TANF DATA COLLABORATIVE

Strengthening Analytics in Government Agencies

A Toolkit for Sustainable Data Use

OPRE Report 2023-148 JULY 2023

Funders

The TANF Data Innovation (TDI) project is funded by the Office of Family Assistance (OFA) and the Office of Planning, Research, and Evaluation (OPRE) within the Administration for Children and Families (ACF). MDRC and subcontractors Actionable Intelligence for Social Policy at the University of Pennsylvania, Chapin Hall at the University of Chicago, and Coleridge Initiative are conducting the Promoting and Supporting Innovation in TANF Data project for the Department of Health and Human Services (HHS) under a contract with the Administration for Children and Families, funded by HHS under a competitive award, Contract No. HHSP233201500059I. The project officer is Nicole Deterding.

Dissemination of MDRC publications is supported by the following organizations and individuals that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: The Annie E. Casey Foundation, Arnold Ventures, Charles and Lynn Schusterman Family Foundation, The Edna McConnell Clark Foundation, Ford Foundation, The George Gund Foundation, Daniel and Corinne Goldman, The Harry and Jeanette Weinberg Foundation, Inc., The JPB Foundation, The Joyce Foundation, The Kresge Foundation, and Sandler Foundation.

In addition, earnings from the MDRC Endowment help sustain our dissemination efforts. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

Strengthening Analytics in Government Agencies

A Toolkit for Sustainable Data Use

OPRE Report 2023-148
JULY 2023

Authors:
Emily R. Wiegand, Emma K. Monahan, Rose Geoghegan, Melissa Wavelet, and Robert M. Goerge

Submitted to:
Nicole Deterding, Project Officer, Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Project Director: Richard Hendra, MDRC, 200 Vesey Street, New York, NY 10281.

Contract Number: HHSP233201500059I

Suggested citation: Wiegand, Emily R., Emma K. Monahan, Rose Geoghegan, Melissa Wavelet, and Robert M. Goerge. 2023. *Strengthening Analytics in Government Agencies: A Toolkit for Sustainable Data Use*. OPRE Report 2023-148. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

This report and other reports sponsored by the Office of Planning, Research, and Evaluation are available at www.acf.hhs.gov/opre.

Connect with OPRE

CHAPIN HALL AT THE UNIVERSITY OF CHICAGO

mdrc BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

COLERIDGE INITIATIVE

OPRE

ADMINISTRATION FOR CHILDREN & FAMILIES Office of Family Assistance

AISP

Contents

List of Exhibits

TABLE

FIGURE

BOX

Acknowledgments

We would like to express our most sincere appreciation to the individuals whose insights contributed to this toolkit's development, including Linda Armstrong, Cynthia Barlow, Amadneep Burns, Dave Dlugolecki, Jason Dunn, Ryan Gillette, Brittney Gossard, Jamie Haskel, Jessica Hinkle, Erin Kidder, Bill Koncelik, Kristopher Jones, Megan Kauffmann, Natalie Knudson, David Mancuso, Jess Maneely, Kelsey Martinez, Michael Martinez-Schiferl, Robert McGough, John McGraw, Lauren Molineaux, Mary Nelson, Britany Orlebeke, Letitia Logan Passarella, Heather Saco, Clarizza Singayao, Jeremias Solari, Kris Stevens, Todd Trapp, Jacqueline "Tina" Turner, Julianna Vignalats, Mischelle Williams, and Don Winstead.

We would also like to thank the TANF Data Collaborative pilot agency team members who provided valuable feedback on the content and tools in this publication.

We would also like to thank our partners, Actionable Intelligence for Social Policy, the Coleridge Initiative, and MDRC, who together with Chapin Hall formed the TANF Data Innovation team. Particular thanks are due to Richard Hendra, Erika Lundquist, Johanna Walter and Chris Homa of MDRC; Amy Hawn Nelson, Sharon Zanti, and Della Jenkins of AISP; Jasmine Jackson of Chapin Hall; and Nicole Deterding, Siri Warkentien, Bethanne Barnes, and Lauren Frohlich of the Administration for Children and Families in the U.S. Department of Health and Human Services for informing, reviewing, and supporting the production of this toolkit. Thanks to Jill Kirschenbaum and Carolyn Thomas for preparing it for publication.

INTRODUCTION 1

Using data to improve human services programs is becoming commonplace. Agency administrators, researchers, and policymakers increasingly turn to applied data analytics to evaluate programs or inform evidence-based policy. The applications of the data may vary but the goal is the same: to improve the lives of the families served.

These efforts often begin with a question, such as what aspects of a state Temporary Assistance for Needy Families (TANF) program help families gain employment, followed by a pilot initiative to address it. While completing a successful pilot analysis like this is no small feat, more recently, service agencies have been pursuing a broader goal: to move from data to knowledge to action not just once, but routinely, creating a reinforcing cycle of evidence-building and program improvement.

This toolkit was created to help agencies build the culture and infrastructure needed to apply data analysis routinely, effectively, and accurately—referred to in this publication as "sustainable data use." While qualitative and survey data have tremendous value for program evaluation and improvement, the scope of this toolkit focuses on the use of quantitative administrative data. The toolkit offers strategies and tools for individuals in government agencies and other organizations with similar needs. The material was sourced from interviews with practitioners who have successfully built sustainable data use into their everyday practices. It covers a variety of subjects—from staffing and technology to collaboration and funding—that can impact the longevity of analytics work in the public sector. While the toolkit was developed with state TANF agencies in mind, many of the techniques offered here may also be useful for individuals in a range of government agencies and other organizations, where the challenges—and potential for impact—are similar.

How to Use This Toolkit

Many readers will find it helpful to begin with Section 2, "Conditions for Sustainable Data Analytics," which describes how strategies and tools to foster sustainable data use are conceptually organized in this toolkit. Some sections include suggested resources and spotlights on agencies with noteworthy data use practices. Other than a short conclusion, the sections and tools in the

toolkit are entirely modular and can be read in any order. Table 1.1 provides a roadmap to help users find the sections and tools that are relevant to their specific needs.

Table 1.1 Tools in the Sustainability Toolkit

FOR HELP WITH:	START HERE:
Getting started thinking about sustainability	→ **Section 2:** Conditions for Sustainable Data Analytics
Identifying existing barriers to sustainability in your agency	→ **Tool 1:** Sustainable Data Use Self-Assessment
Encouraging support for (and use of) your agency's analytics	→ **Section 3:** Strategies to Foster Demand
Recruiting and hiring analytic staff	→ **Section 4:** Strategies to Build Staff Capacity **Tool 2:** Sample Language for Creating Job Descriptions for Data Analytics Staff **Tool 3:** Sample Interview Questions for Hiring Data Analytics Staff
Retaining and training analytic staff	→ **Section 4:** Strategies to Build Staff Capacity **Tool 4:** Identifying the Right Training Program for Your Needs
Building strong external partnerships (such as with universities) to augment analytic capacity	→ **Section 5:** Strategies to Build Capacity through External Partnerships
Fostering internal communication and collaboration around analytics	→ **Section 6:** Strategies to Improve Internal Communication
Improving documentation and management of institutional knowledge	→ **Section 7:** Strategies to Improve Knowledge Management and Documentation Tool 5: Documentation Checklist
Securing necessary funding for analytic efforts	→ **Section 8:** Strategies to Secure Funding

Why Sustainable Data Use Matters

> "The Commission envisions a future in which rigorous evidence is created efficiently, as a routine part of government operations, and used to construct effective public policy.
>
> —U.S. Commission on Evidence-Based Policymaking[1]

> "Episodic work with administrative data, whether due to inconsistent fiscal or executive leadership commitment, can be terribly inefficient. Data may not be available in a timely or readily accessible manner, which creates obstacles to regular use. Data quality and organizational knowledge about specific fields may lapse if enough time passes between analyses. Most importantly, intermittent use of administrative data to inform policy and practice rarely improves the experience of program clients or makes lasting impact on the work of an agency.
>
> —2019 summary of an assessment of data use to improve family self-sufficiency[2]

In 2017, the Administration for Children and Families within the U.S. Department of Health and Human Services launched the TANF Data Innovation (TDI) project. The goal of the TDI project was to substantially expand the routine use, integration, and analysis of TANF and employment data by agency staff to improve program services and outcomes for families with low incomes. The TDI project includes the TANF Data Collaborative (TDC), described in more detail in Box 1.1. TDI is being led by MDRC in partnership with Chapin Hall at the University of Chicago, Actionable Intelligence for Social Policy at the University of Pennsylvania, and the Coleridge Initiative. As part of that project, eight TANF agency teams participated in the TANF Data Collaborative Pilot Initiative, to further expand their data analytics capacity through the design and execution of an analytics project.[3] While all the pilot agency teams successfully completed their projects, both they and the TDI team recognized that they would face challenges making the transition from episodic data work to the creation of evidence as a routine part of government operations. This toolkit was designed to support those efforts.

How This Toolkit Was Developed

This toolkit was developed with information from two types of data collection: (1) interviews with TANF agency personnel and partners, and (2) the available literature on sustainable data use. From there, the research team assembled the

1. Commission on Evidence-Based Policymaking (2017).
2. Allard et al. (2018).
3. Wavelet, Rubino, Morris, and Garner (2022).

Box 1.1 TANF Data Innovation (TDI) Project

The TANF Data Innovation (TDI) project includes a needs assessment, efforts to support federal, state, and county staff use of federally reported TANF data, and the **TANF Data Collaborative (TDC)**.

TDC supports the use of administrative data to inform TANF policy and practice, with the ultimate goal of improving employment and well-being outcomes for TANF families.

TANF DATA COLLABORATIVE

TDC Activities

- Resources open to all TANF agencies for accessing and using participant employment and earnings data
- Coaching, training, and technical assistance for eight pilot agency teams on data-driven projects
- Opportunities for TANF agency staff members to meet and learn from each other
- Resources for using TANF data to support sustained capacity gains can be found at www.mdrc.org.

TDI Sponsors

The Office of Planning, Research, and Evaluation and the Office of Family Assistance in the Administration for Children and Families, U.S. Department of Health and Human Services

TDI Team

MDRC (project lead)

Chapin Hall at the University of Chicago

Actionable Intelligence for Social Policy (AISP) at the University of Pennsylvania

Coleridge Initiative

recommended strategies and tools offered here. Some are broad and conceptual; others are targeted and practical. Not every strategy will be right for every agency, but the goal is for the assembled material to offer something for virtually any challenge.

The research team spoke with representatives from ten states and five other organizations. These interviewees included people at TANF and other government agencies as well as staff members in organizations that partner closely with them. The respondents were recruited because their agencies

had taken analytics beyond isolated successes and had begun to weave data use into regular government operations; the state participants were always individuals directly involved in or leading data analytics efforts in their agencies. The research team identified agencies and organizations that were especially strong in at least one of five areas of sustainability: leadership support for analytics, staffing, communication, documentation, and external partnerships.[4] Each interview focused on one or two of the areas identified as particular strengths for that state.[5]

The team also reviewed the available literature (academic and professional) on change in the public sector generally and sustainable data analytics specifically.[6] Before drafting the toolkit, the team organized information across these sources by developing qualitative memos on the same five areas of sustainability that were used to structure the interviews.

4. These five areas were prioritized for sustainable data use based on findings from the TDI needs assessment. See Goerge, Wiegand, Monahan, and Gjertson (2022).

5. All case studies, state resources, and direct quotations have been approved for inclusion by the interviewees.

6. Sample search phrases included "public sector partnerships data analytics," "building data use in government," "building analytic capacity in government," and "data culture in government"; 193 articles and publications matched the initial search criteria, of which 79 were prioritized based on a review of abstracts.

2 CONDITIONS FOR SUSTAINABLE DATA ANALYTICS

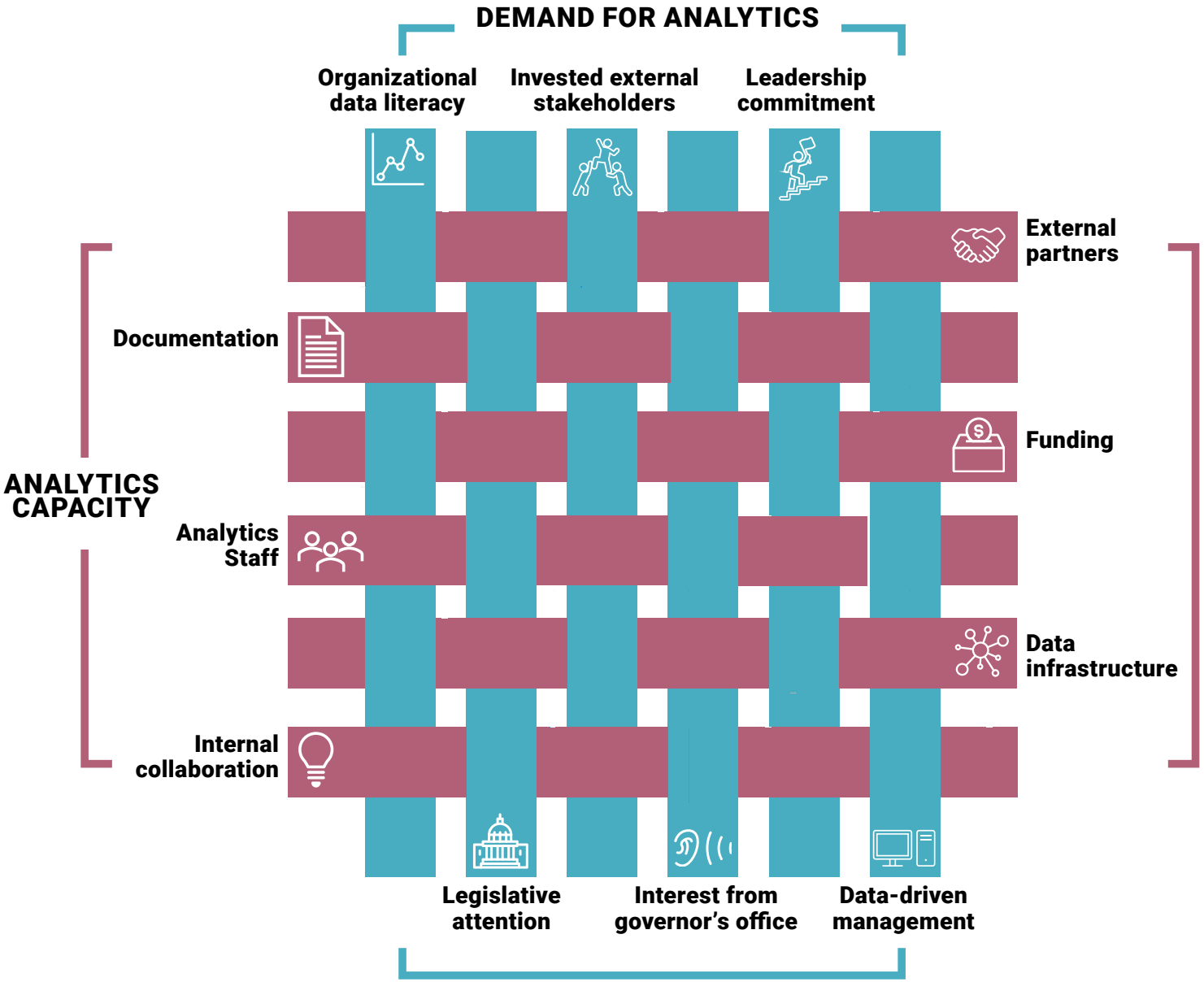
This toolkit presents two essential conditions for sustained data use. First, an agency must have a **demand** for data from decision-makers seeking data to inform their choices. And second, the agency needs the **capacity** to generate analyses—the people, data, and infrastructure necessary for data preparation and analysis.

As shown in Figure 2.1, sustainable data use is like a piece of cloth, with interwoven strands of demand and capacity. The presence and strength of the individual threads looks different in different agencies; there is no single actor who drives demand, and there are many ways to assemble the necessary capacity. Sometimes the demand for analyses is external, from invested parties such as legislators or advocates. Sometimes it comes from executive leadership. With time, many of the most analytically sophisticated organizations develop internal cultures of data use and data literacy, where staff members across levels and units thoughtfully seek and employ data to inform their decisions. On the capacity side, some organizations rely on internal staffing for data management and analysis, while others lean on external partnerships, such as those with universities or research organizations; still others use a combination of the two. Different components of data infrastructure—including data collection, data quality, access to integrated data, and data documentation—vary, and capacity is further reinforced by collaboration and coordination among everyone involved in analysis efforts.[1]

Demand without capacity leads to frustration and a lack of information; capacity without demand means the information does not reach the people who can apply it to improve programs. This toolkit includes more strategies for fostering capacity—not because demand is less important but because the research team heard more diverse responses about how to build capacity and more consistent responses about demand. One thing is certain: When both demand and capacity for data use are strong in an agency, the two are mutually reinforcing, encouraging more consistent practice and, ultimately, more evidence-based policymaking.

1. Goerge, Wiegand, Monahan, and Gjertson (2022).

Figure 2.1. Data Use is Strengthened by a Mix of Data Consumers and Capacities



Unsure of your organization's data strengths and weaknesses?

Consider turning to **Tool 1: Sustainable Data Use Self-Assessment,** on page 37. The assessment invites informal reflection on your agency's data use capacity and guides you to helpful resources in the toolkit based on where your organization excels and where it needs work. The assessment questions will also help you learn more about the components of sustainable data use and the types of practices addressed in this toolkit.

STRATEGIES TO FOSTER DEMAND

3

Demand for data use can come from state government or agency leadership, from agency staff, or from key external constituency groups. External parties, including advocates and legislators, may influence the demand for data in many ways, such as through lobbying and legislation. Data-driven leaders inside government agencies, regardless of level, often serve as catalysts for the growth of analytics. These leaders are important consumers of data, and they are also positioned to directly support analytics capacity building, such as through allocation of funding or through staff recruitment. Individual agency staff members may not exercise the same level of influence, but agencies that develop a strong culture of data use at all levels of seniority and across units are well-positioned to demonstrate the value of data. This is especially true when staff members who do not work directly in data analysis can point to how they have used data to improve operations and benefit clients.

Here are some strategies to foster demand and support for sustainable data use in your organization.

Identify and select a few champions of your team's use of data analytics and then prioritize their needs.

The people you most need to support your work should also understand how they get consistent value from data. Listen carefully to their goals and questions, and then show them how data can further those goals and address those questions. Not all agency leaders bring the same appreciation for using data to inform decisions; nonetheless, most have their own pressing questions. They might be asking a policy or programmatic question, such as "What do these families need?", without recognizing that data could help answer it. For example, by analyzing needs identified in child welfare case management by family characteristics, agencies could learn what is needed most frequently and by whom. This could lead to more successful targeting of resources and services.

An analytics team that shows how to identify and answer such questions with data is one that will earn support:

> "The way we think of our role is that we serve program administrators and policy makers. We also serve leadership and budget staff. Whether or not this kind of longitudinal data capacity is useful depends on their interest and ability to digest what we have to report to them.
>
> —Research Scientist, Bureau of Data Management and Analysis, New York State Office of Temporary and Disability Assistance

Don't overlook federal, legislative, or community audiences.

Don't assume that your champions are all within the state executive branch. Legislators, federal partners, and community members may also be consumers of analytics work, whether through mandates or by interest. They may also recognize the value of data analytics. Conferences and professional associations can present opportunities to reach external audiences with results of data analyses. Capitalizing on interest in analyses from leaders and lawmakers outside of your agency can help motivate growth and investment in data use inside your organization by helping internal leaders recognize the demand for data and analytics. By using analytics to generate external funding and interest, you can demonstrate to agency and state leadership the value of data and how it benefits the agency as a whole:

> "Federal agencies have really been pushing for outcomes, and that push has been a strong motivator for change—it leads to a lot more willingness to discuss data sharing.
>
> —Chief Analytics Officer, Arkansas Department of Information Systems

Even routine compliance activities such as federal reporting mandates and tracking metrics like work participation rates can be leveraged to build an analytics infrastructure that supports broader research efforts. After all, reporting requirements ensure resources for ongoing analytics. They also create an opportunity to engage in discussions about data—what it shows, what it means, and what additional analyses are needed to more fully answer questions, both internally and externally.

Encourage a culture of data use throughout your organization, not just at the top.

If demand for data runs throughout your agency, more individuals from multiple teams and at various levels will be invested in quality data collection, management, analysis, and communication.

Some states have used data literacy training to introduce the importance of data to staff members across an agency. Senior managers can support these efforts by creating expectations for data use at all levels:

Box 3.1
Building a Data Culture: Spotlight on California

When the California Department of Social Services (CDSS) launched its Research, Automation, and Data Division (RADD) in 2020, the aim was to build a sustainable culture of data use within the organization. But generating support for the division among program staff across departments as well as at lower levels of CDSS management was a challenge for agency leadership. In response, RADD spearheaded an initiative to promote engagement with analytics work and build data literacy throughout CDSS. The team found that workers embraced data use once they recognized it could be used to improve CDSS programs.

The initiative began by conducting focus groups with staff members across CDSS agencies. Staff members said they lacked access to data as well as confidence in their ability to analyze and interpret it. In response to the focus group findings, CDSS developed a data literacy learning program to empower their colleagues around data. Staff members with different skill levels and backgrounds were enlisted to help develop the course to ensure that it addressed the needs of the broadest possible group of CDSS employees. The department also launched a series for all managers, supervisors, and data leaders. The DataU Leader Series walked leaders through data literacy learning program components, discussed why data and information are essential to CDSS work, and used organizational change management strategies to support leaders in guiding staff and teams in the culture shift.

In its current design, the data literacy learning program begins with a self-assessment of data skills; course content varies based on the person's knowledge and data literacy goals. The program includes live interactions with data experts that are similar to coaching conversations but that point individuals to a library of training programs. "The structure is focus-group based, pointing people towards resources, getting them comfortable talking about data," a RADD leader explained. "We worked with Sacramento State and their continuing education department on the initial development and rollout of the effort last year. Now we're on our own.... We have two staff that are dedicated about 70 percent of the time to the data literacy initiative."

For more information on California's data literacy course, a description is available in the associated resources on GitHub.

> "Our director has made this conscientious effort to ask the leadership and management in the department to include data when doing updates—asking specific questions about data and what they're seeing in data.
>
> —Assistant Deputy Director, Research Automation and Data Division, California Department of Social Services

If your organization is embarking on a significant shift in data use, make sure all relevant people have a seat at the table.

Changing how data is used in an organization—for example, moving from episodic to sustained data use—requires change management, and communication is critical to these efforts. Allow time and space for everyone involved to voice their interests, concerns, and ideas. One way to do this is to host meetings that bring all relevant individuals to the table, even if that means slowing down the process or sharing control over parts of a change effort:

> "Everyone felt like they contributed to what the Data Sharing Agreements needed to look like; every legal team got to see the template, and we had as many meetings as people needed to have everything resolved. It was a huge tap on resources, but it was worth it.
>
> —Assistant Director of Employment Assistance, Arkansas Division of Workforce Services

Looking for more examples for building data literacy programs? Check out:

Lessons from Leading CDOs: A Framework for Better Civic Analytics
This 2017 report summarizes efforts by chief data officers in several cities and states to build "a culture of data" with workshops on data use for city employees and citizens.*

New Jersey's Manage by Data Program: Changing Culture and Capacity to Improve Outcomes
This 2015 report describes a data literacy initiative of the New Jersey Department of Children and Families and makes recommendations for agencies interested in building data culture.†

*Wiseman (2017).
†Lambert and Atkins (2015).

Be a trustworthy partner.

Human services work happens in a political context. Be mindful of concerns about negative media attention related to research findings. Internal and external end users are more willing to look at data once they know they "don't need to worry about headlines in the local paper that might make the department look poor," as one external partner noted in interviews with the research team. Establishing a reputation as a responsible steward of data—taking care to make sure analyses are accurate and are presented with appropriate context—is particularly crucial for sustaining the work through leadership and administration changes. Be prepared for the need to reestablish this trust when leadership roles turn over, new key partners emerge, the organization reorients its mission or goals, or there is significant staff turnover.

STRATEGIES TO BUILD STAFF CAPACITY

4

Adequate staffing is one of the first areas agencies focus on when building analytics capacity—and staffing is a persistent challenge. Finding, hiring, and retaining high-quality staff members are universal priorities for hiring managers, so many of the strategies the research team heard about in interviews reflect best practices and fundamentals in many industries. However, the combination of a competitive, evolving analytics job market and the bureaucratic stipulations of public sector hiring pose additional challenges. Organizations successful at hiring invest significant time and resources to ensure they excel at each step in the process.

Create clear, well-defined roles for both hiring and managing data analytics employees.

Thoughtfully defined and differentiated staff positions can help align expectations between managers and employees, both during and after the hiring process. For example, roles and job descriptions should clearly distinguish between an analyst who extracts and queries databases and a researcher who designs studies and conducts statistical modeling. Be specific about the discrete skills that are needed for each position and make sure the job description accurately matches what a staff member in that position is expected to do. Avoid looking for what one interviewee described as "Swiss army knife" candidates, who can perform a wide range of data management, analysis, and research functions, wherever possible. In a competitive job market, it is harder to hire individuals with wide-ranging skill sets at state government pay rates.

Thinking about diversity, equity, and inclusion in your workforce? Check out:

Public Sector Jobs: Opportunities for Advancing Racial Equity, a report from the Government Alliance on Race & Equity, includes strategies designed to advance equity goals in the public sector workforce.*

*Nelson and Tyrell (2015).

Accept that some jobs may have high turnover and design those roles accordingly.

Hiring managers in state government often struggle to compete with other organizations and sectors to hire research, analytics, and data staff. Entry-level data positions are often stepping stones for recent graduates and may consistently turn over after two or three years. Be strategic about designing these jobs to reduce the damage from that churn. Prioritize standardization, documentation, and automation for these roles so that tasks can be more easily transitioned to other staff and adopted by new hires. In designing professional development opportunities, concentrate on skills that can be applied in the near-term. Create pathways for promotion to new roles for individuals who stay for longer periods or who demonstrate the ability to thrive in state government. Although individual jobs may have high turnover rates, if the employee is promoted internally the organization can retain and benefit from its investment in that individual.

Make it easy for strong candidates to access your job listings and to picture themselves in the role.

Data and analytics job descriptions include specific language and requirements; make sure each role is accurately described and sounds desirable to the candidates you are trying to reach. Ask individuals in similar roles both inside and outside your organization to comment on the accuracy of

Need help creating job descriptions that will attract the best candidates?

Attracting, screening, and hiring the most qualified candidates begins with a job description that accurately describes the skills and experiences required for each position. For help in creating precise job descriptions, go to **Tool 2: Sample Language for Creating Job Descriptions for Data Analytics Staff**, on page 43 of this publication. It provides language sourced from interviewees and the TANF Data Innovation team that you can pull from as you craft job descriptions for various data analytics positions. You may also find it helpful to use this content in conjunction with **Tool 3: Sample Interview Questions for Hiring Data Analytics Staff**, on page 50, which provides interview questions for screening job candidates.

Still struggling with defining analytics roles?

Check out page 11 of Lessons from Leading CDOs: A Framework for Better Civic Analytics for a helpful table of government analytics roles.*

*Wiseman (2017).

the postings and required skills. Job candidates get an early sense for the hiring team from the technologies specified and what words are used; individuals trained in information technology, data science, and the social sciences often describe the same things very differently. In addition, if the level of technical skill required is clearly out of alignment with the responsibilities of the role, it suggests the hiring team lacks experience in the areas described—and may not have realistic expectations for new staff members.

This need for clarity is complicated by the language and presentation requirements for government job descriptions, which can be difficult for candidates to navigate. Include summaries in more colloquial terms when possible. The sample language for job descriptions included in **Tool 2** may help.

If you are hiring for multiple jobs that require similar skill sets, you might follow the strategy used by one state agency staff member interviewed by the research team: The agency posted related jobs simultaneously, allowing applicants to apply for several positions with a single application. Internally, this approach conserves resources and reduces the workload of hiring; externally, it reduces the need for the applicant to parse individual job postings to try to identify the one that fits them best.

Box 4.1
Getting Creative with Job Descriptions: Spotlight on New York State's Office of Temporary and Disability Assistance

Many state agencies must deal with dated job classifications that don't keep pace with the fast-changing skills and training required for many 21st-century jobs. In addition, exams are required as part of the application process for many government positions, which can significantly reduce the applicant pool. New York's Office of Temporary and Disability Assistance (OTDA) found a creative way to work with one existing classification. The title of "Research Scientist" was originally created for state agency workers doing basic research in fields such as medicine or environmental conservation, but it could also be applied to research and data analytics personnel in the human services. OTDA saw an opportunity: Research scientist positions are "open competitive," meaning no exam is required and any qualified candidate can apply. That meant the agency could recruit from the general public, significantly broadening the applicant pool and making it easier to reach individuals with in-demand technical skills. OTDA recently hired several strong candidates using this approach.

Seek out public service-minded, mission-driven employees by recruiting from universities and capitalizing on expressions of interest from the public, for example at job fairs or conferences.

Individuals who are passionate about the work done by human services agencies may be more amenable to public sector salaries and more likely to stay in their jobs longer. Recent graduates from social science, public health, and public policy programs are likely to be mission-oriented and may be a good fit for early career human services jobs in data analytics. University internship programs and capstone projects can provide immediate capacity as well as candidates to recruit for permanent roles. Giving career talks and attending job fairs at universities are other good ways to develop networks of future applicants. Some agencies use staff members who are also working as adjunct faculty at local universities to connect with current students:

> "We start building networks of people who are invested in the work and have technical skills and are tired of their current profession for whatever reason.
>
> —Data Scientist, Kentucky Center for Statistics (KYSTATS)

Work closely with human resources so the hiring process doesn't disqualify good candidates.

In some states, the human resources (HR) department prescreens job applications to ensure minimum requirements are met. If this is the case in your agency, cultivate a close working relationship with HR to prevent candidates from being disqualified due to miscommunication or a misunderstanding of the experience noted on their resumes. For example, HR personnel may not understand how information on an applicant's resume meets the minimum qualifications stated in the job posting. It can be helpful to talk with HR about your team's needs, to encourage flexibility, and to have a process in place to appeal failed applicants:

> "We're trying to work closely with HR on what changes they need to make and what kind of keywords to look at, and make sure we're partnering with them so we don't lose good candidates that come our way because of red tape.
>
> —Research Data Manager, Research Automation and Data Division, California Department of Social Services

Use written materials and exercises to assess job candidates' hard skills. Focus interviews on more general skills like problem-solving, teachability, and communication.

It can be tempting to concentrate solely on technical skills when recruiting for jobs in data analytics. But these positions in state agencies also require creativity, flexibility, and communication skills to make those technical skills effective for the agency. Use written materials, exercises, or tests to prescreen candidates for their technical skills. An effective screening tool should take a reasonable amount of time (45 minutes to one hour) and gauge the technical skills essential to the position. Save the interview time for evaluating soft skills and motivation. These qualities are crucial for analytics positions in state government but are often hard to teach:

> "We do have a simple test we give most of our analyst applicants—it has a 'tell us how you would approach this research question' kind of item, and you can gauge a lot about communication skills from the answer to that. We try to have interview questions that assess ability to communicate and their analytic thought process: how they would look at data to answer a question and then how they would communicate it.
>
> —Operations Manager, Employment and Benefits Division, Colorado Department of Human Services

Still need help with the hiring process? Take a look at these tools:

Interview questions are a natural extension of the job description, facilitating a deeper examination of a candidate's experiences, knowledge, and skills. For examples of interview questions, see **Tool 3: Sample Interview Questions for Hiring Data Analytics Staff**, on page 50. The sample questions in this tool, sourced from state government employees and the TANF Data Collaborative project team, can provide ideas for your interviews. This resource complements **Tool 2: Sample Language for Creating Job Descriptions for Data Analytics Staff**, on page 43, which offers sample language for data analytics-related job descriptions.

Emphasize growth opportunities and work culture.

A collaborative and supportive work culture promotes job satisfaction and retention by helping staff members identify their professional interests and develop their skills. While this is a best practice for all employers, it is especially relevant for agencies seeking to compete for skilled analytics and technical talent at public sector salaries. Several interviewees discussed the value of working in an agency where managers are focused on staff development:

> "It's making sure that people are seen and recognized and that we're going to give them opportunities to develop, grow their skills, and advance in what they're interested in.
>
> —Research Data Manager, Research Automation and Data Division, California Department of Social Service

> "When we're interviewing staff, we tell them that whether they're with us for a short period or long period of time, we'll be focused on growing and developing them and figuring out how to work with them to develop their skills in a way that aligns with their career path.
>
> —Operations Manager, Employment and Benefits Division, Colorado Department of Human Services

Provide analytics training opportunities to help existing staff members develop new skills.

Additional data analytics training opportunities can increase employee satisfaction and expand the analytics capacity of your team. Staff members with expertise in a programming language or a statistical technique can be tapped to train other members of the organization. Where this is not feasible, individuals can enroll in external training programs. Universities, private companies, and nonprofit organizations offer classes in popular programming languages and on topics in statistics, data analysis, and research. Many of these resources are available online, some with a registration fee and others free of charge. Keep in mind that training programs and skill-building workshops should include adequate time for participants to engage with the new material; this will ensure the largest pay-off for this type of investment.

Box 4.2
A "Great Place to Work": Spotlight on Colorado

As part of the 2020–2023 strategic plan for the Colorado Department of Human Services (CDHS), agency leadership spearheaded a "Great Place to Work" initiative focused on employee retention and satisfaction. The initiative included individualized professional development and career coaching activities, leadership training for senior staff members, diversity and inclusion programming, and wellness resources. As one staff member explained, "We're a culture now where we do everything we can to retain our employees. I ask in the first week, 'Where do you want to be one year from now? Five years from now? How do we grow you into that?' We align their workload to their interests, objectives, and goals to keep them motivated and keep them sane." These investments in employee satisfaction are paying off. In one case, an employee who was offered a higher salary elsewhere decided to remain with the agency because they valued the career development and professional support offered by CDHS.

Recognize how much staff members need to learn that is specific to your program, policies, and data.

Working with data and conducting research in a state policy environment can be highly complex. Analytics staff members need to learn details of program implementation and data collection, knowledge that must be developed on the job. You should expect that even with significant resources, training, and supports in place for new hires, they will need significant time to integrate fully into the department and be able to work independently:

> "We think it takes a couple of years for an analyst to really understand what they're doing—a prolonged period where the training wheels need to be on.
>
> —Research and Data Analysis Division Director, Washington State

Looking for training programs that will boost your team's skills?

There are many classes, both free and fee-based, that individuals can take to develop and advance their technical skills. But which classes are the best fit for your needs? **Tool 4: Identifying the Right Training Program for Your Needs**, on page 54, offers some important questions to ask to find the right programs for your staff and your organization.

STRATEGIES TO BUILD CAPACITY THROUGH EXTERNAL PARTNERSHIPS

5

While some agencies prefer to keep their analytics work in-house, many find that external partnerships can be a valuable complement to internal capacity. An external partner such as a university or a research organization can provide specialized skills for an ad hoc project. A longstanding partnership can also serve as a source of institutional memory and continuity through staffing or leadership changes at the agency. Finally, productive external partnerships often lead to publicly disseminated findings. This not only benefits the partnering agency but also contributes to the development of a rigorous evidence base for the field.[1]

Here are some strategies for creating and managing a successful data-sharing partnership.

Ensure any partnership is mutually beneficial.

A successful partnership will benefit both the state agency and the external partner. When entering into such a partnership, be specific about the expected benefits. Put them in writing and revisit them regularly. Are both parties getting tangible value from the relationship?

Mutual benefit is often clearest when the external partner is called on to apply skills that the agency lacks and does not plan to develop in its workforce. For example, a university partner might publish reports that answer questions the agency wants answered but doesn't have the analytic capacity to address. The partnership provides the university with a steady source of research material (potentially including access to the agency's data) and provides the agency with access to the capacity to answer its own policy and program research questions, such as the proportion of families who leave TANF and enter employment.

Sometimes agency staff members lack a skill that the leadership wants them to develop. Does the agency want the external partner to support the development of those skills? That expectation

1. Goerge, Wiegand, Monahan, and Gjertson (2022).

should be made clear. External partners may be happy to help train agency analysts if it means the analysts can continue to build on existing work. Alternately, some external partners may not find a partnership beneficial if there is an expected end point; a researcher considering investing time in learning an agency's data and policies may not want to do so if the partnership will be short-lived because the agency plans to do its own analytics work in the future.

Create scopes of work but consider leaving flexibility in the tasks and deliverables.

Create clear scopes of work outlining each entity's responsibilities and expectations for the partnership and ensure these are well understood. However, interviewees gave varying advice about just how specific contracts and work scopes should be. That is, *clear* is not the same as *rigid*.

One nonprofit organization that partners with multiple state TANF agencies recommended using formal (but not necessarily legally binding) cooperation agreements. In these agreements, the two entities formally identify the problem they are trying to solve by working together, define what roles and expertise each party brings to the table, and specify some measures of accountability, such as milestones to be accomplished by certain dates.

In other situations, flexibility is an important part of defining work with a partner. At a university with a longstanding, legislatively mandated partnership with a state agency, contracts are generally written on a three-to-four-year basis. Although these contracts do formally define a scope of work, they can be updated as needs arise:

Box 5.1
Streamlining Procurement: Spotlight on Colorado

In 2020, the Colorado Department of Human Services (CDHS) initiated a procurement process to create a list of approved vendors for research and evaluation. This allows the department to quickly identify and contract with the appropriate external partners for a particular data project. At least three vendors are selected initially; the state provides a statement of work and asks for a short, informal budget outline, then selects a partner based on the response. "[The vendors] all have parent agreements with the state—there's a procurement side to this—so we don't have to know as much or go through the same rigamarole to get a project started," explained a lead staff member. CDHS found this process has greatly increased efficiency in contracting with partners, recently allowing them to create four evaluation contracts in the span of four months—a remarkable achievement given the difficulty of moving quickly in state government.

The Colorado team provided its request for proposals as an example for other states. It is available in the associated resources posted on GitHub.

> "We try to be as flexible as possible by developing our contracts with to-be-determined projects and ad hoc clauses, so if something new pops up that is of great importance to our partners, we are in a position to say okay. Decisions to make larger changes to the contract are discussed informally in meetings, but we then need to formally modify the contract to cover the department in the event of legislative audits.
>
> —Director, Center Operations, Family Welfare Research, University of Maryland School of Social Work

There may also be times when it's preferable to keep contracts as flexible as possible to respond to changing analysis needs. This might be the case when two partners have a well-established and trusting relationship, funding is assured, or the scope of work is dependent on other work that has not yet been completed.

Make sure there is adequate communication between staff members with different areas of expertise on both sides of the partnership.

Communication can make or break a partnership's success. Agencies and partners must successfully explain their areas of expertise to ensure effective communication. Creating this level of communication could involve the agency hiring staff members who have experience in research and policy to serve in this translational role. Agencies could also hold informational meetings in the early stages of a partnership to provide basic education in each area of expertise, so everyone involved with the project has the same base knowledge. In general, it is important to learn and understand the other partner's context, speak in plain language, and explain constraints and interests that may not be apparent to the other party:

> "Sometimes it takes a translator to explain what the data can provide and help frame research questions. We respond to requests by clarifying what the state team really want to know, what's feasible based on availability/limitations of administrative data, and whether that information is meaningful.
>
> —Director, Center Operations, Family Welfare Research, University of Maryland School of Social Work

> "Without that translation layer that an internal evaluator is providing, I don't think external work would be as effective. Having someone with that internal program mindset who is close to program staff and understands the data.... The internal piece is really critical to having good external work.
>
> —Operations Manager, Employment and Benefits Division, Colorado Department of Human Services

Assign liaisons to manage relationships and streamline communication.

Advocate for assigning one staff person on each side of the partnership to act as a liaison. Assigned liaisons provide a single point of contact in every meeting and on every email, reducing communications challenges. They also ensure the individuals on each side are dedicated to managing the relationship, including understanding how work cultures and norms differ.[2]

It's also important to make sure there is a good fit between the agency staff members assigned to liaison positions and the partner organizations:

> "Staffing sometimes goes wrong when there's a mismatch between the liaison and the organization. You need to make sure that the liaison likes the organization.... Authentic relationships on the team create better work, so make sure there's a fit between liaison and organization.
>
> — Director of Training and Technical Assistance, Actionable Intelligence for Social Policy, University of Pennsylvania

Prioritize trust between the state agency and the external partner.

Build trust over time and actively work to maintain it. While trust matters on both sides of a partnership, the onus is often on the external organization to demonstrate that it will handle data securely, for example. External organizations also must listen carefully and understand what is important to invested parties at state agencies, operate with discretion to avoid creating negative attention, and communicate results to all interested parties before disseminating work:

> "Everything we do gets reviewed by the agencies we're partnering with. If we do a report that looks at high school to college trajectories for SNAP participants, we send that to the SNAP agency, also to the state Department of Education and the Council on Post-Secondary Education, and everyone gets some amount of time to review, check the numbers, and react. We don't want the first time that an agency leader hears about our SNAP numbers to be when someone asks him about a number that we published. It also helps us to learn the data.
>
> —Data Scientist, Kentucky Center for Statistics

2. For an analysis of how culture impacts the success of partnerships between government agencies and academics, see Gooden, Graham, and Martin (2014).

On the state agency side, it is important to be explicit with external partners about expectations in these areas and to communicate regularly around those expectations as the work unfolds.

Box 5.2
Building a Relationship: Spotlight on Maryland

For more than 20 years, the University of Maryland (UMD) has partnered with the state's Department of Human Services (DHS) to store, analyze, and report on the Maryland Temporary Assistance for Needy Families (TANF) program using administrative data. The partnership is supported by a legislative mandate requiring an annual report to the state legislature on TANF, which UMD produces. The university also performs more targeted analysis to help DHS understand where programs and services are successful and where there is room to improve.

Trust and mutual respect form the backbone of the partnership's success. In interviews for this toolkit, state program administrators expressed deep appreciation for their university colleagues as "historians of [the agency's] programs and systems." With long-term subject expertise dating back to the 1990s, the UMD researchers are able to advise agency staff on the historical and present-day impact of DHS policy. For their part, UMD staff emphasized the value of access to DHS data, and their obligation to return value to the state. "Data is valuable," one UMD partner said. "It's what every researcher is always looking for. So why should DHS provide UMD with administrative data? We need to provide them with something that makes the partnership valuable to them, so we provide as many services as possible to help them understand the data."

STRATEGIES TO IMPROVE INTERNAL COMMUNICATION

6

Internal communication may not seem as essential to analytics capacity as something like staffing or funding, but it is critical in making analytics work both more accurate and more impactful.[1] Analysts who use administrative data without a clear understanding of the context in which the data are collected or stored may misinterpret data points or misunderstand data elements. Regular, open communications among analysts and IT, policy, program, and budget staff members ensure the rigor and relevance of analytics projects. Communication between staff members who have similar roles but work with different social service programs or data sources promotes collaboration and improves institutional memory. It also allows staff to think about the programs and the people they serve more broadly.

A team built around fulfilling one requirement or generating isolated reports may have specialized skills and knowledge, but it is not well-positioned to have an impact on programs and services more broadly or share what it has learned. Breaking down silos—for example, pulling together two separate teams that get to know each other's work and data—helps organizations approach questions with a wider lens. The value of cross-functional teams was apparent during the TANF Data Collaborative (TDC) project, which encouraged the creation of teams that included data, policy, and program staff members. The intentional mixing of technical and nontechnical staff members yielded an exchange of keen insights that contributed to more accurate data analyses and clearer interpretations of results.[2]

Create positions or job responsibilities that are focused on communicating about data across teams.

Staff liaisons who are responsible for managing relationships and communications between teams can be as valuable inside an agency as they are

1. Goerge, Wiegand, Monahan, and Gjertson (2022).
2. An important component of the project was the TDC Pilot Initiative. The 30-month pilot offered technical assistance and training to support cross-disciplinary teams at eight state and county TANF programs in the routine use of administrative data to inform policy and practice. Pilot teams included a mix of agency leadership, research, data, and program staff. See Rubino et al. (2022) for more details.

between agencies and external partners. Internal liaisons are often responsible for streamlining communication between policy and data analytics staff members, typically by being very familiar with the issues and terminology of each team.

Consider centralizing analytics teams and data operations.

Representatives from some state agencies reported bringing together analytics staff from several programs, such as by putting TANF and SNAP analysts on the same team. This kind of structure can strengthen overall technical expertise by improving communication and facilitating knowledge sharing among technical staff members. It also allows for cross-training across programs. By breaking down program-specific siloes, these teams create environments where analyses across programs become more feasible.

However, such centralized analytics teams may run the risk of losing touch with changes in various programs and policies because the analytics staff members are no longer as tightly embedded within program teams. Specific strategies, such as creating working groups or assigning liaisons, are needed to maintain ongoing communication with program and policy staff:

Box 6.1
Communicating Across Teams: Spotlight on Utah

The Utah TANF program has a "gatekeeper" position—a person who organizes the process of requesting and transferring data from the agency. The gatekeeper ensures reports and data requests are prioritized correctly and that work is not duplicated. This means that data analysts can focus on writing queries and producing reports, while the gatekeeper negotiates between the analysts and policy staff members to accurately identify what is needed. The gatekeeper is also well versed in data systems and policy.

A job description for Utah's gatekeeper role is available in the associated resources on GitHub.

> "We are subject matter experts on our programs. We get in the weeds, find errors in data that stand out to us but to someone like [an analyst], they might not stand out. Say the number of adults on TANF seems really high. We might find out that the analyst is counting 18-year-old high school students as adults, so we have to filter those out.
>
> – Public Assistance Program Specialist, Division of Family Support, Kentucky Cabinet for Health and Family Services

Be intentional about incorporating feedback and collaboration in research and analysis projects.

Creating feedback loops during the analysis process is a great way to promote communication and collaboration across teams and staff roles. For example, in one state, the analytics and policy teams collectively set a research agenda that identified key questions and areas of analysis. The analysis process subsequently included a series of analytics-to-program checkpoints, including codesigning analyses with program staff members, holding a review meeting for initial results, and establishing a cross-team peer review of products prior to publication. An interviewee from another state described routine exchanges where research and data staff members turned to the program staff to provide context for their findings. Meanwhile, program staff members learned about how the data being collected were being used, as well as some of the limitations of those data. In a third state, program staff members in different agencies used instant messaging to quickly resolve questions about whether numbers made sense and to highlight previously invisible policy impacts.

STRATEGIES TO IMPROVE KNOWLEDGE MANAGEMENT AND DOCUMENTATION

7

Documentation is a central part of sustainable data use. Preserving knowledge, not only of datasets but also of code, analytics decisions, and policies and processes, is critical to expanding and applying evidence. Yet documentation is also the area where agencies regularly say they are struggling. When knowledge is fresh, documentation feels like an unnecessary luxury; as that knowledge fades, documentation becomes a burden.

In many ways, documentation is a component of communication, and the strategies for both are often complementary. When individuals are communicating effectively inside and across teams and organizations, they are constantly generating new knowledge, often in writing. The challenge of documentation is to create systems that effectively organize and store that information—and then get people to use them. Another challenge agencies face is thinking about everything that needs to be documented to ensure sustainability and preservation of institutional memory.

Does your documentation system support sustained data use?

For assistance in thinking about where your agency may need to focus on improving its documentation, see **Tool 5: Documentation Checklist** on page 58. An organization could have great documentation systems for processes, projects, or data, and still find itself lacking critical information in another area. In fact, strong documentation in one area can even lead a team to become complacent and overlook documentation in other important areas. Tool 5 details questions that highlight the breadth of knowledge that may be important for a state agency analytics team to track.

Here are some strategies for creating and maintaining effective documentation systems.

Share and centralize resources.

Start by making sure everyone knows which resource serves as the "single source of truth" for a given activity or set of logic used in or across

analyses. Information that is diligently documented but stored in scattered or hard to access locations and forgotten may as well not be documented at all. Analytics teams can use documentation resources to track the definitions of analytics and policy concepts, identify key variables, standardize data models, and archive past projects.

Interviewees described putting documentation in databases, documents, code—even in project management systems. The most important thing is that all users have access to the documentation, know where to find it, and understand when and how it should be updated. Regardless of location, clear naming conventions make any documentation system more searchable.

Establish documentation as a routine part of staff responsibilities and analyses.

Codify processes to document not only code and data dictionaries but also staff roles and responsibilities. One agency department includes documentation as part of its strategic plan, to ensure it is valued and prioritized across teams throughout the agency. Maintaining clear, written instructions for routine tasks (such as creating reports) keeps regular work and deliverables going in the event of staff turnover or the unexpected absence of a team member. Documenting tasks can also help people to better understand them.

One data analyst said that in a prior job she had to maintain a manual called "How to Do My Job." The manual included significant details about steps taken to create a certain dataset or product, meetings she was required to attend or coordinate, and a schedule for producing products. Similarly, for the TANF Data Collaborative Pilot Initiative, the technical support team created detailed instructional code notebooks for key data processing steps throughout their pilot analyses, so pilot agency teams could use them as on-demand resources later.

Use automation to support documentation.

In addition to improving efficiency, automation can support the documentation process. Automated tools can be used to profile data or conduct basic validations. For example, a script can review a dataset and identify likely data types based on the contents of each field. Profiler scripts can then run frequencies (for categorical fields) or distributions (for continuous fields) and document quality issues such as rates of missingness. These tools can quickly create basic reports on the contents of a dataset, though the resulting reports require human review to document definitions, translate coded data, and interpret validation results.

Automation can also support process documentation, where an automated report takes

the place of a manual process. An automated report has code that will persist and document the necessary logic even if the staff members involved in creating the report are no longer with the agency. However, to address questions that may arise in the future, it is important that report developers comment throughout the code to document not only what is being done but why.

Box 7.1
Documentation at Scale: Spotlight on Washington

Washington State's Department of Social and Health Services maintains one of the nation's largest integrated client data environments. The department's Research and Data Analysis Division (RDA) has several dozen staff members who are involved in the extraction of information from a broad set of administrative data systems (TANF-related systems among them). RDA is also responsible for the transformation of data into longitudinal, person-level data structures and the analysis of those data through a broad range of analytics methods (for example, quasi-experimental program evaluation methods).

Keeping track of "what things mean" in a constantly evolving program, policy, and data system environment requires extensive documentation. RDA has developed a customized application to track code sets, the logical association between code sets used to build higher-level conceptual definitions, and other metadata. Maintaining this application requires close coordination between program subject matter experts—who are often the first to become aware of program and system changes requiring documentation changes—and a technical team that both manages the analytics data infrastructure and updates the metadata infrastructure.

"In the old school way that I was trained as an analyst—more than 25 years ago—we would rely on our analytic programming code to document measurement concepts, even in some cases embedding smaller-scale value sets in code," the RDA director told the research team. "Now we organize reference information as data in a relational database, to facilitate transparency and ongoing maintenance of definitions, and so that staff don't need to 'find the code' to use the measurement concept. This infrastructure has parallels to what the National Committee for Quality Assurance maintains in its Healthcare Effectiveness Data and Information Set, for those who are familiar with the healthcare quality measurement domain."

STRATEGIES TO SECURE FUNDING

8

Securing sustained funding for data analytics may or may not be a challenge, depending on your agency's specific circumstances. But when continued funding is not secured, gaps can upset progress in all other areas. Here are some strategies for how to make sure funding is available and maintained over the long term.

Include funding for data capacity in new initiatives, pilots, or programs.

Routinely incorporating resources for data and analytics in the budget for new work creates reliable streams of future funding. The cost of conducting data analytics can be treated like other operational costs, such as human resources, finance, or facilities, and built into the operating budget. In this way, even if existing operations do not adequately pay for data analytics capacity, new initiatives can pick up the slack.

Explore federal or private funding if state funds are not available.

State government funding ebbs and flows. If your agency lacks resources for expanding data analytics work, be persistent about trying to secure funding through other available avenues. Monitor the federal agencies funding your program for formula or competitive grant opportunities. There may be someone in your state governor's office or in a state agency executive office that handles grant requests and can help identify philanthropies that might support your work. Individuals in one state agency interviewed by the research team have been highly successful at receiving funding from philanthropies and noted that in recent years, funding requests had been "going in the opposite direction," with philanthropic groups reaching out to the state to assess their needs.

Emphasize that effective investments in analytics can reduce the cost of working with data in the future.

Building data infrastructure and analytic capacity takes resources, but in time those investments can lead to increased efficiencies. Representatives from three state agencies with integrated data systems (administrative data linked across programs) noted the value of streamlining data-sharing agreements and processes. The effort was a significant

time-saver across units and agencies as well as a necessary first step toward building the data infrastructure needed for longitudinal analysis.

One state administrator noted that prior to the centralization of data, every agency was trying to do its own analytics work and and to maintain its own databases, and every project required multiple data-sharing agreements. Creating a centralized data team allowed agencies to pool resources and share the cost of servers, database management, and analysis, while also reducing the administrative burden of doing contract maintenance and applying for grants. Once data are integrated, documented, and ready for analysis, the cost of each additional evaluation or analysis is reduced.[1]

When funding for new titles or positions isn't available, repurpose existing titles and job duties.

Creative efforts to increase analytic capacity within existing personnel structures are a useful way to build and sustain data use. Some states have repurposed existing titles by modifying a job description to accommodate analytic duties, as described in the "Spotlight on New York" case study on page 16. Automating reports and using more efficient technology can also free up significant staff time—which can then be redirected to new data work. Organizational leaders who display

Box 8.1
Finding Creative Funding: Spotlight on Kentucky

In Kentucky, support for more robust data work was very strong among the leadership of the state Cabinet for Health and Family Services (CHFS), but the state legislature was not prepared to offer financial support. CHFS capitalized on the opportunity provided by a 2012 longitudinal data system grant from the U.S. Department of Education and partnered with the Kentucky Center for Statistics (KYSTATS) to link education and workforce data. Since then, CHFS and KYSTATS have built a thriving analytics partnership—one that was more than 90 percent grant-funded for nearly ten years before finally receiving a line in the state budget in 2022.

Although relying on grant funding can sometimes be a roadblock to sustainability, when support is not available from the state it's far better to begin the work using whatever funds are available than to wait indefinitely for more favorable circumstances. In Kentucky's case, the team's persistent pursuit of funding allowed it to build the capacity and data infrastructure to conduct the analytics work it valued so highly. Once that infrastructure was in place, the team was able to focus on producing results, which the legislature began to recognize, rely on, and finally, support financially.

1. See, for example, the case study on Washington State's integrated data system in Dube and White (2018).

little interest in analytics may support efforts to improve efficiency:

> "I don't know that we necessarily got a huge influx of new positions funding when we reorganized our research branch. We took a bunch of older positions that had been collected for concrete purposes in the past, that have been around a long time, and repurposed them. The staff in these roles will still be doing work with TANF, but modern technology is helping us use less staff time since we're able to automate things—we can redirect staff time to other analytic purposes now. We're trying to automate things, especially a lot of our reporting, so that something that used to take a staff member 40 hours a month now takes them three hours a month.
>
> —Research Data Manager, Research Automation and Data Division, California Department of Social Services

Want to learn more about data-sharing agreements and developing an integrated data system?

Actionable Intelligence for Social Policy has an array of relevant resources on its website. The following resources are a good place to start:

- Introduction to Data Sharing & Integration
- Finding a Way Forward: How to Create a Strong Legal Framework for Data Integration
- Building + Sustaining State Data Integration Efforts: Legislation, Funding, and Strategies

U.S. Department of Health and Human Services Office of Planning, Research and Evaluation's *Responsibly Sharing Confidential Data: Tools and Recommendations* project has related resources on its website.

CONCLUSION AND A NOTE OF ENCOURAGEMENT

9

We hope this toolkit helps you build your agency's sustainable data analytics capacity. The number and diversity of strategies offered here illustrate that state agencies that have sustainably incorporated data analytics have taken many different pathways and approaches to that success. While there are many choices to make as you move forward with a plan for your agency, keep in mind the two essential conditions of sustained data use: Which, if any, decision-makers **demand** data to inform their choices, and what is the status of your agency's **capacity** to generate data analyses? Next, take stock of the legal, technological, political, and fiscal climate in which your agency operates as well as the current state executive leadership. Use this information to have conversations about data analysis with your key constituencies early and often and prioritize strategies that can successfully generate quick or easy wins. The important thing is to start somewhere, knowing that making data use a routine part of operations will lead to more evidence-based program and policy decisions that support improved outcomes for children and families.

And remember: Although this toolkit covers many areas and strategies that may seem overwhelming at first, they come from state agencies that have implemented them. It is possible for your organization, too! The operating climate may change, making the design or implementation of certain strategies unpredictable at times. But by prioritizing the ones that are most likely to be successful at *your* agency, you can begin to build capacity for data analytics that will enable your agency to consistently and routinely create evidence-based policies and programs.

TOOLBOX

Tool 1 A Sustainable Data Use Self-Assessment

Tool 2 Sample Language for Creating Job Descriptions for Data Analytics Staff

Tool 3 Sample Interview Questions for Hiring Data Analytics Staff

Tool 4 Identifying the Right Training Program for Your Needs

Tool 5 A Documentation Checklist

TOOL 1

SUSTAINABLE DATA USE SELF-ASSESSMENT

Unsure of your organization's strengths and weaknesses when it comes to data use? This assessment provides a chance to informally reflect on your agency's data use capacity and guides you to helpful resources in the toolkit based on where your organization excels and where it needs work. The assessment questions can also help you learn more about the components of sustainable data use and the types of practices addressed in this toolkit. The assessment was developed based on interviews for this toolkit as well as the research team's prior direct experience with state TANF agencies. It was refined with feedback from members of the TDC Pilot Initiative team.

INSTRUCTIONS

Each of the six sections of the assessment is worth a certain number of points. Based on your score, there is guidance below each section suggesting which areas of the toolkit might be particularly useful for your agency. This self-assessment is not meant to be a rigorous analysis of your organization's data practice. Many of the questions are subjective, and different people might rate the same organization very differently. The goal of the assessment is to help you determine the areas you believe to be stronger and where you think extra attention would be most valuable.

Demand for Data (44 possible points)

1. Is evidence-based policymaking prioritized by leadership in your agency?

   ☐ Not at all (0) ☐ Somewhat (2) ☐ Very much (4)

2. Is the importance and value of using data for evidence communicated from the top down in your agency?

   ☐ Not at all (0) ☐ Somewhat (2) ☐ Very much (4)

3. Who uses data analyses or reports related to TANF? Choose all that apply (2 points per user, up to 14 possible points).

   ☐ Legislature (2)

   ☐ Advocates (2)

☐ Governor's office (2)

☐ TANF management (2)

☐ Other human service administrators (2)

☐ Local government officials (2)

☐ Other (2)

4. Who requests data analyses or reports related to TANF that are outside of regular reporting? Choose all that apply (2 points per user, up to 14 possible points).

   ☐ Legislature (2)

   ☐ Advocates (2)

   ☐ Governor's office (2)

   ☐ TANF management (2)

   ☐ Other human service administrators (2)

   ☐ Local government officials (2)

   ☐ Other (2)

5. How often do data analyses related to TANF directly inform and impact changes to policy and programs?

   ☐ Rarely or Never (0) ☐ Sometimes (2) ☐ Often (4)

6. Do the majority of employees in all roles and at all levels value using data for program improvement?

   ☐ No (0) ☐ Yes (2)

7. Do the majority of employees in all roles and at all levels understand how to use data and analytics for program improvement?

   ☐ No (0) ☐ Yes (2)

If you rate your organization a 20 or less in this area, some of the resources and strategies in Section 3, "Strategies to Foster Demand," may be a good fit for you (page 9). ***SCORE*****: 0**

Staffing (12 possible points)

8. Do you have an adequate number of staff members to complete all the data work needed by your agency?

   ☐ No (0) ☐ Yes (2)

9. Does your staff have the skills needed to manipulate and analyze data (for example, to answer an ad hoc program question based on data)?

   ☐ No (0) ☐ Yes (2)

10. Are you able to recruit and hire individuals who are a good fit for your data analytics needs?

    ☐ No (0) ☐ Yes (2)

11. How would you rate the opportunities in your agency for training your analytics staff and providing professional development opportunities?

    ☐ Poor (0) ☐ Fair (2) ☐ Very good (4)

12. Do you have a hard time retaining staff members with data analytics skills?

    ☐ Yes (0) ☐ No (2)

If you rate your organization a 6 or less in this area, some of the resources and strategies in Section 4, "Strategies to Build Staff Capacity," may be a good fit for you (page 14). ***SCORE*****: 0**

External Partnerships (8 possible points)

13. Has your agency used research conducted externally to inform planning or operations?

    ☐ No (0) ☐ Yes (1)

14. Is your internal capacity for data analysis strong enough that you do not need to regularly work with external partners to complete analytics projects, including publicly disseminating results?

    ☐ No (0) ☐ Somewhat (1) ☐ Yes (2)

15. Have you been able to find compatible external partners to collaborate with on research, evaluation, or data analysis activities?

    ☐ No (0) ☐ Yes (1)

16. How would you rate your communication with external partners, past or present?

    ☐ Not applicable/no external partners (0) ☐ Poor (0) ☐ Fair (1) ☐ Very good (2)

17. Do you establish clear expectations and roles with external partners?

    ☐ Not applicable/no external partners (0) ☐ No (0) ☐ Yes (1)

18. Are you and your external partners able to share and access the data you need to answer research and policy questions about the TANF program?

    ☐ Not applicable/no external partners (0) ☐ No (0) ☐ Yes (1)

If you rate your organization a 4 or less in this area, some of the resources and strategies in Section 5, "Strategies to Build Capacity Through External Partnerships," may be a good fit for you (page 21). ***SCORE*****: 0**

Communication and Collaboration (9 possible points)

19. How strongly does your agency value communication and collaboration across diverse roles, teams, and agencies?

    ☐ Not strongly at all (0) ☐ Somewhat strongly (1) ☐ Very strongly (2)

20. How often do program or policy staff members and data staff members within your agency communicate with each other?

    ☐ Never (0) ☐ Sometimes (1) ☐ Often (2)

21. How often do data staff members at your agency use informal channels other than email (such as Microsoft Teams, Slack, or in-person or telephone conversations outside of meetings) to communicate with others outside their teams, such as with program staff members or external partners?

    ☐ Never (0) ☐ Sometimes (1) ☐ Often (2)

22. How often does your agency communicate and collaborate with other agencies about analytics projects or data?

☐ Never (0) ☐ Sometimes (1) ☐ Often (2)

23. Does your agency promote a culture of sharing data work and projects with other teams and staff for feedback and additional perspective?

☐ No (0) ☐ Yes (1)

If you rate your organization a 5 or less in this area, some of the resources and strategies in Section 6, "Strategies to Improve Internal Communication," may be a good fit for you (page 26).
***SCORE*: 0**

Documentation (8 points possible)

24. If a key data, analytics, or research staff member left on short notice, how difficult would it be to pick up their work where it was left off?

☐ Very difficult (0) ☐ Somewhat difficult (1) ☐ Not at all difficult (2)

25. Does your agency have a standard practice for documenting staff roles and responsibilities for data, analytics, or research staff?

☐ Not at all (0) ☐ Somewhat (1) ☐ Definitely (2)

26. Does your agency have a standard practice for documenting data work and analytics projects?

☐ Not at all (0) ☐ Somewhat (1) ☐ Definitely (2)

27. How would you rate the level of data documentation (such as information about data types, field definitions, code values, or the limitations of certain fields) available to data users at your agency?

☐ Poor (0) ☐ Fair (1) ☐ Very good (2)

If you rate your organization a 4 or less in this area, some of the resources and strategies in Section 7, "Strategies to Improve Knowledge Management and Documentation," may be a good fit for you (page 29). ***SCORE*: 0**

Securing funding (8 possible points)

28. How adequate are the financial resources available for data analytics within your team/agency?

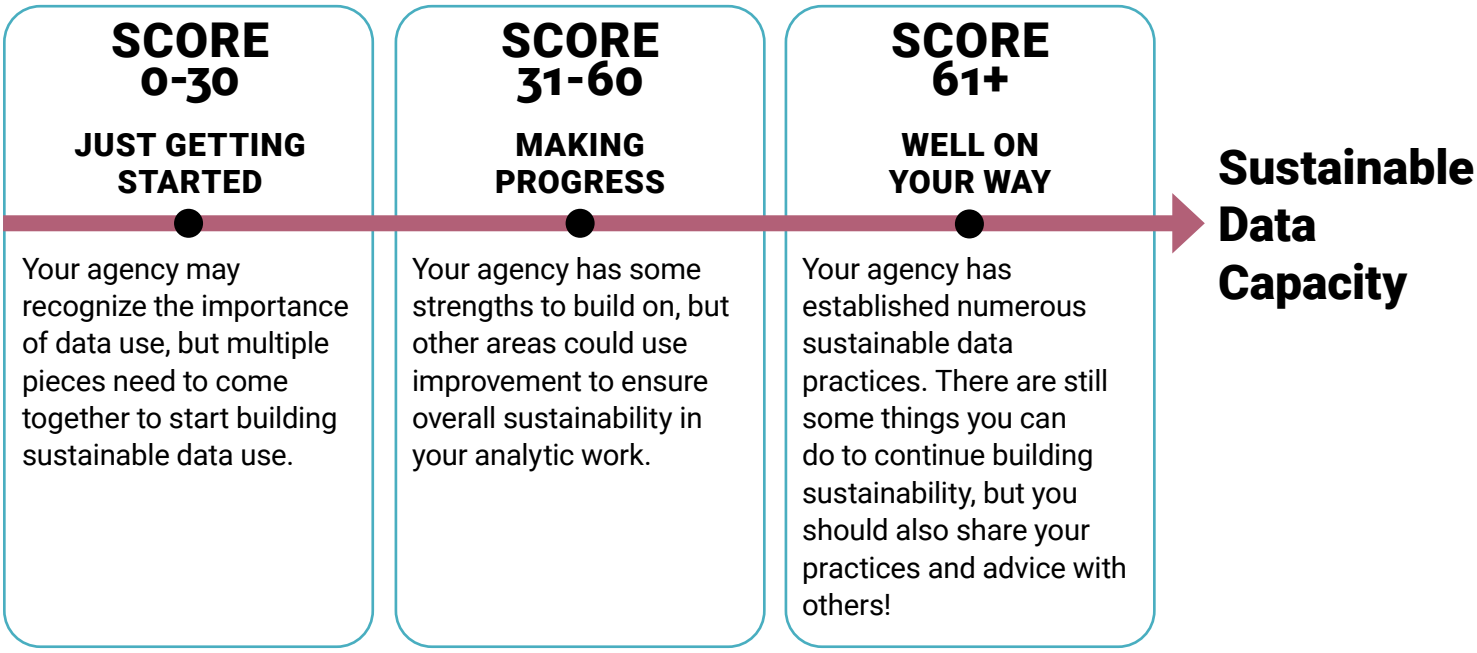
☐ Not at all adequate (0) ☐ Somewhat adequate (2) ☐ Completely adequate (4)

29. How much experience does your team or agency have in securing funding from outside organizations (for example, federal agencies, private foundations)?

☐ None (0) ☐ Some (2) ☐ A lot (4)

If you rate your organization a 4 or less in this area, some of the resources and strategies in Section 8, "Strategies to Secure Funding," may be a good fit for you (page 32). ***SCORE*****: 0**

Now add up your total score to see where your agency is on the path to sustainable data use.



TOOL 2

SAMPLE LANGUAGE FOR CREATING JOB DESCRIPTIONS FOR DATA ANALYTICS STAFF

Attracting, screening, and hiring the most qualified candidates begins with a job description that accurately captures the skills and experiences required for the position. This tool provides sample language, sourced from interviews for this publication as well as from the TANF Data Collaborative project team, that you can use as a master list to pull from as you craft job descriptions for data analytics staff. You may also find it helpful to use this content in conjunction with **Tool 3: Sample Interview Questions for Data Analytics Staff**, which offers interview questions for screening data analytics job candidates.

INSTRUCTIONS

This tool is divided into five categories of information that often appear in job descriptions. In each category, the tool provides sample language in regular text and additional word choice options or prompts in bracketed italics to help you further customize the job description.

Remember that you don't have to include everything listed here. Any single job description should include a tailored version of the sample language offered that is most directly relevant to the position you are filling. Be clear about what the role requires versus what is simply "nice to have" and consider which skills could be supplemented with training after hiring if necessary.

The organizational context

Emphasize positive characteristics of the job, such as work culture, professional development opportunities, and organizational mission to attract quality applicants. Use this space to brainstorm the strengths of your agency in each of these areas. You can then use these ideas to build an effective job description.

WORK CULTURE

ORGANIZATIONAL MISSION

OTHER POSITIVE CHARACTERISTICS

PROFESSIONAL DEVELOPMENT OPPORTUNITIES

The team

Describe the higher-level outcomes or objectives of the unit or department using language such as:

This role contributes to:

- Our proven track record of bridging the gap between academic researchers and government policymakers
- Facilitating data-driven or better decision making and data-informed policymaking and planning
- Gaining insight into problems
- Growing our organization's data analytics capacity
- Improving current processes and finding efficiencies
- Making our organization more responsive, experimental, and informed by evidence
- Transitioning our organization from legacy systems to updated technology
- Providing tailored analytics solutions to address complex management and policy questions and enable improved outcomes for [*insert population*]
- Creating in-depth analyses to convey accurate, meaningful insights that empower [*insert audience*]
- Sustaining our relationships with [*community partners, legislators, executive leaders, higher education institutions*]
- Making our organization more effective and equitable for [*residents, families*]

The role

Describe the overall experience of the individual in this position.

This role is well-suited for people who enjoy:

- Tackling complex issues with an analytical approach
- Understanding the nuances of a problem

- Exploring the role that data can play in [*the public sector, cities, states*]
- Using data science to find innovative solutions to the problems facing [*insert population*]

Competencies

The role requires the ability to:

REQUIRED TECHNICAL COMPETENCIES

- Conduct quantitative/statistical analysis including [*descriptive analyses, statistical modeling, machine learning*]. For roles where particular statistical models or machine learning methods are especially important, it may make sense to specify those methods, such as:
  - Multivariate linear regression, logistic regression, other regression techniques [*Ridge, LASSO, Bayesian*]
  - Mixed modeling techniques
  - Survival analysis, proportional hazards models
  - Methods for modeling non-linear data (for example, generalized additive models)
  - Predictive analytics
  - Sequence analysis
  - Clustering or principal component analysis
  - Random forests
  - Artificial neural networks
- Plan and design [*experimental, quasi-experimental, mixed methods*] studies to [*assess effectiveness of program services, assess client behavioral responses to program services, assess agency messaging and outreach efforts*]
- Evaluate and synthesize external research on [*analytics methodologies relevant to human services programs, evidence-based practices*]
- Manipulate data, including cleaning, standardizing, and transforming raw data

- Develop, manage and use [*large, complex, cross-sectional, longitudinal, relational, administrative] datasets* [*in the public sector, in human services*]
- Manage, mine and analyze [*public assistance records, administrative datasets, longitudinal administrative data*]
- Assess the structure, content, and quality of data
- Maintain [*impeccable, accessible, comprehensive*] documentation
- Exercise judgment and discretion in the handling of sensitive and confidential data
- Program in [*R, Python, SQL, SAS*]
- Use [*business intelligence tools, Tableau, Power BI, SAP Business Objects*]
- Develop maps/conduct spatial analyses in [*ArcGIS, QGIS, other GIS tools*]
- Use record linkage methods to link disparate data sources
- Demonstrate familiarity with other business software including [*Adobe Acrobat, Microsoft Office suite, specific relational database applications such as Microsoft SQL Server or PostgreSQL, web-based project management tools such as Trello or Jira, file sharing platforms such as Box or Google Drive*]
- Use version control tools such as Git and GitHub
- Design, create, and deploy complex and automated data integrations and pipelines
- Support the development of a data warehouse in [*Azure, Amazon WorkSpaces*]

REQUIRED PROFESSIONAL COMPETENCIES

- Collect and interpret data
- Lead or contribute to preparation of reports and communication of findings to various internal and external audiences
- Represent the [*organization, department, division*] to [*community members, legislators, organizational management*] and make presentations of data and findings
- Contribute to shaping data governance policies and practices
- Understand technical challenges that lead to data errors or inconsistencies and make recommendations on a set of standards to improve accuracy of data collection

- Identify, use, and promote best practices for software and data quality control
- Develop standards and best practice procedures for [*data validation, documentation, analysis, presentation*]
- Translate program staff needs into technical requirements to create [*rapid prototypes, minimally viable products*]
- Effectively communicate, both verbally and in writing
- Form productive relationships with [*community members, constituents, funders*]
- Produce written products on time and on budget that rely on data for [*public dissemination, policy decisions, budget decisions*]
- Translate research and analytics findings into actionable evidence and implications for policy and programs
- Develop and deliver presentations of findings to diverse audiences

WORK STYLE

- Work with [*low, medium, high*] degree of independence and autonomy
- Work collaboratively in a [*insert number of staff, budget, or other characterization of team size*] team to accomplish goals on time and on budget
- Effectively manage your own time while balancing competing demands and multiple projects
- Work in a [*fast-paced, unpredictable, interdisciplinary*] environment

SUPERVISION

- Supervise [*undergraduate and graduate student employees, other research staff*]
- Provide mentoring and guidance to [*junior staff, other staff*]

Qualifications

What education and work experience will the candidate need for success? Consider including both required and preferred education or experience. For education, include both degree type and field. For experience, describe both the duration desired and the field or type of work.

- [*Bachelor's, master's*] degree in [*public policy, mathematics, computational social science, economics, computer science, information technology, statistics, or related discipline*]
- [*Insert number of years*] experience working in [*interdisciplinary, government, academic*] environment
- [*Insert number of years*] experience in or familiarity with [*TANF policy and programs, public assistance programs, employment, and training programs*]
- [*Insert number of years*] experience with [*large, complex, cross-sectional, longitudinal, relational, administrative*] datasets [*in the social sciences, in the public sector, in the human services*]
- Other things to note:
  - Who person will report to
  - Details of work environment [*remote, in-person, or hybrid*]
  - Work location
  - Travel expectations [*average number of days, frequency, or percent of time*]

TOOL 3

SAMPLE INTERVIEW QUESTIONS FOR HIRING DATA ANALYTICS STAFF

Interview questions are a natural extension of the job description, facilitating a deeper examination of a candidate's experiences, knowledge, and skills. The sample questions in this tool, sourced from interviews with state government employees and the TANF Data Collaborative project team, can be used as a master list to pull from as you draft interview questions. This resource complements **Tool 2**, which offers sample language for creating data analytics-related job descriptions.

Note that these sample questions address interests and soft skills as well as technical competencies. The individuals interviewed for this toolkit emphasized the value of using interview time to evaluate a candidate's nontechnical qualities and skills.

Finally, don't forget that the interview is also your chance to market the position to potential candidates and ensure they have the right information to understand if the job is a good fit. Make sure that the interview includes some time spent describing the team and the organization in more detail, as well as adequate time for the candidate to ask questions.

INSTRUCTIONS

This tool is divided into four categories of interview questions. Each one contains sample language in regular text and additional word choice options or prompts in bracketed italics to help you further customize the question. These questions will help you carefully assess candidates' qualifications and fitness for the role.

The position

WHY ARE YOU INTERESTED IN THIS JOB?

- What attracts you to this job? This organization, unit, department?
- Why *now?* How does this job fit into your career trajectory?
- What do you want from your next job? Are there things you want to learn or experiences you want to have?

- Do you have medium- or long-term career goals you would like to share with me? How do you see this job fitting into those goals?
- What do you most and least enjoy doing in your current job?
- What matters to you most about your current job or your current employer?
- What concerns you about this job? This organization, unit, department?

Nontechnical skills

WHAT DO YOU KNOW AND WHAT HAVE YOU DONE?

Tell me about a time when you:

- Had to explain technical results to a nontechnical user or audience
- Created a product [*brief, journal article, or PowerPoint*] to disseminate your work
- Learned something new and incorporated it into your skillset
- Learned something new from your colleagues or peers
- Taught or explained something new to a colleague
- Worked as part of a team or a group [*professionally or in school*]
- Managed multiple projects with similar deadlines. What tools and strategies did you use?
- Felt overwhelmed at work or school or had to juggle a lot of competing priorities. How did you handle it?
- Had to do something that you didn't like for work or school. What was it, why didn't you like it, and how did you handle the situation?
- *[For supervisory positions*]: Took an active role in promoting a supervisee's professional development.

Analytics skills

WHAT DO YOU KNOW AND WHAT HAVE YOU DONE?

Tell me about a time when you:

- Were assigned a project based on a data source or domain with which you were unfamiliar. Describe your method of analyzing a problem and providing a solution involving an area where you have limited or no prior knowledge.
- Prepared data for analytics modeling. What were the steps involved, and what are some of the reasons you had for transforming or excluding certain variables?
- Used a complex or challenging data source. What made it challenging? How did you handle those challenges?
- Worked with a longitudinal data resource or conducted research that studied experiences over time.
- Had data that didn't match your expectations. What did you do?
- Learned and applied a new data science model, technique, or analytics technology to solve a business problem.
- Evaluated model performance and what to watch out for when evaluating model performance.
- Used version control and/or code review in managing your analytics work.
- Describe your experience with [*specify applicable methods such as data transformation, data management, administrative data, analysis or modeling, evaluating model performance, research design, program evaluation, data documentation, dashboard design and development, record linkage*].

See **Tool 2** for more ideas about potential methods and technical skills that may be relevant to ask about. Note that it can be helpful to distinguish between experience gained in a work environment versus in a class and to ask about the level of independent work the candidate has done in the topical area.

Software experience

WHAT DO YOU KNOW AND WHAT HAVE YOU DONE?

Given the variety of software and languages used in the field, interviewers should ask candidates about both specific tools and similar tools they have used (for example, other statistical programming languages, business intelligence tools, relational database systems), since knowledge and experience generally translate across similar platforms. See **Tool 2** for ideas about software and tools that may be relevant.

- Describe your experience with [*insert language or application*].

For particularly relevant languages and applications:

Tell me a time when you:

- Used [*insert tool*] to perform complex analytics tasks
- Had to fix a bug in the code or troubleshoot an error
- Found a faster or more efficient approach to existing work. How did you proceed?

Tell me about:

- Your favorite or go-to functions or libraries
- The commands you can't live without or find yourself often showing others

TOOL 4

IDENTIFYING THE RIGHT TRAINING PROGRAM FOR YOUR NEEDS

There are many available classes, both free and fee-based, that your staff members can take to develop their skills in data analytics. These include short interactive courses in popular programming languages, statistical seminars, and online lecture series at major universities. But how do you choose the right training to help your team members advance and further the goals of your organization? This tool is designed to help you do just that. It offers questions to ask to help you identify your staff members' goals, your agency's capacity and needs, and which programs are a good fit for your team.

INSTRUCTIONS

Use the questions below as a guide for finding the right training opportunities for your staff members' professional development and to enhance their analytics skills. The questions consider both employee interests and agency needs. Each question is followed by related considerations in italics to help you develop your answer, as well as space to use for notes. And remember: It is always useful to ask for recommendations for training programs from staff members or external partners who already have strong skills or significant experience working with data.

Understanding your staff

WHAT ARE THEIR INTERESTS AND NEEDS?

What analytics skills are your staff members interested in developing? Which ones would be most valuable in their current jobs?

> *No two analytics courses are alike. Some focus on statistical methods, others focus on understanding programming concepts and syntax, and still others focus on data visualization and communicating results. Consider classes on programming languages or tools, statistical methods, technologies like relational databases, and contextual topics like human-centered design.*

What will you need to do to make sure your staff members can effectively apply the new skills in their current jobs?

Staff members who learn a new programming language but whose work environment (such as data access or the existing technology) does not allow them to work in that language immediately will struggle to apply and retain what they have learned. An individual may learn a new statistical method, for example, but not have the additional skills needed to assemble the necessary data. Think in advance about organizational barriers or skills gaps that could keep the individual from applying the new skills. Consider planning the training in stages so that each new skill builds on the previous one learned (for example, training on analysis after training on querying databases and building data files). Look for internal opportunities that align with the training so the individual can apply and build on that new skill on the job.

How much time do your staff members have available to apply what they learned and strengthen their analytics skills in their jobs?

Consider not only the time needed for staff members to complete their training but also the time needed to incorporate those new skills into their current responsibilities. Individuals who learn skills in smaller pieces, with time to practice and apply each new piece, will be easier to train and support than those who learn a large amount of new material all at once. Supervisors play an important role in working with staff members to adjust their workloads so they can use what they learn.

Ask each staff member: What is your learning style? Do you prefer to learn through written materials, a lecture, or an interactive workshop? Do you want to work through something in a group, or do you prefer being left alone to explore the material?

Training opportunities come in many formats (video, written, classroom) and many modes (in-person, virtual, on-demand). They can range from semester-long classes to boot camps to webinars, and everything in between. Questions of cost and content will drive some of the selection process, but individuals may gravitate toward certain formats based on their individual circumstances or how they prefer to learn.

Assessing your organization

WHAT IS THE CURRENT DATA USE STATE OF PLAY?

What skills does your agency want employees to develop?

What are the skills gaps for the team and the agency as a whole? It may be possible to shift responsibilities among staff members to create opportunities for others to apply new skills they have learned.

What is the existing capacity for data analytics training within the organization? Are there individuals on various teams who have the skills to train others?

It can be challenging to transfer technical skills learned in a classroom to a professional working environment. Staff members need to know where and how to access necessary data in their organization, and the process of using tools often looks different on an organization's actual servers than in a classroom setting. Learning from someone internally who has direct experience applying a skill is an effective way for staff members to get up to speed quickly. Formal training can always complement this in-house training later.

Does your organization's budget include funding for data-related training?

External training may not be immediately feasible if the cost exceeds available resources. When funding is limited or unavailable, alternatives to consider include identifying external partners who may be willing to support the development of staff skills, looking for free training or learning materials, identifying individuals in your agency who have the skills and could serve as trainers or mentors for other staff members, and making a professional development plan with a budget request for future staff data analytics skills training.

Assessing the training programs

WHAT'S THE RIGHT FIT FOR YOUR ORGANIZATION'S NEEDS?

What are the programs' intended learning outcomes or objectives? To what extent are the classes focused on applied or theoretical skills?

> *Theoretical skills can be useful for someone who already has experience but is looking to develop an understanding of best practices. For example, someone who routinely creates reports may appreciate training in data visualization or design theory. On the other hand, someone who wants to learn how to actually do the data visualization work will benefit more from hands-on training for a particular visualization tool.*

What types of materials and tools will participants walk away with?

> *Classes that provide materials that can be taken back and used in the workplace make it easier for staff members to remember and apply what they learned. Sharing these materials (such as code samples or detailed take-home references) might also facilitate learning among other staff members.*

What is the cost?

> *Cost will clearly factor into any decisions regarding analytics training. Weigh the cost of a training against your answers to the previous questions about learning objectives, training format, types of skills, and take-home materials. Additionally, request a course syllabus and ask if there is any prerequisite knowledge participants will need to have, as well as who will be teaching the training. If there is an opportunity for individuals to share any take-home materials and what they have learned with others on their team, a more expensive training may become more defensible.*

TOOL 5

A DOCUMENTATION CHECKLIST

One of the largest challenges agencies face in organizing and documenting information is thinking about everything that needs to be documented to ensure sustainable data use and the preservation of institutional memory. An organization might have great systems for documenting processes, projects, or data, for example, and still find itself lacking critical information in other areas. In fact, strong documentation in one area can even lead a team to become complacent and overlook other important areas.

This tool details questions that highlight the breadth of knowledge that may be important for a state agency analytics team to track. It was developed based on the research team's experience working with state agencies to build analytic capacity and later revised based on feedback from the toolkit interviewees. Depending on the size of your team and the scope of its work, you may choose to prioritize the quality of your documentation in some subset of these areas: Documentation of Source Data Systems; Documentation of Commonly Used Data Extracts, Views, or Datasets; Documentation of Frequently Used Definitions; Contextual Information About Data Collection, Quality, and Policy; Documentation of Analytics Workflows, Decisions, and Products. This tool is intended to highlight those choices for your consideration.

INSTRUCTIONS

Review the questions below and think about the tools and processes used to document this information at your agency. In each area, consider:

- Is there any documentation on this subject in your agency?
- Is there an agreed upon "single source of truth" for this information? Are there processes and training in place so users routinely update the single source of truth when they learn something new or when something changes?
- Is the documentation searchable and interpretable by most users? Do users know about the documentation and consistently access it when they should?

Once you have reviewed the full list of documentation needs, think about what your organization's next documentation priorities should be. **Balancing the importance of the information against the quality of current documentation, which areas are priorities for improved documentation processes or tools?**

Documentation of Source Data Systems

- Which of the following are available? When and how are they updated?
  - Table and field lists
  - Entity relationship diagrams
  - Data dictionaries (including field names, definitions, variable types, and indicators for primary keys)
  - Codebooks

Documentation of Commonly Used Data Extracts, Views, or Datasets

- Are there data dictionaries or codebooks for these resources?
- Where is the source code or logic used to create the data stored? Is the source code well-commented and easily understandable to at least a technical audience?
- How are dataset versions handled? Ideally, there is clear documentation of dataset versions, along with associated code and dates run (for example, if a dataset changes, is it easy to determine who made the change, when, and why?)

Documentation of Frequently Used Definitions

- How are key concepts (for example, case types, exits, program participation status, total earnings) defined in the data?
  - Are there authoritative code snippets or sets of logic to define key concepts? How are these maintained and tracked over time?
  - If the logic to define these concepts varies, is there clear documentation of where and why?

Contextual Information About Data Collection, Quality, and Policy

- What documentation exists showing how data are collected, and what guidance is given to staff members doing data collection or data entry about how different situations are coded? Relevant information might be captured in user manuals and training materials.

- As analysts or policy staff members identify problems in data elements (such as missingness, inaccuracies in data collection, or variation over time), are these findings documented?

- Changes in areas such as program eligibility or requirements impact what patterns are expected in the data. Is there consistent documentation of policy changes and their implications for the data?

Documentation of Analytics Workflows, Decisions, and Products

- Is there documentation of which analyses have been performed, either routinely or on an ad hoc basis? Can staff members easily build on what has been done before, or are they constantly called on to recreate logic, code, or analyses?

- Is there documentation of key workflows (for example, how are scripts, datasets, and result files related)? For a routine analysis, if the person who runs the report is out unexpectedly, could someone else pick it up? For an ad hoc analysis, could someone replicate the analysis months or years in the future? Process maps and clear naming conventions can help make workflow components more accessible.

- Are there log files or other records of when (and by whom) files were last changed or run? Version control tools such as Git can help track changes in code. Are final versions clearly delineated?

- Are there records of key analytics decisions (for example, why it was done the way it was done)?

- Is the original data behind any analysis clearly documented, including both the extract date and the source system?

References

Allard, Scott W., Emily R. Weigand, Colleen Schlecht, A. Rupa Datta, Robert M. Goerge, and Elizabeth Weigensberg. 2018. "State Agencies' Use of Administrative Data for Improved Practice: Needs, Challenges, and Opportunities." *Public Administration Review* 78, 2: 240–250.

Commission on Evidence-Based Policymaking. 2017. *The Promise of Evidence-Based Policymaking*. Washington, DC: Commission on Evidence-Based Policymaking.

Dube, Sara, and Darcy White. 2018. "Washington State Data Warehouse Fosters Comprehensive, Frequent Program Evaluations." PEW Charitable Trusts (blog), December. https://www.pewtrusts.org/en/research-and-analysis/articles/2018/12/06/washington-state-data-warehouse-fosters-comprehensive-frequent-program-evaluations.

Goerge, Robert M., Emily R. Wiegand, Emma K. Monahan, and Leah Gjertson. 2022. "Exemplary Data Use by State TANF Agencies." OPRE Report 2022-133. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Gooden, Susan T., Farrah S. Graham, and Kasey J. Martin. 2014. "Research Partnerships at the State Level: Bridging the Academic-Practitioner Divide." *State and Local Government Review* 46, 3: 184–196.

Lambert, David, and Julie Atkins. 2015. "New Jersey's Manage by Data Program: Changing Culture and Capacity to Improve Outcomes." Washington, DC: IBM Center for the Business of Government.

Nelson, Julie, and Syreeta Tyrell. 2015. "Public Sector Jobs: Opportunities for Advancing Racial Equity." Oakland: Government Alliance on Race & Equity.

Wavelet, Melissa, Stephanie Rubino, Caroline Morris, and Maryah Garner. 2022. "Increasing Data Analytics Capacity in State TANF Agencies: The TANF Data Collaborative Approach." OPRE Report 2022-124. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Wiseman, Jane. 2017. "Lessons from Leading CDOs: A Framework for Better Civic Analytics." Cambridge, MA: Ash Center for Democratic Governance and Innovation, Harvard Kennedy School.

NEW YORK
200 Vesey Street
23rd Flr.
New York, NY 10281
Tel: 212 532 3200

OAKLAND
475 14th Street
Suite 750
Oakland, CA 94612
Tel: 510 663 6372

WASHINGTON, DC
750 17th Street, NW
Suite 501
Washington, DC 20006

LOS ANGELES
11965 Venice Boulevard,
Suite 402
Los Angeles, CA 90066

CHAPIN HALL
AT THE UNIVERSITY OF CHICAGO

mdrc
BUILDING KNOWLEDGE
TO IMPROVE SOCIAL POLICY

COLERIDGE
INITIATIVE

OPRE

ADMINISTRATION FOR
CHILDREN & FAMILIES
Office of Family Assistance

AISP