be accurate than having complete information. In other words, don't be afraid of null or empty data.

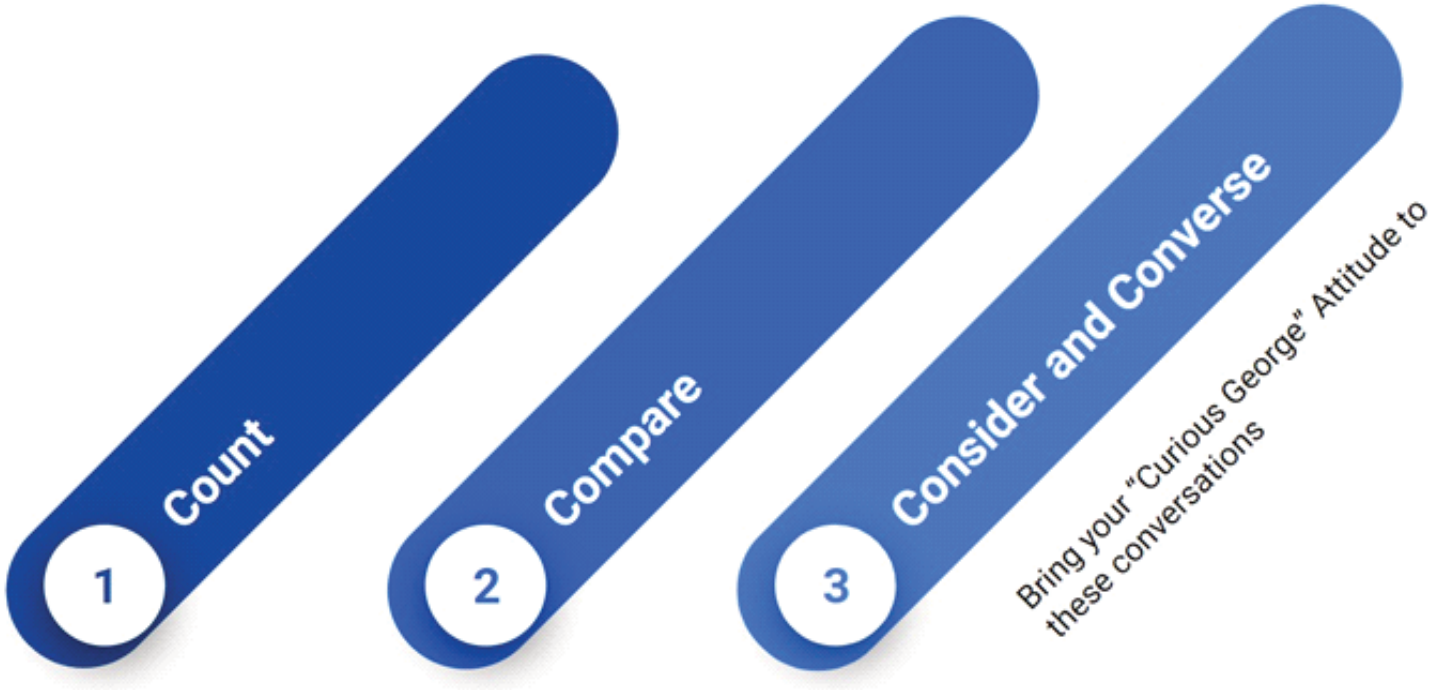
In another example, after months of not seeing increases in cross-departmental referrals and determining that it was not a training or technology issue, YCCAC senior leadership made the decision that frontline staff which have responsibility for entering data about referrals now report referral numbers and the reasons for the numbers directly to their supervisor. This is another form of data monitoring.

Data Retrieval and Analysis

Retrieving and in some cases consolidating data can be challenging for Community Action Agencies. We review some of the methods of retrieving and consolidating data in the next section. For the purposes of understanding this stage in the data management pipeline, it is important simply to recognize and support the skills needed to retrieve data and make sense of that data. This is not solely an analytic skill: It requires engagement with agency and program leadership to understand the types of questions that are trying to be answered, which in turn impacts the type of analysis that might be done. Once there are numbers based on analysis, it will require engagement from programmatic leaders to help make meaning of them.

A simple diagram of the analysis stage is represented below: Count 'things,' Compare the things you are counting over time or across groups, and then Consider what that information might mean or what actions it might suggest.

For example, at ABC, New York, staff was



analyzing data about outcomes achieved by different case workers in the Bridges2Success Program. The program has a mentorship and incentive component. An initial review of the data revealed that two case managers were achieving significantly more outcomes than other case managers. The next piece of analysis was to dig into why, and this required further engagement with staff with program knowledge. In this step, agency staff found that the two case managers with the higher number of outcomes awarded more incentive funds than others did. To truly understand what was happening, staff engaged in discussions about why others didn't have as much incentive funds distributed and to tease apart whether the incentives (and potentially the celebration/recognition they entailed) made a difference in customer behavior, helping them achieve greater outcomes.

Data Presentation/Reporting and Use

Presenting information, sometimes as a report and sometimes as graphs or visualizations, is part of using the data. The data needs to be in a format that people can see and understand in order to make decisions. In some cases, you simply need the data points aggregated in formats that are required by funders, although that should be the minimal use of your data.

To go beyond mere funder reporting, the team looking at the data (which usually, at this point, will be some aggregated count of data) should be looking for interesting numbers and comparisons, trying to figure out how to best present that information to the people who need it. One common example of this type of comparison is actual services or outcomes achieved against agency goals. When these figures are disparate, it is often an indication that something is different from what programs 'thought' would happen and might be worth discussing. This is only one type of comparison. In new programs or programs where there have been significant operational changes or the environment radically shifts (i.e. a pandemic shuts down offices), this kind of 'target vs. actual' comparison can be easily explained instead of being used as an opportunity for further discussion. Some common ways of displaying data and the kinds of questions they can help illuminate are displayed on p. 21.

People Inc, YCCAC, and CAPSLC have worked with senior leadership and departments to create dashboards from their data, which are reviewed each month or quarter. A few specific uses of data information from People, Inc:

- Demographics are used when looking at adding new programs, program expansion, or program location

- Service counts are used to answer questions like: "Is this what was expected? If not, does something need to shift?"

Data Visualization	Example Purpose
Heat maps to spatially display data	Concentration of usage or need, including under-served geographic areas
Bar graphs comparing the same data point across different categories, including shapshots in time	How do groups (defined by time or characteristic, such as housing status) compare in outcomes or services; often will have an idea that you are trying to test by looking at data
Line graphs to present continuous data over time	Average income (i.e., running 7-day total) over time, sometimes multiple lines on same chart, to show different groups and their income COVID cases and hospitalization rates are one ubiquitous example today
Pie charts display the relationship of part to the whole. Requires a small number (up to 5 or 6) of discrete categories	Ease of seeing the relative amount of a part of a whole. Can be used to get a quick picture of the characteristics of customers served, especially if you want to communicate that a particular part is 'large' in comparison to other parts
Median, mean and mode are measures of central tendance, or the "middle," of a set of data	Data must be ordinal. Can help to describe a data set, which is particularly important when comparing to different sets of data

- The Board wants to see outcomes… "Got the job" or "Purchased a house" are common examples. CSBG outcomes are about 50% of the outcomes that are tracked

- Data that documents progress or change tends to be more limited to specific programs

At Wayne Metro, a real-time dashboard of online applications for services during COVID, and the visualization of said real-time data, began the accelerated use of data and information by the agency. It also led to the agency altering programmatic processes to meet customer and staff needs. It showed stakeholders what is possible and how technology, data collection, and data presentation could be tied together to make their work easier.

Data Disposal/Privacy and Data Ownership

The last element of the data management pipeline is ensuring that data is disposed of in ways consistent with the data ownership and privacy expectations of those the CAA has served. Disposal could include de-identifying data for on-going use after some period of time. These processes are often less defined in CAAs and can require engagement with both funders and vendors of software systems. Customers are often receiving multiple services, and will come back for services multiple times over several years. Agencies need to be clear whether older program-related data will be kept

if the customer is still receiving other services. Agencies may want some data to be 'less' visible after a period of time, but still accessible in their systems. Defining the criteria for this can be challenging. Finally, agencies should note, if they completely purge data after a period of time, the agency loses access to a valuable asset. As an alternative, consider deidentifying the data through a data store or warehouse.

TECHNICAL APPROACHES TO ACCESSING DATA

As Paige Teegarden at empowOR says, "You have to have the data (in hand) to use data." This is a comment both on collecting and storing data and on consolidating data from various sources. Since many CAAs are still required by some funders to enter data into a specific software system, most agencies need strategies to consolidate data from different systems into one place in a way that ensures unduplicated customer counts. There is no single 'right' way to consolidate data that is being stored in different electronic systems. The techniques an agency uses will depend on their data infrastructure, organization's direction, resources, and the skills of staff. Ultimately, CAAs are likely to use a variety of approaches.

The first step to planning a data consolidation strategy is to conduct a customer data system inventory so that you understand how data about the people served (the people your agency has defined as 'customers') is collected, where it is stored, and whether different software systems are required or preferred. This is critical, because it will always prove more expensive to try to combine data from different database systems than it is to learn to enter the information into a system that accurately tracks distinct individuals and families. If multiple systems are required, consider a group advocacy effort to allow for the import of data into those systems.

Following are some of the methods used by our interviewees to consolidate data for analysis across various software systems.

Data Warehouses and Lakes / Interoperability Projects

In these approaches, data sources are submitting data regularly into a data lake or data warehouse, usually (although not always) in an automated fashion. These projects require a project lead who has some knowledge of ETL (Extract Transform Load) process and tools, members of the team (including vendors) that have SQL expertise, and program experts who understand what the data being transferred represents. Without a national data standard for CSBG-related data, these projects usually end up being one-offs where each agency is doing the data transformation and mapping work necessary for their specific systems.

Utah Community Action is pursuing a data lake approach to bringing their data together. They have two primary customer databases: A Customer Track for all social services data and Child Plus for Head Start. For Child Plus, they have a direct database access (DDA), which is a service that Child Plus offers costing approximately $2500 annually plus the cost of VPN tunnels needed to access the database. DDA access allows users to write SQL and retrieve data directly from the Child Plus database. Customer Track data is updated every 4 hours using a direct data pull.

The Utah Community Action Data and Policy Manager leads these processes and is responsible for meeting with internal program experts to understand the data from the source system that should be included in the data transfer to Domo, which they use for their data lake. This includes conveying sufficient information about how the data is stored in the source system so that the Data and Policy Manager can write SQL to retrieve the required data. Developing this understanding of systems data model and the data needed is the crux of this approach.

Manual File (Batch) Import / Exports to Business Intelligence Tools, Data Stores, or Customer Tracking Software

Exporting and importing data is another approach that is used in some cases, usually as part of a hybrid model. In these cases, agencies usually have limited access to the data in required funder systems, but have been able to export sufficient customer details (i.e., name, date of birth, address, partial social security number, etc.) to determine whether a customer matches a customer record in another data system, as well as information on services or outcomes with a unique customer ID across all exports. They then import this information into a business intelligence tool, a data warehouse/ store, or a primary customer data system.

YCCAC in Maine is pursuing a hybrid strategy that includes manual and automated exports and imports to a data store. Specifically, they have a required LIHEAP system which provides an export of .csv files that can be dropped into an SFTP server. There, it is automatically imported into empowOR's Data Store (called CADS). There is a similar arrangement with their Child Plus data. The LIHEAP process has required working with Maine State Housing (MSH), which holds the contract with the LIHEAP System. MSH has direct access to create reports to develop the .csv reports that meet the data store's csv standard, or create exports and have the data stores transform process manipulate those reports and import them. The Service Integration and data-informed Project Manager at York leads the team working on this project, and is responsible for mapping the data from the LIHEAP system to the data store .csv standard, while staff at Maine State

Housing are responsible for generating the csv files from the LIHEAP system. empowOR staff are responsible for validating the files and, in tandem with York staff, ensuring that the imports into the data store include all expected data. Costs include the expense for the Data Store, which is based on total number of customer records and number of transactions, as well as a small initial configuration cost.

Sometimes funder-required systems allow import of data. In these cases, agencies can collect data within a single software system and simply export data (according to funder specifications) for import into the funder's required system. These approaches require clear data standards from the funder along with data dictionaries, sample files and validation processes. HUD's 9902 and Housing Counseling systems are an example of these standards in federal programs. Within states, some Aging/Senior services (AAA/NAPIS) systems allow limited imports, some HMIS systems allow imports, and some CSBG Reporting systems allow imports.

Central Intake System with Duplicate Entry Into Program Systems

In this approach, customer-identifying information and demographics is entered into a central database at the time of intake. In some cases, customers are literally entering their own information into the data system through a public-facing page or portal.

This becomes the de facto 'master customer record.' Data is then entered by staff into all other required program-specific data systems. When manipulating this data, it can be helpful to have multiple monitors for staff so they can cut and paste information into the program-specific system. Sometimes the central intake system will include reports with data ordered in the same order as the program-specific system to help ease the data entry time. This approach can be combined with other approaches by allowing some data to be exported from the program specific systems and re-imported back into the 'master system.'

Wayne Metro is utilizing this strategy, building off a project initiated during the COVID-19 pandemic to provide customers with online access to apply for assistance. Wayne Metro started using Google Forms to collect basic intake / demographic and identifying information. They have now moved to using Form Assembly to create forms and CASPIO to process data. In the future, they will create forms directly in CASPIO. A staff member notes that "for most programs, we actually obtained waivers from the State Offices so that we can produce the required data and reporting straight out of CASPIO and not have to enter it elsewhere."

For some programs, Wayne Metro staff

members re-enter data. Leadership has ensured that staff who will be entering into required program specific-systems have two monitors so they can cut and paste from the Central Intake system into their systems and then add additional data. Currently there is no pulling of data from any program-specific systems into CASPIO. Wayne Metro's Director of Data Strategy leads the team for this work and is responsible for building out the CASPIO system. They are working with program directors to ensure that the staff who enter into other systems understand expectations and have the training to enter the data properly properly. The agency's annual cost for CASPIO last year was about $8,000.

Mixing and matching strategies is likely. Below are critical considerations if a CAA is consolidating data across systems and wants to be able to understand data associated with customers who might be in multiple systems:

"Matching Customers" Across Systems

CAAs need accurate and consistent identifying information in all systems: For example, name, date of birth, last 4 digits of social security, and in some cases address, phone number or email information.

Another way of accurately identifying people across systems is to ensure easy access to the internal IDs that identify a customer in all systems. Names are often misspelled, nicknames are entered, and date of births and SSNs are ripe with opportunities to reverse digits. It is also important to note that because many CAA customers have frequent address and phone number changes, these fields are not ideal for use as unique identifiers for customers.

Transforming and Mapping Data Options

In many cases, what a particular field means is slightly different across different data systems. For example, in some cases, "Housing Type" might be asking whether the house structure is a single-family home, and in other systems it might be asking whether the customer rents or own their home. Agencies need to make sure staff understand what fields like "Housing Type" mean in different systems and that those fields are 'mapped' to one another. Then, within each data point, there may be multiple options and those options in different systems may be equivalent or not. This is a second level of mapping that is necessary.

Transforming and Mapping Data Structures

In some cases, the data in one system is stored very differently than in another system, and in order to get to a single data point in one system, you may need to combine multiple pieces of information in specific circumstances to get to that single data point.

For example, in one data system it may be possible to record amounts of money by person and by type of income over time. In another system, it may only be possible to record a total household income and denote that it included different income sources. Fitting these two data sources together requires decisions and clear documentation to help communicate those decisions.

In another example, two systems may both record income sources and amounts by person, but one may do it literally – by type of income and date – while another may wrap all income sources and people into a snapshot of income at a particular time. In this case you may have comparable information, but it is recorded so differently that the transformations again require consideration and documentation.

Every piece of data across each data system – especially when there may or may not be guiding data standards – will increase the cost of these projects. Consequently, CAAs need to be clear about what data is needed or wanted and how it will be used.

CONCLUSIONS

It's a Journey and It's Worth Taking

The COVID-19 pandemic has disrupted lives and communities. It has also created an opportunity for CAAs to rapidly adopt technology in new ways. This opportunity beings with it an additional opportunity to re-think how CAAs collect, analyze and use data. Moving toward being data-informed is a journey that requires building data capacity and investment of resources, but can yield a stronger, more resilient organization better able to successfully navigate turbulent times, providing the information and capacity necessary to better serve its customers. But don't do data for data's sake - Do it for the customers.

Get the Right People Involved and Support Them

Organization leadership must understand the value of data and how it can benefit an agency and its customers, and be willing to prioritize the systems and resources necessary to appropriately collect, manage and use customer data.

There are a wide range of reasons that agency leadership decides to invest in building their data capacity including **both strategic and operational considerations**. There are a few common themes motivating those decisions, including: **new leadership**, questions raised **through strategic planning or assessments**, needing new support from **funders and stakeholders**, **crisis/disasters** and CSBG **reporting**.

In addition to leadership, CAAs will need **a staff**

person who can lead the charge on data collection, management and use. These staff will need to have the organizational standing and reach to work across the agency, the skills and knowledge to understand programmatic considerations, and the resolve to motivate successful data collection and use.

Develop Right Processes

Even with a supportive leadership and 'the right' person in the data position, it is critical to have **supporting infrastructure** within the organization. Think of this as the mechanisms through which expectations about data are communicated and analyzed, and where decisions using data happen. These include a well-defined leadership team that meets regularly and who examines agency data and regular meetings within programs and across programs.

Further examination of the data management pipeline or lifecycle reveals the importance of understanding and managing data through these processes, including identifying who is responsible at different stages and considering the question of data ownership (i.e., who is responsible for the integrity and accuracy of specific data). Agencies need to ensure that their systems, both technology systems and human systems, which manage these stages for different types of data are secure, consistent and understood.

Be Curious

Play with the information you have, wonder why information is showing what it is showing and ask what it means and what can you do with it. Having an open and inquisitive attitude toward the data the CAA collects can yield new insights for agency staff and the customers and communities it serves.



The Promise of Community Action

Community Action changes people's lives, embodies the spirit of hope, improves communities, and makes America a better place to live. We care about the entire community, and we are dedicated to helping people help themselves and each other.

ACKNOWLEDGMENTS: We are deeply indebted to the following individuals and agencies for their frank and lively discussions about using data in their agencies. Asterisks denote that the individual is a Data Ambassador with the National Community Action Partnership: **Ciara Alvarez** at Community Action Partnership of Lancaster and Saunders Counties (CAPLSC) in Nebraska; **Meaghan Arzberger** at York County Community Action Corporation (YCCAC) in Maine;* **Rand Clark** at Douglas County Department of Community Development, Community Services Division in Colorado; **Andy Agardy** at Utah Community Action;* **Amanda Shelton** at Gulf Coast Community Services Association (GCCSA) in Texas;* **Nadeem Siddiqi** at Wayne Metro Community Action Agency (Wayne Metro) in Michigan;* **Sue Kervin** at Action for a Better Community (ABC) in New York;* **David Hagen** at Hawkeye Area Community Action Program (HACAP) in Iowa; **Christine Shaver** at Tioga Opportunities in New York; **Kelli Smith** at People Inc. in Virginia.

This publication was created by the National Association of Community Action Agencies - Community Action Partnership, in the performance of the U.S. Department of Health and Human Services, Administration for Children and Families, Office of Community Services, Grant Number 90ET0469-01-01. Any opinion, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. Department of Health and Human Services, Administration for Children and Families.

CSBG
Community Services Block Grant