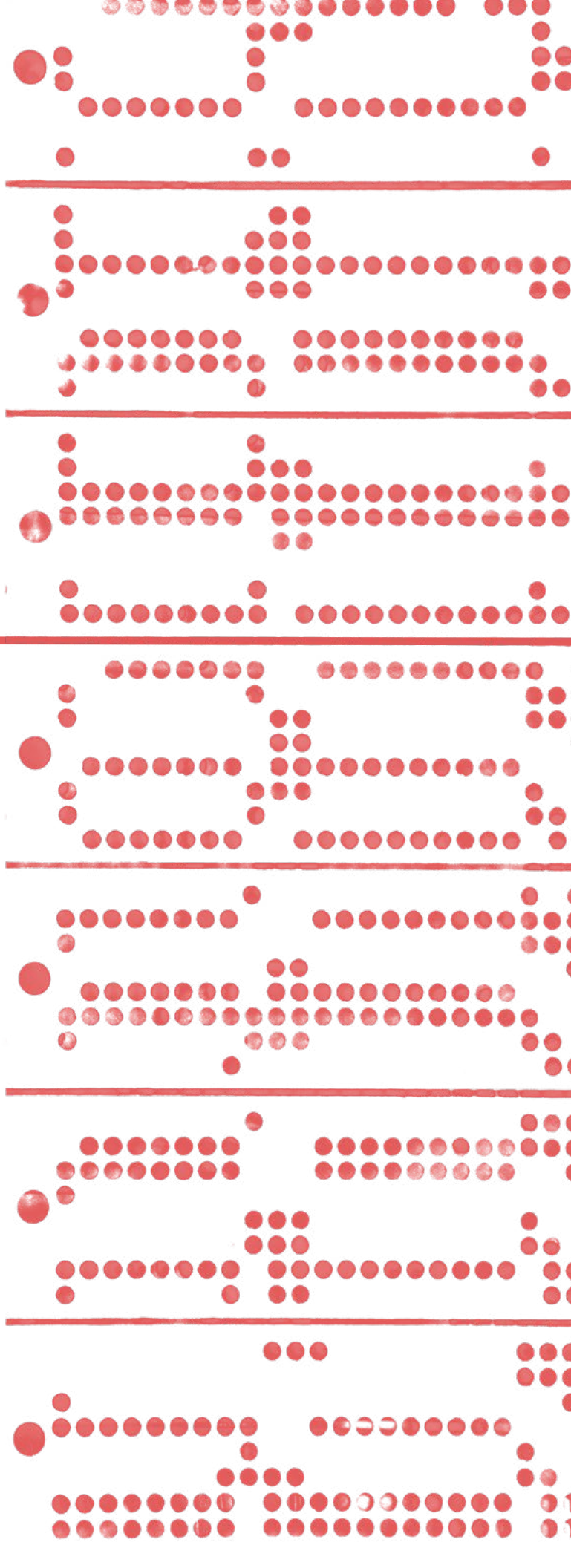
Evaluating Organizational Data Capacity & Needs in the Social Sector

Building community capacity for data ownership, analysis & communication



August 2023
Susan Paykin and Dylan Halpern

Open Spatial Lab,
Data Science Institute,
University of Chicago

Contents

Report Licensed under Creative Commons CC BY 2.0.

Evaluating Organizational Data Capacity & Needs in the Social Sector by Susan Paykin and Dylan Halpern, University of Chicago. CC BY 2.0.

Cover image based on "Punched cards from a Jacquard loom" by pedrikFollow - CC BY 2.0

Executive Summary

Data-driven projects illustrate key insights, opportunities for impact, and strategies for resilience for mission-driven organizations in the public or social impact sectors. However, without the organizational capacity to own and manage data, insight does not always translate to action. The Open Spatial Lab (OSL) at the University of Chicago's Data Science Institute works to address this gap by co-building accessible, low-code or no-code tools that make spatial data science more accessible and affordable.

From January through April 2023, OSL engaged 18 nonprofit and community-based organizations to meet with, learn from, and assess each organization's data capacities. Organizational data capacity is how organizations collect, analyze, and use data in a way that increases their overall effectiveness and creates greater stability within the organization. Our data capacity evaluation process included organizations participating in a quantitative survey assessment followed by in depth interviews. Following the assessment and interview, we synthesized our learnings in detailed and customized reports for each organization. These reports summarized data capacity areas for growth, and recommended resources to begin addressing identified challenges or pain points for each organization. A select group of organizations were also invited to participate in the next stage Data Tool Collaboratory, based on data project needs and their own interest and capacities.

Process

Inputs

Data.org Data Maturity Assessment (DMA): The first step in the evaluation process was for organizations to complete the Data Maturity Assessment (DMA). The DMA is a web-based survey tool designed to help organizations understand their "data maturity", defined as an organization's ability to effectively use data to meet the needs of their mission. The survey includes 30 multiple choice questions across three main categories: Purpose, Practice, and People. The results are numerical scores on a scale of 1-10, which is calculated based on a simple weighting scheme.

Qualitative Interview: After organizations completed the DMA and OSL reviewed the results, organizations and OSL engaged in a 1-hour qualitative interview to unpack the DMA questions and better understand their data capacities and needs. In this interview, OSL asked questions about their overall data goals, specific reactions to the DMA process and results, current capacity needs, and any other data- or project-specific pain points. We took detailed and comprehensive notes which were then synthesized into a customized report.

Outputs

Customized report: Following the DMA and interview, OSL prepared a customized report for each organization. These reports included recommendations on next steps to reach data goals, including recommending free or low-cost data tools, training, and resources. It integrates the DMA responses with the qualitative interview responses so that documentation is clear organizations can refer back to specific areas. The report was shared internally with staff that engaged with the evaluation process but several groups reported sharing it with their leadership and board members.

Selection in OSL's Data Tool Collaboratory: Following the evaluation process, we shared the opportunity to participate in the Open Spatial Lab's Data Collaboratory, a 6-month collaborative data tool building program. This was limited to 8-10 organizations who completed both stages of the process and were interested in public-facing data tools. In the Collaboratory, the organization commits to working with OSL to "cobuild" a custom built data tool. This engagement process includes iterating with OSL leads and regular meetings through fall 2023.

Areas for Growth

Across the 18 organizations engaged in the assessment and interview process, several common themes emerged relating to their data capacity needs. The most commonly shared topics fell under two broad categories: Culture & Capacity and Data Use & Application. These include baseline capacity questions around data management, opportunities to advance analytics across temporal and spatial dimensions, and approaches to decentralizing data skills and talent beyond a single team. These insights are useful for anyone working to diversify representation in data science and integrate these practices into nonprofit structures. These are also critical for industry or academic partners to be aware of when approaching nonprofits for community engagement or research projects.

Culture & Capacity

- Baseline capacity: Building and maintaining databases, data management, warehousing, and security
- Data culture: Strengthening culture that supports the importance of data within and across the organization
- Decentralization: Distributing data capacity beyond a single individual or team

Data Use & Application

- Data visualization: Integrating easy-to-use tools or workflows, such as automated report generation
- Analysis: Advancing analysis over temporal or spatial dimensions
- Aspirational: Increasing use of scripting for streamlining, replication, and consistency

This report details these common areas for growth and highlights the challenges and opportunities in building community capacity for data ownership. It also provides a comprehensive overview and key insights from our organizational data assessment process.

Finally, it highlights how the Open Spatial Lab is using these insights to directly inform our current collaborative data tool building program with a select group of organizational partners and our work ahead.

Introduction

Data-driven projects can illustrate key impacts and inform evidence-based strategies and practices for mission-driven organizations, particularly in the public or social impact sector.

However, without the organizational or community's capacity to own, manage, and analyze data, insight does not always translate to action. The Open Spatial Lab (OSL) at the University of Chicago's Data Science Institute works to address this gap by co-building accessible, low-code or no-code tools that make (geo)spatial data science more accessible and affordable. Through reducing barriers to entry for data analysis and visualization, more people can use data to understand the challenges and opportunities in their communities. Through reducing long-term costs and making updates easier, data tools and infrastructure can exist for longer and be more responsive to changing data and communication needs.

OSL works closely with partner organizations, such as nonprofits, community-based organizations (CBOs), and government agencies, to understand data needs and to develop low-code and no-code tools that can help them activate data to achieve their missions. Our programs emphasize ownership of tools and infrastructure so that organizations can continue to use and update these tools effectively after the initial development phase. By making geospatial data science more accessible and affordable, we work to broaden the perspectives that can analyze and communicate data in its many forms.

Evaluating Organizational Data Capacity

From January through April 2023, OSL engaged 18 nonprofit and community-based organizations to meet with, learn from, and assess each organization's data capacities. Organizational data capacity is how organizations collect, analyze, and use data in a way that increases their overall effectiveness and creates greater stability within the organization. The evaluation process included engaging organizations through a survey followed by in-depth interviews. Following the assessment and interview, we synthesized learnings in detailed and customized reports for each organization. These narrative reports summarized the assessment results, areas for growth, and recommended resources for addressing challenges or pain points in each organization.

Through the assessment process, we identified common themes across diverse perspectives from organizations of varying scales and focus areas. These include baseline capacity questions around data management, opportunities to advance analytics across temporal and spatial dimensions, and approaches to decentralizing data skills and talent beyond a single team. These insights are useful for anyone working to diversify representation in data science and integrate these practices into nonprofit missions and organizational structures. It is similarly critical for industry or academic partners to know when approaching nonprofits for community engagement around data or research projects.

This report shares insights from our organizational data assessment process, highlights common emerging themes across nonprofits and organizations, and discusses the challenges and opportunities in building community capacity for data ownership.

We employed mixed quantitative and qualitative methods for evaluating organizational data capacities in the social sector. To capture a standardized numerical baseline of organizational practices, we used Data.org's Data Maturity Assessment (DMA). After organizations completed the DMA survey, we met with organizations for 1:1 in-depth interviews to discuss their results and dive deeper into the details around their current data capacity challenges and aspirations.

Figure 1. OSL Data Capacity Evaluation participating organizations

Organization	Location	Focus Area
Air Alliance Houston ♥	Houston, TX	Health, Energy & Environment, Racial and Social Justice
Californians for Pesticide Reform ◘♦+	California	Energy & Environment
Charleston Community Research to Action Board ♥	Charleston, SC	Energy & Environment, Other
Circulate San Diego ♥+	San Diego, CA	Urban Development
Data You Can Use ♥	Milwaukee, WI	Education, Health, Urban Development, Energy & Environment, Public Safety, Economic Development, Racial & Social Justice
El Buen Samaritano ♥	Austin, TX	Social Services
Equiticity +	Chicago, IL	Urban Development, Public Safety, Racial & Social Justice
Ethiopian and Eritrean Cultural Resource Center ♥+	Portland, OR Metro Area	Education, Health, Social Services, Energy & Environment, Public Safety, Economic Development, Racial and Social Justice
Greater Greater Washington ♥	Washington, DC Metro Area	Urban Development
Groundwork Bridgeport ♥+	Bridgeport, CT	Urban Development, Energy & Environment
Higher Expectations for Racine County ♥+	Racine, WI	Education, Social Services, Economic Development, Racial and Social Justice
Kids First Chicago	Chicago, IL	Education
Maryland Center on Economic Policy ♥	Maryland	Education, Health, Social Services, Urban Development, Economic Development, Gender Equality, Racial and Social Justice, Other
Milwaukee Succeeds +	Milwaukee, WI	Education, Racial and Social Justice, Other
My Block My Hood My City +	Chicago, IL	Social Services, Urban Development, Racial and Social Justice
National Nurse-Led Care Consortium	National	Health, Social Services
Plan4Success	Cook County, IL	Education, Health, Social Services, Urban Development, Racial and Social Justice
Prevent Blindness ♦+	National	Health

♥ RWJF-Urban Institute grantee
◘ UChicago DSI 11th Hour Project grantee
♦ Capacity Accelerator Network (CAN) Project
+ OSL Data Collaboratory Partner

Data Maturity Assessment (DMA) Quantitative Survey Tool

In this first year of co-building data tools with social sector organizations, it was important to us to have a comprehensive understanding of their data challenges. In order to evaluate across organizations in a systematic way, we wanted to capture baseline indicators of the actions or structures in data work that organizations are putting in place.

There is an established field of organizational capacity evaluation tools and frameworks with a few starting to focus more on data capacity specifically. Not wanting to "reinvent the wheel", we first conducted an initial landscape analysis of existing nonprofit and social impact organizational evaluations. We reviewed dozens of tools and frameworks for evaluating organizations capacity and honed in on seven existing options that aligned with our goals. These are identified in Figure 2 below.

Figure 2. Existing Tools and Frameworks for Evaluating Organizational Capacity

Evaluation Tool	Organization or Author	Year
Data Maturity Assessment	data.org	2022
Community Research Activity Assessment Tool (CREAT)	Debbie Humphries et al (2019)	2019
Organizational Capacity Assessment (OCA)	Pact	2017
Assessment of Learning Outcomes and Social Effects (ALSE) of Community-Based Education Project	NYU Institute of Human Development & Social Change	2016
Participatory Monitoring and Evaluation System Assessment Tool	FHI 360	2013
What Makes an Effective Coalition? Report	The California Endowment	2011
Organizational Capacity Assessment Tool	Marguerite Casey Foundation	2005

After reviewing results from the tool landscape analysis, we selected the Data.org Data Maturity Assessment (DMA) because it most aligned with our proposed research goals and engagement process. The DMA was developed by Data.org, an organization that seeks to democratize and reimagine the use of data to tackle challenges and improve lives across the globe. Prior to this project, colleagues at the Data Science Institute at the University of Chicago consulted with Data.org on developing the DMA, including drafting question topics and resources.

The DMA is a web-based survey tool with a simple user interface; it is significantly less time-consuming compared to other evaluation tools oriented toward evaluating organizational culture, including but not limited to data work. The DMA focused on data culture and practices with 30 questions with multiple choice or short answer responses.

The goal of the DMA is to provide a snapshot of organizations' data "maturity" - defined here as an organization's ability to effectively use data to meet the needs of their mission - across three main categories: Purpose, Practice, and People. It takes 10-15 on average to complete. At the end, the DMA gives a report / summary page, reporting mean scores from questions across the three main categories and subcategories. It also provides an overall score on the scale of 1-10, with a categorical description for different levels (Figure 3).

Please see Appendix Figure 1 for a complete list of the DMA's question topic categories and subcategories, as well as the averages scores across all organizations.

Figure 3. Data Maturity Assessment Scoring Rubric

Category	Overall Score
Data Curious	0 - 1.9
Data Informed	2.0 - 3.9
Data Guided	4.0 - 5.9
Data Driven	6.0 - 7.9
Data Led	8.0 - 10.0

These scores provide a benchmark that organizations can use to measure their progress going forward. We reviewed organizations' assessment results, including their scores and responses to individual questions. We used the DMA scores and responses as a foundation upon which to develop clarifying questions to better understand their organizations' challenges through the next stage, in-depth interview.

Interviews: Qualitative Assessment Tool

We met virtually for one hour in-depth interviews with each organization after they completed the DMA. The goal of these sessions was to provide a forum for more detailed or nuanced discussions around the organizations' data goals and to contextualize the quantitative results of the DMA. This understanding of organizational capacities and use cases directly informs our tools' baseline features. We asked open-ended questions in a standardized order and format. We used a racial equity framework to develop these questions, focused on establishing a common language around data practice, allowing organizations to identify current assets and frame their own challenges. We followed the same template for each organization, asking the most of the same questions across organizations and customizing details on specific questions or issues based on results from their DMA survey. See Figure XX for the interview questions asked of all organizations. The full interview question template is available in the Appendix.

Figure 4. Interview Questions

Interview Questions	Questions
Big Picture	How does data-driven or data-informed work contribute to your core mission?
Assessment	What are some of the key takeaways for you from the Data Maturity Assessment? Do the assessment results align with what you expected? After completing the assessment, how might data better serve your organization? Observations and follow up questions: From reading your assessment, here are a few things that stood out to us: • Example: On the data completeness question, "Our data is complete - we have all the data we need", you answered 2/5, compared to data accuracy scored at 4 and data consistency across the organization scored at 2. What opportunities to do see for additional more data to fill the gaps and keep things consistent?
Capacity	What data resources or training do you wish you had? What specific tools, expertise, or analysis do you wish your organization had?
Projects & Pain Points	What are some data-related projects your organization is working on? What do you find missing or are pain points or limitations? Is there anything else you'd like to share, or do you have any questions for us?

Customized Report & Recommended Resources

Following each organization's DMA survey and follow-up interview, we prepared a written narrative report outlining the context and relevance of the results. The report was shared directly with the organizational contacts within 1 week of the interview. These reports identified key findings, areas for growth, and recommended tools and resources to address immediate data-related challenges, as well as an Executive Summary and Next Steps sections with information on continuing to engage with us. It also contained the full responses to the DMA as well as all the notes taken during the interviews. See Figure 5 for an example of the customized reports where we shared resources and tools to address specific pain points that were identified by the organizations in their DMA results and interview.

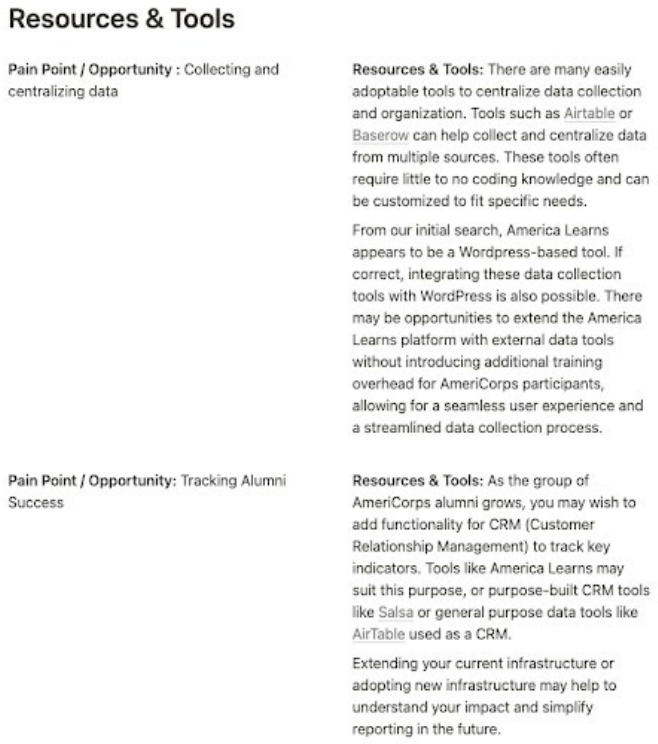


Figure 5. Customized Resources & Tools recommendations in the Data Capacity Report for organizations (right)

Key Findings

OSL engaged 18 organizations in the data capacity evaluation process between January and April 2023. The organizations represented many different missions and focus areas, sizes, budgets, demographics, and communities served. Below are some key details and context on the organizations and their participating staff. Taken as a whole, these trends matched our initial target of reaching small- to mid-scale social impact organizations. These organizations often lack the resources, institutional infrastructure, or staffing for internal capacity for data management, analysis, and visualization.

Overview of Organizations

All organizations engaged in the OSL evaluation were based in and their work focused on U.S. communities. Most organizations worked across multiple fields; the most common selected were Racial and Social Justice, Urban Development, Health, and Education. Eleven organizations chose more than one field; nine of these indicated three or more fields.

The majority of organizations' total annual budgets were below $5 million; eight organizations reported budgets below $1 million and eight with budgets falling between $1-5 million. There did not appear to be a significant difference in the overall scores based on annual operating budgets across the 18 participating organizations. Only two organizations had annual budgets larger than $5 million. Individuals that completed the self-guided DMA assessment came from different staff roles and departments from their organizations: 8 executives, 6 data professionals, 2 program staff, and 2 others (non-specified).

Mission Areas

Field	Number of organizations
Racial and Social Justice	9
Urban Development	8
Health	7
Education	7
Social Services	7
Energy & Environment	5
Economic Development	4
Public Safety	3
Other	3
Gender Equality	1
Humanitarian Relief	0
International Development	0

Organization Locations

Regions	Locations	Number of organizations
Midwest	Racine, WI; Milwaukee, WI; Chicago and Cook County, IL	8
South	Charleston, SC; Austin, TX; Houston, TX	3
West	Portland, OR; San Diego, CA; California (state-wide)	3
Northeast	Bridgeport, CT; Philadelphia, PA	2
Mid Atlantic	Maryland (state-wide); Washington, DC	2

Average DMA Scores Across All Organizations

	Mean	Median	Min	Max
Overall Score	4.5	4.3	1.7	6.9
Purpose	4.9	5.2	2	7.7
Application	5.2	5	2.2	8.9
Analysis	4.2	3.8	0.9	7.5
Strategy	5.3	5.4	0.8	9.6
Practice	4.4	4.2	1.4	8.9
Quality	4.1	3.3	0.8	7.9
Security	2.9	1.8	0.5	9
Responsible Use	4.8	4.7	0.6	8.4
Infrastructure	5.7	6	0	10
People	4.4	4.4	1.7	7.9
Leadership	5.1	4.4	0.6	10
Talent	4.1	3.8	0	8.8
Culture	4.1	4.1	0	8.8

Average Scores and Operating Budgets

The majority of organizations' total annual budgets were below $5 million, with eight with budgets below $1 million and eight with budgets between $1-5 million. There did not appear to be a significant difference in the overall scores based on annual operating budgets across the 18 participating organizations.

	< $500k	$500k to $999k	$1M - $4.9M	$5M - $9.9M	$10M - $24.9M
No. of orgs	3	5	8	1	1
Overall Score	4.3	3.9	4.8	6.7	4.1
Purpose	4.8	4.5	5.2	5.1	4.9
Practice	3.6	3.9	4.4	8.9	4.4
People	4.9	3.5	4.9	5.0	3.1

Average Scores and Staff Size

The organizations engaged ranged from small community-based organizations (CBOs) to national nonprofits. All organizations reported having fewer than 100 employees on staff, with a majority (10 organizations) with 10 or fewer employees. There was no meaningful difference across the spectrum of organizations of this size in terms of their overall scores.

	< 10 people	11 - 100 people
No. of orgs	10	8
Overall Score	4.9	4.2
Purpose	4.9	4.8
Practice	5.1	3.8
People	4.5	4.7

Areas for Growth

Across the 18 organizations engaged in the data capacity evaluation process, several common themes emerged.

These themes could be grouped into two main categories: Culture & Capacity and Data Use & Application. They include baseline capacity questions around data management, advancing analytics across temporal and spatial dimensions, and decentralizing data skills and talent beyond a single team. These insights may be useful for anyone working to diversify representation in data science and integrate these practices into nonprofit structures. These are also critical for industry or academic partners to be aware of when approaching nonprofits for community engagement or research projects. We expand on each of these themes and share examples from organizations in the section below.

Culture & Capacity

- Baseline capacity: Building and maintaining databases, data management, warehousing, and security
- Data culture: Strengthening culture that supports the importance of data within and across the organization
- Decentralization: Distributing data capacity beyond a single individual or team

Data Use & Application

- Data visualization: Integrating easy-to-use tools or workflows, such as automated report generation
- Analysis: Advancing analysis over temporal or spatial dimensions
- Aspirational: Increasing use of scripting for streamlining, replication, and consistency

Culture & Capacity

Baseline capacity: Building and maintaining databases, data management, warehousing, and security

Many organizations that we spoke to had a wealth of data already collected about people who engage in their programs or communities impacted by their work, and they already understood the importance of collecting data to better understand internal and external trends. But most were not familiar with what to do beyond saving data and information in one-off spreadsheets. We saw that for many groups, the foundational process of centralizing data in a single location would allow for easier collection and insights. Furthermore, ensuring the appropriate data security policies and practices are in place is necessary for continuing public-facing work and collecting data in a safe and secure way.

Data visualization: Integrating easy-to-use tools or workflows, such as automated report generation

Organizations largely know there are many visualization platforms or softwares available, but are not sure how to choose one or then get started. They reported lacking the capacity to do the deeper learning required to onboard a new platform or tool systems-wide. The most common use case they reported finding visualizations helpful was for communicating information with leadership or foundations/funders. Data visualization templates or tools that enabled automated processes to populate program or funder reports were often at the top of their data wishlists. Furthermore, while there are many advanced solutions, most are paid subscription models. Free tiers are often limited to 1-2 users with limits on data, and meet some but not all of the needs of a long-term sustainable solution.

Data culture: Strengthening culture that supports the importance of data within and across the organization

Many organizations shared that they were lacking a strong data culture, or an environment that supported the importance of data within and across the organization. Particularly for smaller organizations that lack time and resources as well as organizations that did not previously work with data, it can be harder to take up new technologies or practices. Building culture requires planning, resources, and investments. Investing in creating new staff roles or hiring staff with the training and skills required for working with data can be a significant barrier for smaller organizations. When asked what kinds of data training or resources were needed, many organizations responded that an on-staff data role would make the largest difference. One partner described a "Catch 22" in fostering a stronger data culture -- they needed more training and expertise on their staff to build culture, but needed a stronger data culture to support and hire a role with the training and expertise.

Decentralization: Distributing data capacity beyond a single individual or team

Many organizations have siloed processes when it comes to data work, with one person or one team having the bulk of the responsibility and/or skills-base required. Outside of this core group, however, programmatic or operations staff are often involved with data collection or reporting. Staff often lack the context of the data they are collecting or reporting, which may affect the accuracy and consistency in their inputted responses. For example, are surveys used for data collection asking the right questions to get the information the organization needs to understand the current picture of food insecurity in their community? How might more internal data skills sharing help staff change data collection processes? We heard calls for expanding transparency and investment to build a culture of responsible data use shared across the organization.

Furthermore, the capacity to conduct more complex analyses on data available is currently limited by the small number of nonprofit staff members with analytic expertise. There are clear needs to build out more capacity for conducting and managing data science and analytics workflows. There is also a need to clarify the role data plays in leadership decision-making, and where possible formalize relationships between data-driven processes and outcomes with organizational goals and decisions.

Data Use & Application

Analysis: Advancing analysis over temporal or spatial dimensions

Many organizations were engaged in data collection and reporting in some way, but reported that they often did have enough time to analyze and reflect on data to drive their strategy. For example, descriptive analyses of subset communities within the overall population served, or time-series analyses capturing program changes year-over-year, could help better understand the impact of policies or programs over time, place, or groups of people. These more complex analyses are often beyond current capacity. There is an opportunity to make these more accessible by setting up standardized workflows that automate analyses for exploring and uncovering trends.

Programming and Automation: Increasing use of scripting for streamlining, replication, and consistency

We learned that previous conventions of spreadsheet-based approaches to data management and analytics often led to work being singular or "one-offs" rather than a reproducible and reusable workflow. Especially for more complex analyses, this means that reviewing and validating results can be an onerous and time consuming task. Establishing a modern data science workforce is part of the dream for several of these organizations. Most analyses are performed in desktop environments, and validating or sanity-checking outputs rely on manual methods to confirm results. This is a valid and standard approach to analysis, but there are opportunities to streamline analyses or frequently used methods. Additionally, more technical data output validation checks remain challenging. Utilizing more formal workflows, generated through visual or scripted programming languages, may reduce the administrative burden of validating results.

Overall, organizations were enthusiastic about expanding or growing their work with data, but lacked the knowledge or resources to get started beyond basic spreadsheets. Small nonprofits are capacity-constrained in terms of the staff time, funding, and talent needed to invest in the strong foundations in data practice.

Considerations

There is no industry-wide standard to evaluate the large and changing field of organizational data capacity in the nonprofit or social impact sectors, nor is there a one-size-fits-all approach. The specific questions we sought to ask and answer were: What are the barriers and pain points to organizations being comfortable with using data? And what, aspirationally, do organizations want to do with data? We knew that some basic level of standardized, quantitative survey responses would be helpful in trying to seek answers, but likely not as helpful to the organizations participating. The follow-up in-depth interview was just one approach to bridging the quantitative and the qualitative. There may be other research methods or combinations of methods that may be appropriate in different contexts or different populations.

This hybrid qualitative and quantitative approach required significant time and effort from the OSL team and the participating organizations. We developed this approach tailored to our target audience of small to mid-scale organizations typically without their own data teams or tech staff. There may be formats better suited to larger organizations or those working in a specific topic area. In assessing results across organizations, it is important to consider the diverse organizational positions of survey respondents. Future applications of this approach could encourage multiple organization members to participate in the initial survey, or emphasize a specific organizational role when engaging.

The interview enabled us to learn many more details than the DMA alone. We limited the interviews to 60 minutes; however, many of these interviews spurred conversations and questions that could have continued for much longer. Generally organizations were eager to discuss their challenges and data goals. We heard repeatedly that the assessment process and interview conversation were helpful to them. Future evaluations may benefit from two-part interviews, for example, or allocated more time.

Next Steps

Co-building low-code data tools

The insights from these assessments are directly informing Open Spatial Lab's data tool development and capacity-building work. Using responses from the DMA as well as insights uncovered in our interviews, we identified 9 organizations who met the criteria for joining OSL's Data Tool Collaboratory program. See Appendix Figure 4 for a full description of current Collaboratory partners and projects. In the Data Tool Collaboratory, we work closely with organizations to build a customized tool for their data, analysis, or communication needs, free of charge. In turn, we ask for our partners' engagement and feedback to inform tool development, via monthly meetings and prototype trials to iterate along the way. At the end of the Collaboratory program, the organizations leave with a customized data tool and expanded data infrastructure capacity and training to own and update their data going forward. Through this co-building process, we are integrating learnings from the evaluation into Nectr, a low cost and no-code data management, analysis, and visualization web-based software. Nectr is licensed as open source software and will be made publicly available beyond the Collaboratory project partners in late 2023.

Conclusion

Data is one of the most valuable tools for any social impact organization.

Data can help organizations better understand the communities they serve, identify gaps in their services, or inform how many people a proposed policy might affect. In light of this, we wanted to understand why so many groups report struggling with integrating data processes, and how better data tools that address these challenges. The organizations that we engaged with were all subject matter experts in their respective fields doing mission-driven and community impact work. They were all interested in learning more about data, but more often than not, weren't sure how or lacked the means to start.

After engaging 18 environmental and social impact organizations in the Data Maturity Assessment quantitative survey and the qualitative interview process, we collected and analyzed our results to understand common roadblocks to adopting or expanding technical data work. Several common themes emerged. This evaluation found that organizations lacked resources (primarily financial but also human resources), organizational culture (data collection or analysis is a new aspect of work that hadn't been a previous focus), and staff capacity (current staff at capacity and new roles need to be created and funded). The key findings and common areas for growth identified in this assessment are directly informing OSL's current tool building efforts with organizational partners in the Data Tool Collaboratory program. In the next year, we will continue to use these learnings to build and share open source tools and work to expand and diversify the audiences participating in data science.

Acknowledgements

This work would not be possible without the support of several partner organizations and supporting institutions. We want to acknowledge and thank the Robert Wood Johnson Foundation for supporting OSL's program and activities. Thank you also to the National Neighborhood Indicators Partnership at the Urban Institute for connecting us with organizations and your network. The Open Spatial Lab is grateful to the Center for Spatial Data Science and the Data Science Institute at the University of Chicago for their ongoing support and collaboration. Lastly, we thank the nonprofit and community organizations who shared their time and efforts with us in this process.

THE UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO DATA SCIENCE INSTITUTE

THE CENTER FOR SPATIAL DATA SCIENCE THE UNIVERSITY OF CHICAGO

Funding provided in part by:

Robert Wood Johnson Foundation

Appendices

Appendix 1. Summary statistics of scores across all categories

n = 18 organizations

	Mean	Std. Dev.	Min	Pctl. 25	Median / Pctl. 50	Pctl. 75	Max
Overall Score	4.5	1.6	1.7	3.2	4.3	5.8	6.9
Purpose	4.9	1.7	2	3.6	5.2	6.2	7.7
Application	5.2	2	2.2	3.5	5	6.6	8.9
Analysis	4.2	1.8	0.9	2.6	3.8	5	7.5
Strategy	5.3	2.6	0.8	2.8	5.4	6.7	9.6
Practice	4.4	2	1.4	2.8	4.2	5.4	8.9
Quality	4.1	2.4	0.8	2.2	3.3	6.8	7.9
Security	2.9	2.5	0.5	1	1.8	4.5	9
Responsible Use	4.8	2.2	0.6	3.8	4.7	6.2	8.4
Infrastructure	5.7	2.6	0	4.6	6	7	10
People	4.4	1.8	1.7	3.2	4.4	5.2	7.9
Leadership	5.1	2.4	0.6	3.8	4.4	6.8	10
Talent	4.1	2.3	0	2.5	3.8	6	8.8
Culture	4.1	2.4	0	2.5	4.1	5	8.8

Appendix 2. Average scores across all categories by operating budget

n = 18 organizations

	Less than $500 thousand (3 orgs)	$500 thousand to $999 thousand (5 orgs)	$1 million to $4.9 million (8 orgs)	$5 million to $9.9 million (1 org)	$10 million to $24.9 million (1 org)
Overall Score	4.3	3.9	4.8	6.7	4.1
Purpose	4.8	4.5	5.2	5.1	4.9
Application	5.6	4.5	5.4	7.5	4.2
Analysis	3.8	4.5	4.4	2.5	5.0
Strategy	5.0	4.4	6.0	5.0	5.8
Practice	3.6	3.9	4.4	8.9	4.4
Quality	3.5	3.6	4.6	7.5	0.8
Security	1.7	2.5	2.6	9.0	5.0
Responsible Use	5.4	4.4	4.4	8.4	4.7
Infrastructure	3.9	5.3	6.1	10.0	5.8
People	4.9	3.5	4.9	5.0	3.1
Leadership	4.8	4.5	5.6	6.3	3.8
Talent	4.2	2.9	4.5	6.3	3.8
Culture	5.6	3.1	4.7	2.5	1.9

Appendix 3. Average scores across all categories by staff size

n = 18 organizations

	Fewer than 10 people	11-100 people
Overall Score	4.9	4.2
Purpose	4.9	4.8
Application	5.6	4.0
Analysis	3.6	4.8
Strategy	5.9	4.8
Practice	5.1	3.8
Quality	4.9	3.4
Security	3.8	2.2
Responsible Use	5.2	4.5
Infrastructure	6.5	5.1
People	4.5	4.7
Leadership	5.7	4.7
Talent	4.6	3.7
Culture	3.3	4.8

Appendix 4. Data Collaboratory 2023 Project Partners & Descriptions

Californians for Pesticide Reform ◘♦

California

OSL is developing a new data tool to understand and track pesticide use across the state of California and that enables users to query, visualize, and export data. CPR and its partner organizations previously used a tool developed by researchers at UC Davis called PURwebGIS to help inform the coalition's advocacy work. However, PURwebGIS' server access and instability were challenges to long-term use. This project aims to make a pesticide tracking data tool that remains sustainable and stable online, and transparent to reuse, update, or move to a new environment.

Circulate San Diego ♥

San Diego, CA

OSL is developing an interactive map and data platform integrating multiple data sources to be used for sidewalk audits, community meetings, and advocacy efforts. Circulate San Diego aims to use this tool to facilitate public engagement through programs including sidewalk audits and community planning meetings, and identify infrastructure issues that have persisted over time to inform advocacy work. It will aim to uncover insights into where safety or infrastructures issues have remained in the area over time, which communities are most impacted, and how institutional data sources compare with sidewalk audits and direct community feedback.

Equiticity

Chicago, IL

OSL is developing a custom data explorer and interactive visualization tool to analyze insights around automated traffic enforcement and racial equity in the City of Chicago. Equiticity intends to use this data tool in its program and advocacy work toward reframing public safety and building effective campaigns for racial equity, transportation safety, and mobility justice. Specifically, this data tool will be used in part to analyze inequitable impacts of automated enforcement policies and practices through linking socioeconomic and demographic data with automated enforcement (red light cameras and speed cameras) spatio-temporal data.

♥ RWJF-Urban Institute grantee ◘ UChicago DSI 11th Hour Project grantee ♦ Capacity Accelerator Network (CAN) Project

Groundwork Bridgeport ♥
Bridgeport, CT

OSL is developing an interactive data explorer for its tree planting activities spanning multiple programs. Groundwork Bridgeport aims to develop metrics and strategies to help prioritize their tree planting and maintenance efforts and maximize impact. The explorer tool will be integrated with the organization's existing tree database to track and visualize metrics related to impact, including environmental and social impact trends over temporal and spatial aggregations. The tree data tool will also integrate socioeconomic data to highlight communities disproportionately harmed by environmental racism or opportunities for more equitable tree planting and coverage.

Higher Expectations for Racine County ♥
Racine, WI

Milwaukee Succeeds
Milwaukee, WI

Working with two organization, Higher Expectations for Racine County and Milwaukee Succeeds, in the Strive Together Wisconsin Partnership, OSL is developing an automated data pipeline and data explorer tool that consolidates Wisconsin (state-wide) early childhood education data into accessible, interactive dashboards and one-pager reports. Technical goals of this project include developing an open source, easily-maintainable tool that runs on automated triggers. The audience is Higher Expectations & Milwaukee Succeeds' partners engaged in early childhood education programs, advocacy, and policy.

My Block My Hood My City
Chicago, IL

OSL is developing a centralized, interactive data and visualization tool that captures, tracks, and communicates the impact of M3 programs and activities. The goal of this engagement is to help uncover insights into how effectively M3 is achieving its mission and where it should focus more efforts and resources. Technical goals of this project include developing an open source, easily-maintainable (low-code) tool. The audience is My Block My Hood My City's staff across the organization, leadership and Board, as well as outside funders.

Prevent Blindness ♦
National

OSL will work with Prevent Blindness to develop a database, interactive data explorer and map tool that can help identify gaps in vision care provider locations and promote equitable access to eye care. Prevent Blindness seeks to identify, understand, and map currently available provider data in order to align their service and advocacy efforts and target resources where they are currently lacking or under-serviced in vision care. This data tool will also link the socioeconomic and health conditions in communities (such as zip codes and counties) currently under-served or disproportionately impacted by a lack of vision providers.

♥ RWJF-Urban Institute grantee

◘ UChicago DSI 11th Hour Project grantee

♦ Capacity Accelerator Network (CAN) Project

Appendix 5. Interview Questions

We used the following standardized questions to interview organizations in the Data Capacity Evaluation process. All interviews were 1-hour long, conducted over Zoom, and occurred between January and April 2023.

Big Picture

- How does data-driven or data-informed work contribute to your core mission?

Assessment

- What are some of the key takeaways for you from the Data Maturity Assessment?
- Do the assessment results align with what you expected?
- After completing the assessment, how might data better serve your organization?

From reading your assessment, here are a few things that stood out to us:

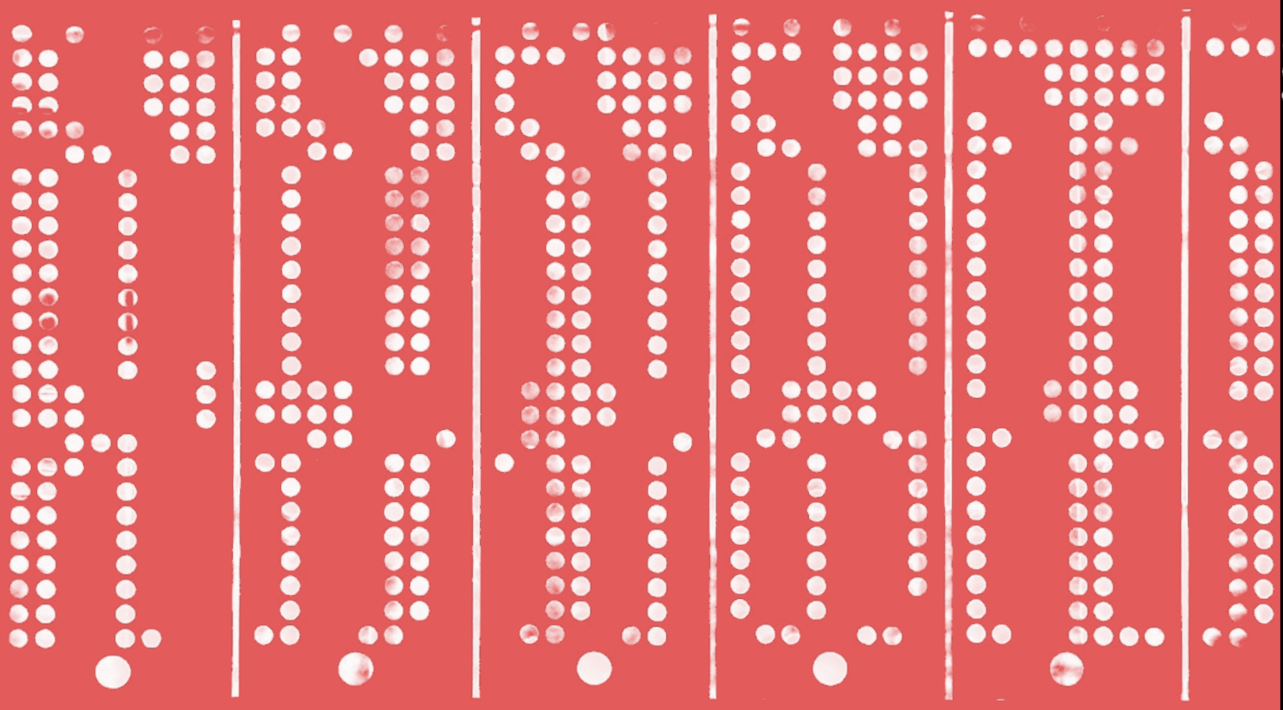
- Example: On the data completeness question, "Our data is complete - we have all the data we need", you answered 2/5, compared to data accuracy scored at 4 and data consistency across the organization scored at 2. What opportunities to do see for additional more data to fill the gaps and keep things consistent?
- Example: Your "People" score is overall very positive, with a strong data culture score in particular. However, the leadership score is lower (2.5 / 10). This relates to questions like: "Does your organization have at least one trusted data expert on its senior leadership team?" What does this look like at your organization? Are there opportunities for leadership to be more involved in data processes, culture, or data-driven work?
- Example: You noted that most of your data processes are manually input into reports or dashboards. Do you encounter any issues with this? Are there repeated tasks that could be automated?

Capacity

- What data resources or training do you wish you had?
- What specific tools, expertise, or analysis do you wish your organization had?

Projects & Pain Points

- What are some data related projects your organization is working on?
- What do you find missing or are pain points / limitations?
- Is there anything else you'd like to share, or do you have any questions for us?



Evaluating Organizational Data Capacity & Needs in the Social Sector

Building community capacity for data ownership, analysis & communication

August 2023 :: Susan Paykin and Dylan Halpern
Open Spatial Lab, Data Science Institute, University of Chicago

Report Licensed under Creative Commons CC BY 2.0