



USING DATA TO POWER SCALE

Innovation Investment Alliance, Skoll Foundation, and CASE at Duke

ABOUT THE SCALING PATHWAYS SERIES

A Google search for “social enterprise” calls up over 400 million links. Indeed, there are hundreds of thousands of new ideas for mission-driven ventures emerging around the world. And there are some notable social enterprise organizations that have started to solve social and environmental problems at scale. What can we learn from the experiences of these organizations? Their hard-won lessons can benefit other social enterprises, funders, and the surrounding ecosystem.

Social enterprises often work on problems that are deeply entrenched, depend on cross-sector collaboration, and require multiple pathways to scale their impact and create systems-level change. The road to scaled impact is a nonlinear, complicated one. Along the way, the organizations have to overcome many challenges and roadblocks, including the following:



FINANCING FOR SCALE: Determining which financing strategies best support their plan for impact at scale.



GOVERNMENT PARTNERSHIPS: Effectively cultivating and managing partnerships with government and other actors in order to increase impact.



TALENT: Defining the different talent strategies needed to identify, train, and retain the human capital needed for scale.



DATA: Understanding how to best use data to drive performance, impact management, and decision-making at scale.

The Scaling Pathways Theme Studies Series dives into each of these topics in depth, bringing to light lessons learned by successful social enterprises that have navigated these challenges on the road to scaled impact.

Acknowledgments:

The authors wish to acknowledge the insights and feedback of many stakeholders who helped to shape this theme study. The organizations highlighted within the study shared their insights and lessons to inform the paper. Chris Walker and Amanda West of Mercy Corps, Liz Diebold, Elyssa Lewis, Brittany Boettcher, Lucien Chan, and Theresa Chen from the Skoll Foundation, and others from these organizations offered guidance throughout and provided feedback on drafts. Finally, we are grateful to members of the CASE team, including Cathy Clark, Robyn Fehrman, Carrie Gonnella, Tara Schoenborn, and Erin Morant for advisory and programmatic support.

This study was funded by the Skoll Foundation, and was prepared by CASE at Duke University in collaboration with Mercy Corps Ventures, the Skoll Foundation, USAID, and the organizations described within. The findings of this report are the sole responsibility of Duke University and do not necessarily reflect the views of USAID or the United States Government.

The lessons shared throughout this paper are most applicable for nonprofit or hybrid social enterprises that are developing their scaling strategy or are in the process of scaling their impact, as well as for the funders that aim to financially support such enterprises. Though not the primary audience for this paper, for-profit social enterprises and earlier-stage ventures not yet actively scaling can also glean valuable advice.

TABLE OF CONTENTS

Credit: David duChemin for
The BOMA Project



1	Introduction
2	Executive Summary
4	Lay the Foundation: Ensure data efforts drive toward equity and inclusion
8	Set the Building Blocks: Answering the how, what, and who of data for scale <ul style="list-style-type: none"><i>Create a learning culture as you scale</i><i>Decide what data to collect to drive action</i><i>Determine how to collect data with scale in mind</i><i>Build data infrastructure to bolster scale</i><i>Engage the right people to support data goals</i>
25	Power Scale: Using data to drive specific scale strategies <ul style="list-style-type: none"><i>As You Grow: Activate data use at all levels</i><i>As You Partner: Adapt your approach to data</i><i>As You Drive to Systems Change: Shift your conceptions of data</i><i>As You Have More Data Available: Consider advanced methods like machine learning</i>
41	Funder Implications
49	Key Takeaways for Social Enterprises
52	Appendix A: Project Overview
53	Appendix B: Featured Organization Descriptions
57	References

INTRODUCTION

Big data, small data, lean data, data privacy, machine learning. These are just a few of the current data trends that social enterprises² have to comprehend while advancements in data science continue to evolve around them at a rapid pace. And all of these come as we witness an explosion of data from mobile phones, satellites, social media, and so much more. Navigating this ever-evolving data terrain is a universal challenge.

Given new pressures (e.g., funders and partners wanting to see data and how social enterprises are using it, competitors threatening to eclipse them with fancy dashboards and advanced analytics) and opportunities, social enterprises are hiring data experts and investing critical resources into data strategies and infrastructure. **At best, these investments can help drive better decision-making and increase the pace of scale. At worst, they can waste time and money, amplify biases, and lead to poor decisions and missed opportunities.**

Unfortunately, it seems like the field is not yet stepping up to this challenge. According to a survey conducted by Kathleen Kelly Janus, 75 percent of organizations reported that they collect data, yet only six percent felt that they were using that data effectively.³ In our decades of experience working with and studying social entrepreneurs, we have seen both nonprofit and for-profit organizations become hindered by too much or too little data, underinvest in the infrastructure—people, culture, and platforms—needed to support data, and not take advantage of data as a critical lever to navigate roadblocks and pivots on the path to scale.

Luckily, the field now includes social enterprises that have been driving to scale for decades and have learned valuable lessons about how data can accelerate that path. We are thrilled to be able to share many of their hard-won insights, advice, and stories with you through this work.

Data is a vast topic, so we do not intend to cover it all in this report. We focus on data that drives operational decisions, not data that is created to respond to funders or for the sake of academic study. Furthermore, we cover the topics that a sample of leading social enterprises (see Appendix A: Project Overview) felt were most critical as they reflected on their scaling journey. The recommendations and advice these leaders provide are sophisticated and timely, but you will not need a PhD or advanced data degree to follow along.

While the lessons herein are relevant for many audiences, this paper is targeted primarily at social enterprises, with a separate section with takeaways for funders (see page 44). We expect that your enterprises have laid the groundwork by collecting data to help validate your business model and impact, and now you want to take data to the next level. Read on to learn how you can leverage data to power scale.

Organizations and funders always want to go for the new shiny tools. But my advice would be: don't start the conversation focused on machine learning and artificial intelligence. Instead, start by asking 'what is the question that we want to answer?' Then determine what data you have and what the right methodology and tools are to move us toward an answer.

SHARMI SURIANARAIN
Chief Impact Officer, Harambee Youth
Employment Accelerator¹

EXECUTIVE SUMMARY

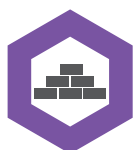
Drawing on the perspectives and experience of some of the world's leading social enterprises, this paper lays out key strategies and advice on how to use data to more effectively and efficiently scale impact. Our interviewees' advice was three-fold: 1) to lay the foundation for data efforts by carefully considering equity and client voice; 2) to set the data building blocks—the how, what, and who of data—that will be critical for driving scale; and, with all that in place, 3) to pursue more advanced data approaches that align with the specific scaling strategies that you are pursuing.



LAY THE FOUNDATION

Ensuring data efforts drive toward equity and inclusion

In order to unlock the potential of data to scale impact, your approach must be grounded in equity, ensuring that all stakeholders have a voice and that you are not scaling bias as you go. This means engaging stakeholders in determining what gets measured, identifying biases in how data is collected, and empowering diverse stakeholders to access, interpret, and act on the data.



SET THE BUILDING BLOCKS

Answering the how, what, and who of data for scale

Once the equitable foundation is laid, there are critical building blocks—the how, what, and who of data—that will power your drive to impact at scale. These two-page memos will give you the top tips to get it right:

- 1. Create a learning culture as you scale.** The road to scale is rarely linear, therefore, enterprises must create a learning culture that uses data to adapt along the way. This means shifting the purpose of data from accountability to learning; routinely interrogating findings; ensuring “bad data” is learned from (and not disincentivized); and, embracing data-driven experimentation.
- 2. Decide what data to collect to drive action.** The wrong data can stifle a leader's abilities to make decisions—or worse, lead to poor decisions. Therefore, prioritize a limited number of key performance indicators; monitor the tension between impact, cost, and reach; collect quantitative and qualitative data; and identify data that will drive scaling decisions—not merely satisfy funders.
- 3. Determine how to collect data with scale in mind.** More manual or resource-intensive data collection may work well in early stages but will be inefficient as the work scales. Scalable data collection finds the balance of being simple and repeatable while still equitable. This approach often means leveraging technology or other partners to make data collection more efficient.
- 4. Build data infrastructure to bolster scale.** Data infrastructure—particularly the systems that store and communicate data, ranging from excel files to enterprise-wide data systems—can help manage data for scale. Building data infrastructure for scale means starting scrappy while also considering future needs; seeking out dedicated capital; creating platforms that go beyond one organization's needs; and ensuring data privacy and security.
- 5. Engage the right people to support data goals.** Data is only as good as the people that collect, analyze, and maintain it. Deep and broad people strategies may be required (e.g., hiring technical experts and people with experience in different methods to dive deep in to data), while also integrating data broadly across the organization through ongoing trainings and open dialogue.



POWER SCALE:

Using data to drive specific scaling strategies

With the foundation and building blocks in place, more advanced data approaches can be layered on. Depending on the scaling strategy you are pursuing, consider the following approaches. Each memo includes a framework to follow as you use these approaches to superpower your scaling strategy:

- 6. As You Grow: Activate data use at all levels.** As organizations grow, the number of people that can (and should) collect, analyze, and use data increases. Empower these users by designing processes that provide the right data at the right time, often facilitated by technology; incentivizing data collection and communication through tailored dashboards; and ensuring users are equipped to take action (often through regular trainings and automation).
- 7. As You Partner: Adapt your approach to data.** Achieving impact at scale often involves various types of partnership, all of which will inevitably diminish the level of control over what data is collected and how it is collected. Therefore, approaches to data must adapt for partnerships: working to set a shared intention for data learning and improvement; aligning on what to collect and how; and promoting streamlining and simplicity in shared data efforts.
- 8. As You Drive to Systems Change: Shift your conceptions of data.** Systems change efforts require a different data approach. Track data related to critical systems change levers, such as FSG's Six Conditions for Systems Change: policies, practices, resource flows, relationships, power dynamics, and mental models. How that data is measured must also change: prioritizing contribution over attribution; being bold in selecting long-term measures; working with and through partners; and going beyond counting to identify meaningful outcomes.
- 9. As You Have More Data Available: Consider advanced data methods like machine learning.** Machine learning can be a powerful tool, but first determine if it is the right tool for the problem. If so, you must have some high quality outcome and reliable predictor data in place. From there, make sure algorithms are not set up to perpetuate inequity and the right people are in place to drive success.

Within each of the sections outlined above, the social enterprises we interviewed share tactics and tips they have applied in their work—or learned through the wisdom of hindsight. This paper also includes key implications for funders (see page 44) about how they can best support social enterprises as they use data to power scale as well as a summary of key takeaways for entrepreneurs (page 48).

As always in our Scaling Pathways series, we acknowledge that this “*Using Data to Power Scale*” theme study is yet another chapter in a story that will continue to be written by the intense experimentation that undergirds the global field of social entrepreneurship. We are honored to share what enterprises have learned and hope this will help guide other social ventures—both nonprofit and for-profit—to achieve greater impact at scale.



LAY THE FOUNDATION

Ensuring data efforts drive toward equity and inclusion

Interested in data for scale? You probably want to talk about fancy dashboards or advanced analytics. But before we get there, our interviewees emphasized how critical it is to first lay a foundation of equitable and inclusive data practices on which to build.

An equitable and inclusive data approach will help answer questions, such as, what are the best practices for listening to and collecting data from clients? If this is done well, you will be able to deliver more value to your clients and make your business model stronger and more sustainable in return. Or, how can we most effectively give voice to the diverse stakeholders who are experts in their own lives? If successful, better solutions will surface and, through data, you can empower and uplift the communities your mission seeks to serve.

While embedding an equity lens in your data is important at all points throughout your development, scale introduces a new pace of growth and reach that has the potential to exacerbate bias and demands a close consideration of equity. Many organizations, although they are faithfully serving their missions, may not have interrogated their data to ensure equity and inclusion. This foundation will help you do just that and will serve as a critical lens through which the remainder of the strategies in this paper can be viewed.

The organizations we interviewed shared three key questions they ask themselves to illuminate possible bias and identify opportunities for improvement in the collection and analysis of data:

1. **Who determines what gets measured?**
2. **How is it measured?**
3. **Who gets to analyze and interpret the data?**

To make it useful for your organization, below we break down each of these questions and provide advice, tactics, and examples of how organizations have approached them.

“Data has even more power than we might recognize for poverty eradication, provided it is shared back with communities. Let’s transform an extractive exercise into an interactive one. Let’s restore to people their fundamental right to own, understand, and control their data and information. Such a matter of moral right is universal and never contingent on economic status.... To be ethically defensible, data must be gathered through transparent social understanding and relationships.”

ANN COTTON
Founder, CAMFED⁴

What do we mean by Equity and Inclusion?

While many definitions exist, for the purposes of this paper, we are focusing on the following:

Equity: the “fair treatment, access, opportunity, and advancement for all people, while at the same time striving to identify and eliminate barriers that have prevented the full participation of some groups.”⁵

Inclusion: an environment which supports, serves, and values all individuals such that they can fully participate and that their perspectives are used to inform action.

This memo focuses on what U.S. Monitor describes as “two important strands that are interconnected but not identical: a focus on [client]⁶ voice and an emphasis on diversity, equity, and inclusion.”⁷ Like U.S. Monitor, we see these two “strands” support and reinforce each other in the context of data to drive impact at scale, especially with the populations with which social enterprises work, and thus include them both in the points below.

QUESTION 1:

Who Determines What Gets Measured?

Too often, social enterprises collect data about their clients without consulting them in the process or even truly understanding what is important to them.

TOP TIPS & TACTICS:

- **Ensure your data is keeping you accountable to your clients** **Tactic:** Take stock. Go through your existing metrics and ensure that the things you are measuring—and thus that are driving decisions—are holding you accountable not just to yourself and your funders, but also to your clients.

Harambee Youth Employment Accelerator (Harambee) recognized the tendency to create metrics that merely hold the organization accountable to itself and its funders for activities delivered. Thus, instead of creating measures such as “have had x conversations with a young person on phone,” which center on the organization, it collects data that holds itself accountable to the clients that the mission is intended to serve: youth that are currently unemployed. Data centers the young person, even being written in their voice, such as “I [young person] am supported in the network” and “I [young person] secure a work opportunity.”

- **Provide opportunities for your clients to help define what data points matter** **Tactic:** Bring M&E closer to the front-line, ensuring closer ties to staff or stakeholders who can help to amplify the voice, concerns, and interests of the ultimate client and who may be best positioned to identify ways to directly engage clients for their input.

“Poor people best understand their own lives. Let’s listen hard and learn from them as to how, together, we can solve the colossal and growing problems facing us all,” says Cotton of CAMFED. For example, Health Leads has leaned into a listening approach in recent years. Initially it collected data on the wrap-around resources that its patients accessed (i.e., food pantry)—but, in digging deeper, recognized that, from the patients’ point of view, “wellness” is not so narrowly defined. Health Leads has begun to experiment with measurement approaches that are embedded in a more holistic inquiry into whether patients have the essential resources they need to be healthy, thus holding itself accountable to wellness as defined by clients.

QUESTION 2:

How is it Measured?

The methods employed to collect data can easily perpetuate inequity and bias if not thoroughly interrogated.

TOP TIPS & TACTICS:

- **Identify biases in data collection** **Tactic:** Test data questions with different constituent groups to ensure questions account for different life experiences, interpretations, and cultural nuances.

Chicago Beyond’s Equity Series shares examples of how inequity can be built into metrics, such as “measuring housing ‘overcrowding’ for participants from a culture that values extended family,” or “measuring ‘progress’ on self-actualization for participants from a culture that prizes interdependent families over independent individuals.”⁸ In another example, one interviewee reported discovering that a question in a key pre-program survey referenced a male-dominated activity, and thus the results favored men over women.

-
-
- **Ensure that all voices are captured**
-
-

Tactic: Use multiple collection methods (quantitative and qualitative) to capture different voices. Make sure these methods do not leave anyone out by asking whether some groups will have difficulty accessing technology, feel less comfortable with format, or are unavailable during data collection.

Harambee uses multiple methods to gain insights from the youth it serves, acknowledging that while some youth may be available in person for a focus group, others will be available only by phone at certain times of day or certain days of the week. Health Leads has an individual with expertise in community-based participatory research and qualitative methods on staff, which is core to its ability to advocate for alternatives to more traditional quantitative research and to challenge existing narratives. As we discuss in more detail in the systems change section (see page 33), qualitative data methods (e.g., outcome harvesting, comparative case studies, etc.) are as important and rigorous as quantitative and offer alternative means of capturing stories of different stakeholders.

-
-
- **Dismantle power dynamics that can bias data collection**
-

Tactic: Interrogate your data collection methods to identify power imbalances: Who is collecting the data? How could that be perceived by those from whom data is being collected?

For Harambee’s on-the-phone experience (HOPE) calls—one of its primary data-gathering tools and key touch points with youth—interviews are conducted by young work-seekers themselves. Data about youth are gathered by youth, placing youth at the center of data collection and product design, and decreasing the likelihood that those being interviewed will feel uncomfortable with the interviewee.

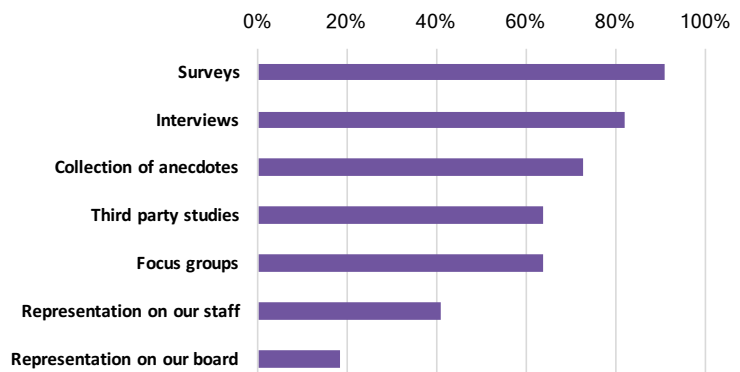
-
-
- **Think about what biases your data collection processes might be hiding from view**
-

Tactic: Work with the program and data teams to articulate the stakeholders that you are seeking to serve and ensure that they are being included in data collection and/or create a plan to address that.

When Harambee began to use machine learning to suggest work opportunities to youth, it used parameters that included cost/time/distance to work location (shorter distance to work was shown to correlate with retention). However, Harambee realized that youth living further away from economic centers would be left out using this data, as they were further from potential jobs; therefore, different interventions—that made use of work-seeker location and locally relevant market intelligence—were required. Learn more about machine learning and how it can perpetuate equity on page 42.

How do you formally capture your client’s voice as a part of your dataset?

Scaling Pathways Data Survey⁹ respondents selected from a list of practices. Of note, few reported client representation on their staff or boards.



QUESTION 3:

Who Gets to Analyze and Interpret the Data?

The individuals and groups who are the subjects of data collection should have the opportunity to actively use and interpret the data, and their voice and experience should be brought to bear in decision-making stemming from that data.

TOP TIPS & TACTICS:

- **Empower stakeholders by providing access to data** **Tactic:** Work with stakeholder groups to understand what data they would like to have, when, and in what format. Communicate and share the data in ways that are relevant and easily understood by impacted communities.

The Family Independence Initiative, an organization that supports families to move out of poverty, not only collects data from families to identify trends and make relevant resources available to those families and communities, but also compiles and analyzes that data to share back with the families. This sharing of data enables families to track their own progress and better understand the trends in their communities. CAMFED finds that sharing data creates opportunities to celebrate progress and reinvigorate community efforts. Theresia Moyo, Head of Education at CAMFED Tanzania, reflects, “By sharing the results of community action at local level—for example through graphs on the increased recruitment of qualified female teachers as a result of a community initiative to build teacher housing—young women help communities celebrate their achievements, and aim higher.”¹⁰

- **Give diverse stakeholders a voice in interpreting the data** **Tactic:** Slow down to engage people at all levels of your organization and of the impacted community to develop ideas about what the data could mean and to bring these stakeholders’ unique perspectives to the analysis.

To ensure its clients have a voice in interpreting the data, Harambee regularly convenes focus groups of members of its target demographic and leverages its regular communications with youth through call center conversations to illuminate additional insights from the data.

- **Commit to taking action on client insights** **Tactic:** At an organization level, measure and track the integration of client input and insight into decision-making to hold the organization accountable for not only collecting the data but using it.¹¹

Damon Francis, Chief Medical Officer of Health Leads, warns that merely providing an opportunity for and tracking client insight is insufficient if it does not lead to shared decision-making power. “People get very tired of giving advice that isn’t followed. A pathway to increased authority via participation in governance is really important.” For example, Federally Qualified Health Centers have long been required to have governing boards in which patients constitute a majority of members, and this design ensures that insights from those most affected by inequities are closely connected to decisions about budgets and organizational strategy.

- **Present data in ways that are accessible for all** **Tactic:** Find out what familiarity different groups have with various presentations of data, work with them to allow unbiased, productive access, and test formats and channels with target audiences.

Chicago Beyond shares considerations for those making data more accessible, such as ensuring the language used is easy to understand (and testing that in early report drafts), working with those who will be consuming the data to determine the best format for its presentation, and determining which forums are best for sharing data (e.g., social media, community discussion).



SET THE BUILDING BLOCKS

Answering the how, what, and who of data for scale

Once the equitable foundation is laid, there are critical building blocks—the how, what, and who of data—that will power your drive to scale. These two-page memos will give you the top tips to get it right. First, you'll need to ensure that you have created a learning culture within your organization that uses data to improve and inform the pivots that often occur on the path to scale. You'll then need to build out your data with scale in mind—collecting the right kind of data to guide scaling decisions and crafting data collection processes that are simple and repeatable. To effectively leverage that data, you will need to invest in creating the data infrastructure—both platforms and people—that can power scale.

BUILDING BLOCK	KEY QUESTION	PAGE #
Create a learning culture as you scale	<i>How can we create a culture that drives learning and experimentation and not just accountability?</i>	9
Decide what data to collect to drive action	<i>How do I collect just enough of the right kind of data to power scale?</i>	12
Determine how to collect data with scale in mind	<i>What simple and repeatable processes do I need to have in place to enable data collection on a larger scale?</i>	15
Build data infrastructure to bolster scale	<i>What systems do I need in place to manage and maximize my data?</i>	17
Engage the right people to support data goals	<i>How can you best resource your data strategy with deep expertise and broad buy-in and engagement?</i>	21



SET THE BUILDING BLOCKS

CREATE A LEARNING CULTURE AS YOU SCALE

IN SUMMARY

As we wrote in *Scaling Pathways' Pivoting to Impact*, "reaching systems change and transformative scale involves disrupting the status quo, which is not a linear process but entails pivots along the way."¹² If you are a leader of a social enterprise wanting to achieve impact at scale, you must create an adaptive learning culture that uses data to question, improve, iterate, and pivot. Without this, staff will likely default to accountability mode—driving toward targets but missing key insights, avoiding experimentation for fear of failure and, most importantly, missing opportunities to achieve greater scale. This learning culture is captured by VisionSpring Board Chair, Reade Fahs, who advised, "Don't fall in love with your 'pretty little solution.' Ask yourself: Is it really working or is it not going to achieve your goals? If not, keep pivoting back to solutions that will be able to achieve impact at scale."¹³



KEY QUESTION

How can we create a culture that drives learning and experimentation and not just accountability?

IN ACTION: Health Leads, shifting from a march toward targets to learning from data

After focusing for a number of years on meeting the targets for its operational and performance metrics, Health Leads made an intentional shift in its data approach to fuel its next phase of scaling. A new leader began to ask questions beyond the targets, trying to understand variation within the program and using that to ask questions. For example, why was enrollment so different across cities, and what could they learn from that?

Health Leads created a quality improvement position and conducted a "listening tour" to understand how people in the organization were thinking about program and quality improvement. It realized that the culture often defaulted to "ready, aim, fire," with little room or time to understand and learn from data. Sara Standish, VP for Strategy, Learning, and Impact, shares, "Most of our staff had been exposed to data for accountability—that's the way most of us use data on a day-to-day basis." But, she adds, "Using data for accountability makes it difficult to authentically ask why something is the way it is." Recognizing that in order to expedite impact at scale it was necessary to surface the root causes of challenges versus defaulting to the easiest answer, Health Leads leadership worked to slow down the ready-aim-fire culture and make room for bringing others to the table (including front-line workers) to understand context around the data and position themselves to learn from it. As Standish explained, "It is important to be sure you are presenting many different storylines that could explain the data versus the one that just validates your current thinking." For example, early on in the Collaborative to Advance Social Health Integration, many improvement teams from primary care clinics were reporting unusually high scores on measures of health confidence among their patients (i.e., "How confident are you that you can manage or control your health problems?"). An accountability approach would celebrate this as an achievement and move on, but the learning approach uncovered that patients might not feel comfortable reporting their true health confidence when first getting to know community health workers;

some feared that low scores might invite coercive or punitive actions of some sort. Understanding this required creating physical and virtual space for frontline staff to safely share data and stories with one another without fear that it would lead to loss of funding or low performance ratings. In implementing such learning approaches, Health Leads has had to accept that the answers to some of these questions could require significant shifts in strategy—which have since occurred—creating sometimes major pivots but keeping the organization scaling effectively.

TOP TIPS

1.1

Ensure your data is keeping you accountable to your clients



To effectively use data to scale impact, a shift in the purpose of data from accountability (e.g., *are we hitting our target number of people served?*) to learning (e.g., *why are we seeing fewer women engage in our program?*) is critical. As leaders did in the Health Leads example above, ask yourself: “Are the questions we are asking and data we are collecting driving only toward accountability or toward learning and improvement?”

1.2

Routinely interrogate your data to move beyond convenient narratives



While new insights can be disruptive to a program and require slowing down to create space for discourse, it is critical to routinely interrogate your data to push beyond existing or convenient narratives. As Harambee expanded into Rwanda, the data indicated that the program was meeting all targets and thus was successful. When interrogating the data further, and questioning whom the organization was serving, Harambee realized that it had met its targets by serving youth with fewer gaps to fill, versus the more vulnerable and hard-to-place youth it also wanted to serve. By interrogating the seemingly positive data further, Harambee was able to tweak and amplify its scale efforts to reach a broader population of youth. For Pratham, routinely interrogating data involves monthly review meetings where regional teams share field experience and data, synthesize data into common themes, and develop action plans to address any concerns.

1.3

Model learning from “bad results”



Individuals and teams can be disincentivized to share data reflecting less than stellar performance if they fear punitive action. This reluctance stifles learning and improvement. Last Mile Health, an NGO that partners with governments to build and sustain community health systems, recognized this dynamic as it brought together county health teams, national government representatives, donors, and partners to regularly review data on the national community health worker program in Liberia. “Understandably, people are worried about being penalized for ‘bad data,’ since this type of data has traditionally been used as an audit mechanism. People of course fear having their funding and/or autonomy taken away,” shares Last Mile Health President and COO Lisha McCormick. To ensure that the data was used for quality improvement, Last Mile Health, the Liberia Ministry of Health, and partners “encouraged county health teams to showcase their challenges, as it would not be held as a mark against them,” as McCormick said. To read more about Last Mile Health and other best practices around using data in government partnerships, see page 30.

1.4 Embrace continued experimentation and the data that guides you

-
-
-
-
-
-
-
-

Knowing that scaling will require continued experimentation and pivots, data and a learning mindset become critical to guide the journey. Each year, One Acre Fund runs more than 40 experiments to test scaling approaches through its Scale Innovations Team. The data from these experiments is used to assess whether a new innovation—such as increasing the farmer to field officer ratio—can support One Acre Fund’s scaling ambitions by either increasing impact per client or increasing the number of customers One Acre Fund can cost-effectively serve.¹⁴



DECIDE WHAT DATA TO COLLECT TO DRIVE ACTION

IN SUMMARY

Given the ubiquity of ever cheaper data collection methods, there is a tendency for organizations to collect, and funders to require, more and more data. However, data is most powerful when it is able to support decisions or actions that drive scale. VisionSpring and EYelliance Co-Founder Jordan Kassalow shares, “You are usually only really addressing the top one to three pain points. Take care of the hearts and lungs, and you can live with an achy knee. There are always trade-offs.” Too much data or the wrong kind of data can stifle a leader’s abilities to make decisions—or worse, lead to poor decisions because the data collected drives the wrong kinds of actions as you scale.



KEY QUESTION

How do I collect just enough of the right kind of data to power scale?

IN ACTION: Harambee, always asking “how will this data serve us?”

Harambee’s leadership starts every Monday with a weekly stand-up call during which senior leadership reviews a dashboard of the organization’s key performance indicators (KPIs). The call is an opportunity to have a conversation about where the organization is today, where it is heading, and what priorities are on the radar. In order to ensure these calls drive the right kind of action, Harambee knows it must be reviewing the right KPIs that keep it laser-focused on the youth it is serving. When deciding which KPIs are important to collect, Harambee considers two things—the first is the number of KPIs it is collecting and the second is the alignment of these KPIs with its mission. For Harambee, the magic number of KPIs to collect is under 10. Having a small number of KPIs allows it to focus on these data points, assign ownership, track the KPIs regularly through an organization-wide dashboard (for more on dashboards, see page 29), and use those KPIs to drive further inquiry and ultimately important strategic decisions. It has the discipline required to move from collecting data that is “nice to know” to collecting data that “will drive action.” Harambee asks itself three questions when it chooses to collect new data:

- 1. Why are we tracking that?**
- 2. What behavior will it drive internally?**
- 3. Does it keep the young person (i.e., Harambee’s client) at the center?**

With these three questions, Harambee is able to prioritize and clarify what is useful to collect versus what is more likely to sit on a shelf.

TOP TIPS

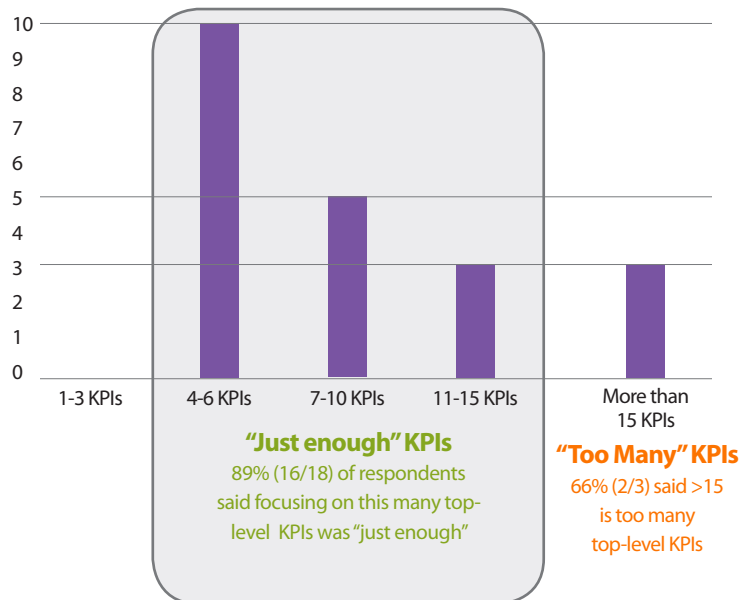
2.1

Limit the number of key performance indicators (KPIs) and prioritize those that will drive scaling decisions

Bart Houlihan, Co-Founder of B Lab, acknowledges, “Data is incredibly powerful to a point. It can also paralyze an organization.” Houlihan encourages enterprises to “**get good enough and go.**” And Asif Akram, CTO of Living Goods, emphasizes the need to understand what data is of highest priority, encouraging enterprises to ask themselves, “Are we a data surveyor, or are we providing a core service?” For Living Goods that meant collecting only data that related to its core service—providing health care to pregnant women and children under the age of five—not collecting additional health or demographic data that the government or funders were interested to know about the population Living Goods is reaching.

Survey Results: Somewhere between 4-15 is “just enough” KPIs

Three-quarters of respondents reported regularly reviewing 4-15 KPIs at the executive team level, which they largely felt was “just the right number” of KPIs.



2.2

Identify data points that will help manage unique tensions of scale

While each organization has unique data points based on its own theory of change, it is also important to track the critical drivers of scaling efficiency: how quality of impact and cost per unit change as the work scales. Emily Bancroft, President of VillageReach, often speaks about the tension between impact, financial sustainability, and scale, and how organizations must often make trade-offs within this “three-legged stool.” She says that, with enough money, an organization can increase its reach and achieve high quality impact—but it is challenging to maintain that level of financial input over time. Or, it could focus on tightening costs while broadening reach and allow for some dips in impact outcomes—a decision that sometimes makes sense for the short-term while scaling. For the purposes of data, then, it is clear that organizations must closely monitor the balance between costs, reach, and quality of impact—and put in place trip wires (i.e., data that can serve as early warning signals) to determine when the desired balance is off track.

2.3

Recognize that data that drives scale is both quantitative and qualitative



There is a natural tendency to provide more weight to quantitative data gathered through the scaling process. But interviewees stressed repeatedly that qualitative data was just as important and helped them ensure that the context of quantitative data was well understood and that client experiences were centered in an equity-driven path to scale. For those reasons, it is essential to balance both quantitative and qualitative data collection and analysis. Our interviewees used focus groups and interviews to develop and test hypotheses, leveraged stories to provide more nuanced context, and ensured that qualitative data was helping to uncover biases, trends, and experiences that add depth to analysis. Qualitative data approaches are well studied and offer rigorous approaches to data use for scale.¹⁵

2.4

Advocate for the data that you have shown makes sense for your organization



“Many organizations approach data in a fractured, reactive way—often driven by data that is needed to answer a question for a funder. Instead, organizations should think proactively about the data that they need to drive scaling decisions and the types of data that they can collect in scalable ways,” says Bob Filbin, Co-Founder & Chief Data Scientist of Crisis Text Line.* Filbin’s comment reflects VisionSpring’s experience as well. In VisionSpring’s early years, it worked to quickly respond to its donors’ every data request. But as VisionSpring scaled and developed a deeper understanding of its impact, it tried to compel its donors to accept the metrics that were critical to VisionSpring for managing its business. As VisionSpring Co-Founder Kassalow explains, “We tried to say, ‘this is data that is important for our business. We assume you don’t want us to create data that’s only for you?’” Kassalow shared that once its donors saw that it was asking the right questions and collecting the right data, the vast majority of them agreed not to add on top of that.

**The mentions of Crisis Text Line throughout this paper are based on interviews the co-authors conducted with Crisis Text Line staff Jaclyn Weiser, Lead Data Scientist, and Bob Filbin, Chief Data Scientist, in January 2020. In June 2020, the Crisis Text Line Board of Directors terminated Co-Founder and CEO Nancy Lublin (who was not interviewed for this series) following a series of staff allegations.*

3

DETERMINE HOW TO COLLECT DATA WITH SCALE IN MIND

IN SUMMARY

As social enterprises ready for scale, they must consider the types of data they will be able to collect through repeatable processes that can be cost-effectively and efficiently implemented by staff, partners, or others. Without this focus, more manual or resource-intensive data collection processes that may have worked well in pilot stages will slow an enterprise down and require unsustainable resources (whether financial, human capital, or otherwise) as the work scales. Finding the right balance of scalable yet robust data collection is critical for scaling success, as Dichter, et al. emphasize in “The Power of Lean Data”: organizations should move toward “use of methods and technologies for data collection that favor efficiency and speed while maintaining rigor.”¹⁶



KEY QUESTION

What simple and repeatable processes do I need to have in place to enable data collection on a larger scale?

IN ACTION: Pratham, using simple, repeatable, and timely data collection

With an M&E team of more than 70 people in India alone, one might think that Pratham’s data approach is highly complex. On the contrary, because Pratham has so many people involved in data collection, it has strongly prioritized a data approach that is simple, easily repeatable by multiple stakeholders, and scalable. For example, one of Pratham’s core programs, Teaching at the Right Level (TaRL), involves simple learning assessments that are quickly and easily used by teachers to place students in appropriate learning groups and track progress throughout; this same data is also aggregated at various geographic levels for teacher monitors and education officials to use. For the teachers, the simple assessments are highly aligned with accomplishing their work (i.e., in delivering value to students), so it does not feel like an added or irrelevant exercise. While Pratham is directly involved with delivering TaRL through its own learning camps, it has also scaled by partnering with state and district administration to have TaRL delivered by government teachers at government schools. To ensure that the relevant information is collected consistently and in a timely manner across this growing cohort—and to ensure that the other stakeholders have timely access to relevant aggregated data—Pratham developed an Android application for teachers to input information and receive reports.

TOP TIPS

3.1 **Keep it simple & repeatable**

Where possible, data collection processes should be simple and easily repeatable by someone with limited training (as different stakeholders may be collecting the data). For example, MiracleFeet has identified five simple indicators that can be easily collected and reported from the clinics in the field, such as average number of casts per patient (target of 4-6), percentage of patients receiving a specific surgical procedure as part of the treatment process (target >80 percent), and treatment completion rate. These metrics are easy to collect, do not require subjective

interpretation but rather a simple yes/no for surgical procedure and a tracking and count for number of casts and completion, and are easily reported at appointments in real-time through a mobile app used by patient-facing personnel. Living Goods emphasizes the importance of incorporating elements in data collection tools that minimize data entry mistakes—such as setting ranges that a reported number must fall within and limiting open-ended questions.

3.2 Leverage technology to simplify data collection

Many organizations have taken advantage of the prevalence of mobile phones to create phone-enabled templates or apps to ease and standardize data collection. CAMFED turned to a mobile monitoring app in 2012 when “paper-based data collection [had become] less and less viable” due to its scale. The primary challenges mobile collection addressed for CAMFED were reducing time to receive data and ensuring data did not get lost. The tool was able to replace paper records which had to be sent from rural schools to CAMFED offices. The mobile app also makes the step of inputting data (from paper records) redundant. Efforts can instead be focused towards unlocking data insights. Find more on leveraging technology to activate data collection at all levels of the organization on pages 26-29.

3.3 Make data collection a passive activity

Rather than collecting data only through surveys which require active engagement from both its team and texters, Crisis Text Line tries to find data that is inexpensive and easy to collect, tied to existing customer touchpoints, and automated, when possible. In practice, this means that Crisis Text Line defaults to automated data collection, including text messages with metatags (e.g., time stamps, assigned risk level). Other organizations can embed data collection into their customer touchpoints, such as collecting data while onboarding a customer or automatically tracking purchasing behavior. Leveraging technology is one key way that data can be collected more passively.

3.4 Reduce costs by leveraging existing resources to support data collection

Rather than bringing in M&E staff or external enumerators, CAMFED relies on its volunteers—young women in the community—to act as its data collectors. Not only is this cost effective, CAMFED also feels it is valuable as volunteers have significant trust of the community and are able to immediately share back insights with the community through their data collection.

3.5 Assess collection methods & tools for equity

The above recommendations must take into account an equity lens which often means striking a balance between scalable processes and nuanced approaches to address systemic inequities. For example, simplification shouldn't mean there are no opportunities to capture client voice directly through more qualitative methods. Attention must also be paid to any power imbalances inherent in data collection processes—for example a teacher asking for assessment of their teaching from students. For additional tips on assessing collection methods and tools using an equity lens refer to the Lay the Foundation section on pages 4-7.

4

BUILD DATA INFRASTRUCTURE TO BOLSTER SCALE

IN SUMMARY

As organizations begin to scale their impact, they often find that formalizing and bolstering their data infrastructure is required to power scale—particularly the systems that store and communicate data for operational and performance management, which can range from excel files to cloud-based, enterprise-wide data management systems.



KEY QUESTION

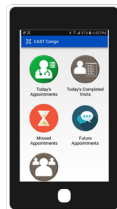
What systems do I need in place to manage and maximize my data?

IN ACTION: MiracleFeet, building a data platform to propel scale

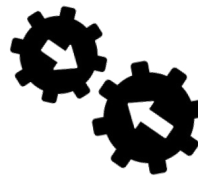
MiracleFeet works with a network of over 500 local partners and treatment providers to supply the technical and financial resources needed to create and support high-quality clubfoot clinics. After working for six years with a legacy data management system, MiracleFeet knew what its data requirements were and that it needed to develop a more efficient and responsive system to support scale. MiracleFeet emphasized that having sufficient implementation and field experience before making such a major investment was critical: “If we had developed our data system even a year or two earlier, it would have been a disaster,” said Chesca Colloredo-Mansfeld, Executive Director of MiracleFeet. “We needed to have enough diversity of experience in implementing our program—in different country contexts and facing different logistical challenges—to be able to design a system that would meet our needs.” MiracleFeet was also fortunate to find a dedicated source of capital to create a system that could propel scale, when it won a Google Impact Challenge Award. With the \$1 million award, MiracleFeet undertook a process to work with other key clubfoot treatment programs globally to ensure it designed a system that could meet all of their needs as well. After two years of development and testing, MiracleFeet rolled out its CAST mobile data application, integrated into Salesforce for real time monitoring and reporting, as shown below.



1. CommCareHQ: Access management and phone monitoring. Online platform to manage CAST app deployment, including creating and monitoring users and accessing data for cleaning as necessary. Managed by HQ.



2. CAST Mobile App: Core data collection platform. Mobile app to guide clubfoot treatment and collect patient data at clinics; includes automated SMS and call features for patient communication. Data input at clinic level and supported by regional Tech Coordinators (limiting burden on HQ).



3. Open Function: Integration system. Automates the flow of data between CommCare and Salesforce in real time. Managed by HQ.



4. Salesforce: Data management and real-time reporting. Platform available online and via mobile app to access individual patient and visit records and see aggregate key performance indicators at the clinic, country, and global levels. Includes standardized dashboards as well as custom reporting at any level (e.g., patient, clinic, county, global). Managed by HQ; used at all levels.



TOP TIPS

4.1 Start scrappy and simple

Survey respondents echoed Harambee's first non-negotiable with **55 percent reporting that they would have made their data systems less complex if they could do it again.**

Enterprises are often tempted to build out complicated, best in class, data systems early in their journey. But as noted in the MiracleFeet story above, interviewees stressed the importance of taking the time—often several years implementing and refining a solution in different contexts—to truly understand data system requirements before making a major system investment. Interviewees reported using basic data collection and analysis tools (often Excel) or inherited legacy systems for the first five to six years of implementation and, once they did start to build a new system, warned not to overbuild or be seduced by flashy technology. Asif Akram, Living Goods' Chief Technology Officer, advises, "start simple and make sure you are meeting immediate requirements. Don't get tempted by what others are doing or try to meet future requirements too soon. But keep improving as you go, building on experience in more of a snowball effect, and do what you can to set yourself up for the future." Similarly, Harambee shares its four non-negotiables when it comes to building out data-related technology.

Harambee's Four Non-Negotiables for Data Technology

1. **Simplicity**
2. **Data integrity**
3. **Enabling for the organization (i.e., helps us to take action)**
4. **Maintainability (i.e., we need to be able to fix broken things quickly)**

Data Platform Hindsight

So, your data platform doesn't really meet your needs. What would you have done differently?

The Scaling Pathways survey respondents who reported that their current data systems did not fully meet their needs told us what they would have done differently.



People: Ensured we had the right skills and bandwidth on our team to use and manage the system.



Less Complex: Made the system less complex.



Future Needs: Spent more time considering our future data needs.



Buy-In: Put more effort into getting buy-in from the data system users.



New Capital: Put in more capital (that we did not have at the time but needed).



Outgrown: We've just outgrown our current system.



Existing Capital: Put in more capital (that we had but did not use).



Nothing: We could not have anticipated the strategy and/or data need changes that have made the system less relevant.

4.2 Spend time considering future data needs

While Akram and others advise keeping a data system simple, they also acknowledge the importance of “setting yourself up for the future.” While you will never be able to anticipate all future needs, by spending time considering them you may be able to make your data infrastructure more flexible for the future. For example, building a data platform that could accommodate different alphabets in the future if you plan to expand globally; having storage capacity that could flex over time to accommodate larger amounts of data; or selecting widely-used technology platforms to help keep systems relevant over time. As Colloredo-Mansfeld of MiracleFeet explained, “Our choice to use established platforms [such as Salesforce and CommCare] has been key to our ability to grow and be sustainable. These companies already have a vested interest in critical areas such as data security and privacy. And choosing a system with a growing network of users and user forums has allowed us to access knowledge about the platform from so many others using it.” **Forty-five percent of survey respondents reported that, had they spent more time considering future data needs during development, their current platform could have been more useful.**

4.3 Seek out a dedicated source of capital...ideally

Without a dedicated infusion of capital for data and technology, many interviewees reported that they would not have taken the resources away from programs to invest in data infrastructure and therefore would make incremental changes over time. Those with dedicated capital (e.g., Google Impact Challenge Award for MiracleFeet, and Salesforce Force for Change Grant for BOMA Project) have been able to create platforms that they report radically changed their ability to drive impact at scale. So, even if you do not have the resources now, think about what an investment in data infrastructure could help you do, articulate your requirements and costs, and be prepared to pitch current and new funders should the opportunity arise. **According to survey respondents, if they were given a dedicated infusion of capital to develop and test a data system to meet their needs, 67 percent would spend USD\$250,000 or less on a data system (broken down as 29 percent spending less than USD\$100,000 and 38 percent spending USD\$100,000-250,000). Twenty-four percent reported that they would spend USD\$250,000-500,000, and nine percent would reportedly spend \$750,000-1,000,000).**

4.4 Build infrastructure that can be applied to the problem, not just the organization

With the goal of driving impact at scale, especially as it relates to systems change (see page 33 for more), organizations should consider how others working on similar issues might be able to leverage the platform and data being collected. With its Google Impact Challenge Award, MiracleFeet was able to create a system that worked not just for itself but for anyone working in clubfoot. The process was beneficial as it forced MiracleFeet to identify common ground among countries and among organizations and distill the most important needs and requirements. While it was more complicated and time intensive, it resulted in a product with greater potential impact and partners with more faith in the platform and in MiracleFeet as a collaborator. (On the flip side, you should consider how you can leverage any systems that have already been developed by others, if they exist.)

What about data privacy and security?

As organizations scale and build out data infrastructure, data privacy and security must be central considerations. At a minimum, this means ensuring your organization understands and complies with the legal requirements around data privacy (standards such as the EU's General Data Protection Regulation), has a privacy policy, and brings in the right expertise around data security. For useful information on legal and policy considerations, including standards, see USAID's Considerations for Using Data Responsibly Guide.¹⁷ Responsible Data, a grassroots network of social impact organizations, implores organizations to go beyond these legal requirements and take a rights-based approach to the use of data—which entails centering equity considerations, including people's right to consent, privacy, ownership, and security of data.

What does this look like in practice? Crisis Text Line chooses to scrub all of its data of any texter identifiers in order to protect its texters' privacy. This leaves them with data that, from an academic standpoint isn't wholly "pure;" but Crisis Text Line's mindset is that "scrubbed data give us 95 percent of the insights and any small improvements you might get from a 'pure' data set is not worth the risk of loss of privacy." Crisis Text Line has an independent Data, Ethics & Research Advisory Board which helps guide its policies around data privacy and security. Responsible Data has also compiled a list of resources to support organizations seeking to follow this approach.¹⁸



SET THE BUILDING BLOCKS

ENGAGE THE RIGHT PEOPLE TO SUPPORT DATA GOALS

IN SUMMARY

As demands for and uses of data continue to evolve, some organizations will hire more data experts whose jobs are focused on data and who can go deep on certain competencies (e.g., data science, M&E, machine learning). Yet, simultaneously, data must be integrated broadly throughout the organization so that all staff are comfortable understanding and using data as the work scales. Ultimately, the exact make-up of a data team is determined by the organization’s unique strategy, model, and infrastructure. The advice below outlines some key considerations in creating that team.



KEY QUESTION

How can you best resource your data strategy with deep expertise and broad buy-in and engagement?

IN ACTION: The BOMA Project’s people behind the data

Data management is housed within BOMA’s M&E department, which is overseen by a dedicated M&E manager and supported by two data analysts. Closely aligned with this department is the core IT team which ensures that the data collection tools and equipment, including the data management system, function optimally. This team also now includes staff with Salesforce expertise to support BOMA’s Salesforce-based data platform. The ultimate responsibility for all programmatic data collection lies with the head of program—currently the Kenya Program Director. Beyond these dedicated data resources, BOMA provides regular data-use training to all program staff and many of BOMA’s people touch the data collection process, as illustrated below:

Dedicated data-related position on your executive team? Not quite yet...

Less than half of survey respondents have a data-related position on their executive team.

Of the 21 respondents, three have an executive level M&E or data position, and seven have an executive level technology position.

BOMA Village Mentor	BOMA Field Officer	Regional M&E Officer	Data Analysts (Head Office)
Responsible for collecting data using the standardized tools	Manages and supports Village Mentors and ensures primary data quality	Provides regional support for monitoring and data quality checks	Oversee data quality assurance and cleaning, analysis, and reporting of data

TOP TIPS

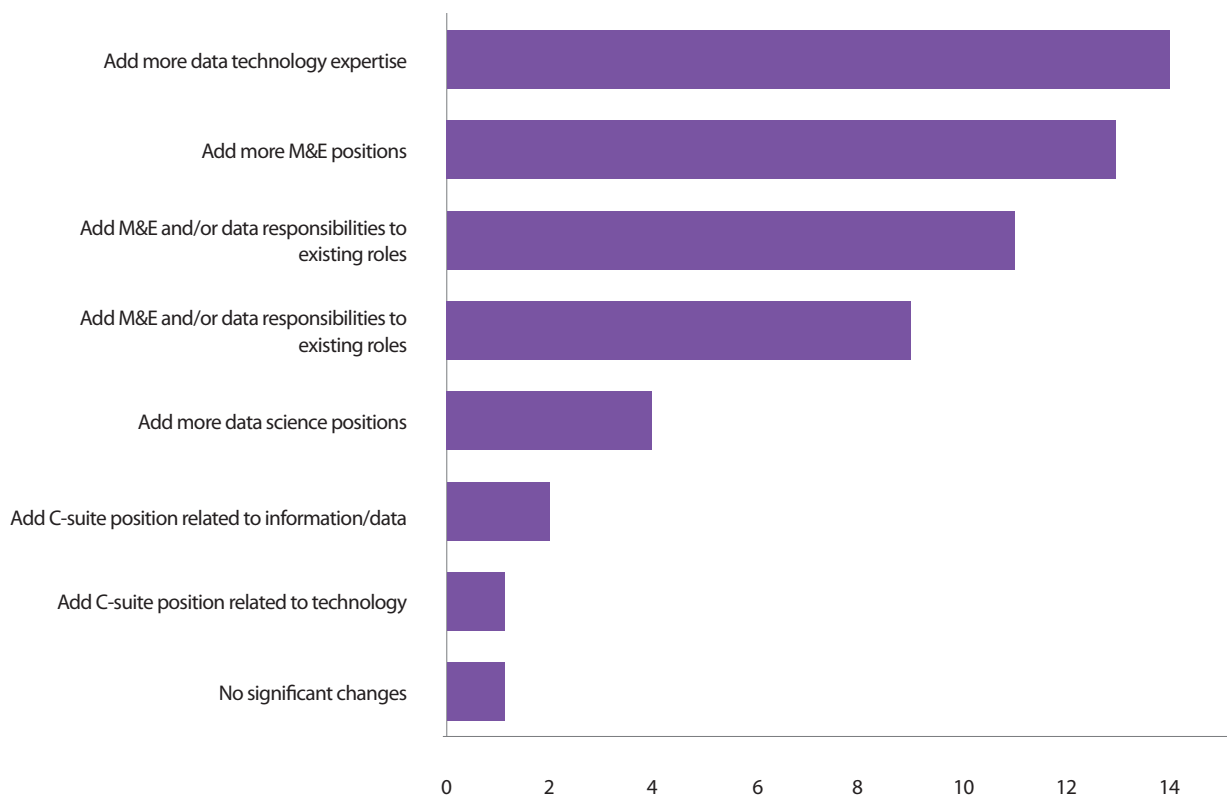
5.1

Use a “dosing strategy” to go broad and deep as needed

As Health Leads has continued to evolve its data infrastructure over time, it has adopted what VP for Strategy, Learning, and Impact, Sara Standish, calls a “dosing strategy” with respect to its internal staff’s data capacity. For example, all staff are exposed to the basics of the continuous improvement frameworks of Model for Improvement and Results Based Accountability, with a handful of staff members being deeper experts. As projects require specific expertise or as complexity increases, Health Leads can deploy those experts to supplement the effort. In our survey, we also asked respondents whether and how they plan to deepen their data expertise in the future, see below for results.

In the Forecast: More data technology and M&E experts

We asked survey respondents about the changes they anticipate in their organizations’ data and M&E capacity over the next five years, and this is what they forecasted:



5.2

Bridge the divide between data and non-data talent

In The Data Lab’s guide to data teams, it stresses that “it is not enough to simply enable the [data] team and leave them to it; the rest of the organization must also adapt to the capability and absorb the new insights into their processes.”¹⁹ To ensure organization-wide buy-in of the data vision and strategy, and demystify the capabilities of data, The Data Lab suggests the organization clearly articulate (and share) the key components of the strategy. *See next page for an additional consideration in integrating data experts within a social sector organization.*

Integrating Data Experts within Social Enterprises

In *Code Switching Across the Social/Digital Divide*, Asha Curran and Julia Rhodes Davis explore generalized differences between data experts and the social sector that could lead to challenges in working together if not intentionally bridged. They write, “Too often, the information (the data, not useful in and of itself) exists on one side, and the people who want to use evidence to make change and drive impact are on the other. The key is identifying the problems that need solving and which questions are answerable.” They share the following table that describes common characteristics of each group, so that social sector organizations ensure they engage “code switchers” and translators who can bridge these divides.²⁰

SOCIAL SECTOR PROFESSIONAL	Collaborative	Independent	DATA SPECIALIST
	Emotional intelligence	Logical mindset	
	Qualitative	Quantitative	
	Passionate	Dispassionate	
	Asks existential questions	Asks answerable questions	
	Complex systems	Linear systems	
	Diligent to morality	Diligent to accuracy	

Source: Asha Curran and Julia Rhodes Davis, *Code Switching Across the Social/Digital Divide*, SSIR 2018.

5.3

Build a team with experience in diverse methods and local context

Given the many and often changing demands on the team leading an organization’s data strategy, our interviewees spoke to some of the experiences that they have found essential. For Living Goods, CTO Asif Akram shared that while there is a temptation to recruit data and technology experts from the West, Living Goods deemed it essential to recruit and invest in local staff (in Uganda and Kenya) who have a better understanding of the local context, live near the work, and can build the required technical competencies. With respect to experience in different methods, Damon Francis, Chief Medical Officer for Health Leads shared, “One thing that has been really helpful for us is having a person with community-based participatory research and qualitative methods (she’s an anthropologist) on staff.” There is also a translation role to play for Francis. “I think my role as a physician with research training has also been helpful. Both of us can be strong allies to front-line staff in technical conversations that drive decisions because we have the language and credibility to advocate in ways that are more influential to people in power who value research-based evidence and often dismiss other types of expertise. Both of us will often illuminate the limitations of research-based and quantitative evidence, and the reasonableness of alternative interpretations based on other types of expertise.”

5.4

Institutionalize regular trainings

As data availability expands and expected use increases, our interviewees spoke about the importance of regular training throughout the organization to maximize ability to use the data. BOMA provides continued needs-based refresher and new function training to its M&E and program staff; “systems champions” within BOMA departments and regions provide decentralized support that can be delivered quickly to end-users. This support structure, alongside a clear problem escalation matrix, helps to ensure that the central IT team can focus its efforts on system optimizations and developments. At the field level, MiracleFeet established regional technical support for users of its CAST app, bringing support much closer to its heaviest users.

5.5 Maintain & evolve your data platform



Most organizations engage external expertise to support the development of a data platform. But once it is up and running, organizations must determine how to resource ongoing maintenance and development. **Our survey respondents reiterated the importance of this step; when we asked those who said their data platforms didn't sufficiently meet their needs what they would have done differently to make it more useful, the top answer (64 percent) was that they would have ensured they had the right skills and bandwidth on their team to use and manage the system.** A number of interviewees reported creating internal positions to administer and continue development of their data systems; this move helps to decrease dependence on expensive external consultants and allows for more flexible, timely, and proactive efforts to evolve these systems which are now integral to program operations and ability to scale. Living Goods partnered early with an external partner (Medic Mobile) to develop its digital tools, but as the tools became a critical part of operations, Living Goods found it had development and customization needs beyond what an external partner could meet on a short timeline. At that point, Living Goods invested in the internal expertise to be able to manage and modify the tools and decreased the external partner's role to focus on platform hosting and maintenance.



POWER SCALE

Using data to drive specific scaling strategies

Once you've set up your data to power your journey to impact at scale, you must consider how your data needs and methods may change depending on your scaling strategy. If you are continuing to expand and grow your product or service, you will likely want to empower your people at all levels to use and extract actionable insights from data. If you are partnering with others, particularly government, you will need to think about adapting your KPIs, systems, and process to ensure their buy-in. If you are pursuing systems change, you will need to shift your expectations of what kind of data you can collect and how. And, as you scale and collect more data, you may want to consider advanced analytics such as machine learning. Each of the memos below provides a framework for you to walk through as you build your data to drive your chosen scaling strategies.

MEMOS	PAGE #
AS YOU GROW: Activate Data Use At All Levels	26
AS YOU PARTNER: Adapt Your Approach to Data	30
AS YOU DRIVE TO SYSTEMS CHANGE: Shift Your Conceptions of Data	33
AS YOU HAVE MORE DATA AVAILABLE: Consider Advanced Data Methods Like Machine Learning	39



POWER SCALE

AS YOU GROW: Activate Data Use At All Levels

For organizations that pursue a growth strategy to scale, the number of active users of data will naturally increase as the organization grows. In addition, while more centralized decision-making is often necessary during earlier stages when a team is trying to validate its model, scaling that model often requires more decentralized decision-making. Therefore, as an organization's work grows, it will need to empower more people at all levels to collect, analyze, and use data to drive to action. But how can organizations ensure that the right data is getting into the right hands at the right time, and that the people using the data are prepared to extract insights and drive action?

"If we can make data an essential tool for frontline workers, rather than a post hoc distraction, we will get more and better data. With that data, we will be able to ask better questions, increasing our efficiency in delivering our products and services."²¹

JIM FRUCHTERMAN
in *"Using Data for Action
and for Impact"*

This was a challenge that Living Goods faced as it scaled its work. As of 2019, Living Goods had over 10,000 active CHWs, serving more than 7.8 million people in Kenya and Uganda.²² Managing this growing footprint required layers of staffing, each with its own data needs. For example, each CHW manages 150 households (with an average of five people per household, so 750 individuals served), supervisors manage 35-50 CHWs, branches manage 200-450 CHWs, and so on through the country level and up to global headquarters. Living Goods knew that its legacy paper-and-pen-based system would not be able to keep pace with this level of growth and so, in partnership with Medic Mobile, it created the Smart Health app and data collection process to collect data on patient health and quality of care, manage CHW performance, and provide dashboards.²³ By carefully designing a system and process that identifies data needs for various stakeholders, leverages technology to speed collection and analysis, and utilizes dashboards to communicate relevant data in real time to all levels of the organization, the Smart Health app allows Living Goods to "exponentially increase the scope of data available to improve performance" and use data to drive action.

The lessons learned from Living Goods and other interviewees showcase critical steps to follow to ensure data is used effectively to empower all levels of the organization as it grows:



DESIGN

1. Identify data needs at each level: Who needs what, and when?

The first step is identifying the specific data that each level of the organization needs to execute its work most effectively. Start by answering the following:

- a. **What programmatic questions does each level need to answer?** For example,
 - Front-line workers often need data to drive actions and identify operational improvements;
 - Supervisors require data to drive performance management;
 - Regional leads need data to understand performance across units/teams;
 - Headquarters leadership use data to see overarching trends and drive strategy.
- b. **How often does it need to be able to answer those questions?**
- c. **What is the simplest and timeliest data to help answer those questions?**

Example: Once MiracleFeet developed and tested a robust model to identify and treat clubfoot through existing facilities in low-resource settings, it recognized the opportunity to decentralize data to drive scale. While MiracleFeet’s central team was already a heavy user of clinic data, the organization made it a priority to push data towards the frontline clinic teams so that it could act quickly to address performance and quality issues. As an initial step, MiracleFeet worked with stakeholders at each level of the organization to understand their key questions and how often they needed the corresponding data. Sample data questions by level are included in the table below:

Individual Clinic (Local Partner)	Country Coordinator (MiracleFeet)	Executive Team (MiracleFeet)
<p>Weekly:</p> <ul style="list-style-type: none"> • What supplies do we need to have ready this week? • How are my patients’ treatments progressing? • Am I taking all of the right treatment steps? <p>Monthly:</p> <ul style="list-style-type: none"> • Am I meeting my quality and enrollment targets? 	<p>Quarterly:</p> <ul style="list-style-type: none"> • Are there any clinics that may need extra support? In what areas? 	<p>Monthly:</p> <ul style="list-style-type: none"> • Are we meeting enrollment targets? • What are the trends we can learn from as we plan future scaling and resource allocation? <p>Quarterly:</p> <ul style="list-style-type: none"> • Are we meeting quality targets? • Are we meeting our donors’ expectations?

2. Consider leveraging data technology solutions to drive down cost and increase quality.

As founder and former CEO of Benetech, Jim Fruchterman, advised wisely in 2013, “Social entrepreneurs also need to embrace technology and data as indispensable parts of our enterprises.”²⁴ Technology can make empowering all levels of the organization with data realistic even in the most remote areas. Many of our interviewees reported using mobile phones and tablets (which feed into larger software programs) to ease and improve data collection at the front line; in the energy and agriculture sectors, social enterprises are also using sensors to automatically collect data that would have previously been measured by hand and self-reported. See the Data Infrastructure Building Block memo (page 17) for additional tips on data systems and technology.

Example: According to Living Goods CTO, Asif Akram, “It was never a question for Living Goods whether we would use digital tools, it was just a question of when.” Living Goods determined when to invest in digital tools by carefully considering the conditions which would drive sustainability and scalability of the tools. For Living Goods that meant investing in the Smart Health app only once mobile phones were prevalent and network coverage reliable in the regions where Living Goods worked. Using technology has allowed Living Goods to eliminate costly and often inaccurate paper-based systems and has increased the amount of data available in real time, shifting “from bicycle-powered delivery of health data to light-speed dissemination.”²⁵ But Akram also cautioned, “Don’t create a problem just because you can solve it. Technology people can think they know more than the user. Sometimes we think we have a solution, but the user is always king. If he/she is not satisfied with the product, nobody will use it.”

COLLECT & SHARE

3. Incentivize staff and partners to collect and use data.

Many of our interviewees reported that their front-line staff and partners were initially hesitant to include or add data collection to their existing responsibilities. To get these staff and partners on board, our interviewees advised that data collected must not become burdensome and must provide value to the data collector. To ensure this, ask yourself:

- a. *Does it save the data collector's time?* MiracleFeet ensures that the data its clinics collect aligns with those clinics' reporting requirements, therefore dramatically decreasing the clinics' time spent on reporting.
- b. *Does it immediately give them information that they need?* Pratham ensures that the data its already-busy teachers collect for Teaching at the Right Level is seen as creating immediate value. The regular student assessments provide teachers with information they can use immediately to adjust that student's learning plan; the assessment data is also aggregated at the classroom, school, regional, and national levels for other stakeholders to act upon. "Data needs to be empowering for people to collect and use it," emphasizes Pratham's Devyani Pershad.
- c. *Does it make their work more efficient?* The BOMA Project's digital platform, Performance Insights, uses the data its front-line workers collect to provide them with immediate feedback so they can take quicker and more informed actions in their work—as opposed to having to go back to supervisors for guidance. For example, BOMA's front-line business mentors can use the data collected about the performance of each business in the BOMA network to inform the types of ventures they suggest that new BOMA program applicants pursue.²⁶

Ensuring data quality as collection expands.

As the responsibility for data collection spreads across many individuals and groups, organizations must double down on data quality measures. For Pratham, whose data is collected by many thousands of classroom teachers, it institutes audit processes and double verification. Teacher mentors regularly conduct sample verifications, retesting small groups of students at a subset of schools and comparing results to the teacher's results for those same students. These verification checks are used as teaching moments for the teacher to improve quality of future assessments so that children are receiving learning at the correct level.

4. Utilize dashboards to effectively communicate data.

Even with the plethora of available data, it would be overkill—and ineffective—to provide data on all metrics to all team members. Interviewees shared the following advice for creating effective dashboards:

- a. *Tailor for use, but keep it simple.* Dashboards tailored for the needs of each user help the user to home in on the metrics that matter most to their workstream and provide a way to quickly identify areas for further interrogation. Akram, Living Goods' CTO, cautioned, "Organizations often think about fancy dashboards, but don't appreciate or realize that people don't always have the ability to interpret or absorb." He went on to explain that when Living Goods designs dashboards, it keeps them relatively simple and thinks about the perspective of the team using it. At each level, data is displayed, interpreted, and utilized differently and dashboards shift from simple at the front lines (Akram again: "People will say that's not a dashboard, it's more just showing information.") to multiple dashboards of increased complexity at the headquarters level.²⁷
- b. *Keep clients central.* Marzanne Collins, Harambee's Chief Information Officer, stated, "What I find valuable in our dashboard is not just what we're counting, but what it says about what counts." Importantly, Harambee asks itself if the metrics included in its dashboards keep young people—its key constituent—at the center of decision-making for the organization. Harambee also ensures that it accompanies the dashboard indicators with stories—making the discussion less about the target itself (e.g., number of phone conversations with young people) and more about the service (e.g., including stories to better understand if these are quality engagements, and what the young people are getting out of these calls).
- c. *Do not forget users without internet or mobile access.* Keep in mind that not all intended users of the data have access to the dashboards and visualizations. While Last Mile Health has deployed digital tools for reporting and data collection in partnership with the Liberia Ministry of Health, it realized that technology and bandwidth constraints would prevent the data from reaching some key stakeholders. To address the gap, Last Mile Health uses paper reports to connect health workers and supervisors in these areas.

Beware the tendency to interpret data visualizations as fact.

Another tool to communicate insights curated from dashboards is data visualization. Common tools such as Excel, Tableau, Power BI, or others create visual elements (charts, diagrams, etc.) that help various stakeholders understand patterns in data. According to Jake Porway, Founder and Executive Director of DataKind, “The best data visualization helps people investigate a topic further, instead of drawing a conclusion for them or persuading them to believe something new.”²⁸ Although these tools can be powerful, Porway also shared the following warning, “The public still treats data and data visualization as ‘fact’ and ‘science.’ I believe the public has gained enough visual literacy to question photojournalists or documentary filmmakers’ motives, aware that theirs is an auteur behind the final piece that intends for us to walk away with their chosen understanding. We have yet to bring that same skepticism to data visualization, though we need to. The result of this illiteracy is that we are less critical of graphs and charts than written arguments because the use of data gives the sense that ‘fact’ or ‘science’ is at work, even if what we’re doing is little more than visually bloviating.”²⁹

TAKE ACTION**5. Ensure data users are equipped to draw actionable insights.**

“You can’t just make a dashboard and expect people to understand how to use the data,” stresses Pratham’s Pershad. Ensuring that data users are prepared to draw actionable insights means that appropriate training processes must be in place to teach users how to understand, interpret, and act on the data provided. Interviewees also advised embedding data discussions and dashboard reviews in regular meetings and allowing for time in those meetings to walk through interpretations and provide additional context behind the data.

Example: A major role of Pratham’s M&E team is to bolster its users’ capacity to draw actionable insights from the data—above and beyond the basics of ensuring that users understand foundational concepts such as median, outlier, etc. During training and review meetings, Pratham shares case studies (based on actual program data) with participants and has participants practice drawing out insights to develop strategies for giving inputs to teachers and supervisors and to create priority plans for mentors. Pratham also works to incorporate data into regular meetings. For example, regional teams conduct monthly review meetings to discuss data trends, insights, and anecdotal experiences. The teams organize the meetings into three parts: sharing field experience and data, synthesizing the data into common themes, and developing an action plan to address any concerns.

6. Build in systems to automate and drive action where possible.

How do you ensure that your data is not a passive asset for an already busy team member, but rather spurs individuals to further inquiry and action? Interviewees talked about building alerts and automated processes into the system to help all levels improve performance.

Examples: BOMA sets thresholds for where it expects to be on each of its key metrics and has programmed its data system to generate an alert if a metric deviates from the threshold. For each alert, BOMA has standard operating procedures for who should first respond and how. MiracleFeet’s CAST app automatically generates a list of appointments at the beginning of the clinic day and a list of no-shows at the end of the clinic day, so the staff can prepare the right medical supplies and can follow up with those who missed appointments. The app also alerts the user if a clinician has missed a treatment step or the treatment is not progressing well, prompting the clinician to reassess or seek help. Living Goods uses patient data to create “Daily Data-Driven Task Lists” that prioritize follow-up actions for CHWs based on the severity of patient diagnoses and daily action lists for supervisors noting issues that need follow-up (e.g., CHWs missing targets or field offices that are low on stock).



POWER SCALE

AS YOU PARTNER: Adapt Your Approach to Data

Achieving impact at scale often involves various types of partnerships, whether it is a more tightly controlled relationship (e.g., contracting with a local organization to implement your program) or loosely controlled (e.g., providing resources to an organization to incorporate aspects of your methodology into its own work). Regardless, your approach to data must adapt as you pursue these partnerships, given an inevitably diminished level of control over what is collected, how it is collected, and the access you are allowed. One key scaling partnership that many social enterprises pursue is with government entities—in fact government partners came up most frequently in our interviews. Therefore while the advice below can apply to most any partnerships, the examples are primarily about government partnerships.

SET A SHARED INTENTION TO USE DATA FOR LEARNING AND IMPROVEMENT.

The Risk

Many partners, particularly those providing funding or other resources, often default to using data for audit or accountability purposes—which can lead to a hesitancy to share data, and thus a stunted opportunity for data-driven improvements.

Hard-won Advice

- **Ensure all partners understand and agree upon which data is for accountability, and which is for learning and improvement.**
- **Establish that partners will not be penalized for the data they share that is for learning and improvement.**
- **Model a culture of learning, particularly from those who others perceive as holding the most power (e.g., funders, government).**

Example: *To continuously monitor and bolster the quality of community health worker programs, Last Mile Health co-designed the Implementation Fidelity initiative (IFI) with Liberia's government at both the national and county levels. County health teams, NGOs, and national government representatives review the IFI data together quarterly as part of a learning agenda rather than as an audit mechanism that penalizes partners for 'bad data.' "Understandably, people are worried about being penalized for 'bad data,' since this type of data has traditionally been used as an audit mechanism. And when you're sharing this data with the national government and donor partners, people of course fear having their funding and/or autonomy taken away," shares Last Mile Health President and COO Lisha McCormick. In order to ensure that the IFI data was used for quality improvement to promote program improvements and collective knowledge sharing, Last Mile Health, the Liberia Ministry of Health, and donor partners "encouraged county health teams to showcase their challenges, as it would not be held as a mark against them," as McCormick said. This agenda is reinforced at quarterly meetings, where county health teams are encouraged to share stories alongside numbers and where Ministry officials model a learning approach.*

ALIGN ON WHAT DATA TO COLLECT AND HOW.

The Risk

While a social enterprise will understandably create metrics that help drive its own work, those metrics may prove meaningless to partners and could get in the way of identifying a shared agenda. Additionally, any data collected without a partner's engagement and buy-in may be viewed skeptically. [See also "Incentivize other Systems Actors to Collect and Use Data" on page 37 for tips on engaging partners in systems-level data collection.]

Hard-won Advice

- **Align with the metrics that matter to your partner, which may just be a different way of analyzing and communicating data that you already collect.**

Example: *One Acre Fund regularly speaks about its impact in terms of dollars—the incremental income gained by a farmer when engaging with 1AF's model. But the organization recognized that agricultural ministries wanted to understand how work from partners contributed to the specific targets and objectives in their national strategies. In Rwanda, for example, the government wanted to know how many fruit trees were planted. In Kenya, the government was more interested in the total size of One Acre Fund's farmer loan portfolio. One Acre Fund thus took on the responsibility to understand and provide government partners with the specific data relevant to their goals.*

- **Be proactive about ensuring partner buy-in to external data so that they are willing to trust and use that data.**

Example: *Mo Adefeso-Olateju, Managing Director of The TEP Centre in Nigeria, spoke about beginning with the end in mind—knowing that the organization would want to use the data it was collecting to influence its government partners. TEP decided to engage an agency of the government, the National Bureau of Statistics, in developing the sampling methodology and in implementing the research in the field. When TEP went to share the results with other government entities within the country, there was little dissension on the results because they had confidence that a fellow agency had been part of the process.³⁰*

- **Create shared data infrastructure elements, whether you build off of government data systems or ensure your own systems are able to connect with them.**

Example: *Last Mile Health supported the Liberia Ministry of Health to integrate community health into the country's existing health management information system. MiracleFeet deliberately used common platforms (e.g., CommCare and Salesforce) to build its data technology because it was already proven, was developed for the unique contexts in which MiracleFeet works, and could connect to health management information systems being developed at the federal level in many countries.*

"Demonstrate a commitment to the government's vision. They need to see you as a development partner, and that what you're doing matches their priorities."

COLIN CHRISTENSEN
One Acre Fund

PROMOTE STREAMLINING AND SIMPLICITY.

The Risk

Each partner may have different expectations of and experience with the amount and types of data collected for use—often with a bias toward over-collection.

Hard-won Advice

- **Promote lean data practices that focus on timely, relevant, high-quality, simple measures.**

Example: *Pratham, which works to spread the gospel about “simple” (i.e., lean, relevant, and actionable) data, has faced challenges with some government teams who believe that “more is better” when it comes to data. When working through government systems to collect data, Pratham works to demonstrate the need for simplicity so that the data is collected quickly, is visible at all times, and is immediately actionable. Living Goods, which shares its data with the national health information systems, often receives pushback from government partners that its data is “incomplete” because it covers only pregnant women and children under five—but not other critical data on HIV, tuberculosis, diabetes, etc. Living Goods works to explain to government partners that collecting data for services beyond those which it provides would only distract the CHWs from providing their core services—but it is a continuing source of friction. Our interviewees shared numerous tips and considerations for homing in on only the most important data and simplifying collection processes so that they are highly repeatable; see Building Blocks 2 and 3 for more.*

“The tendency in the measurement space globally is to collect lots of information and process information without knowing how it will be used. So one of Pratham’s goals is to understand what data is important and useful and share this understanding with partners.”

DEVYANI PERSHAD
Pratham

Deep dive studies on partnership impact.

In addition to the simple, timely day-to-day performance data, many organizations need to undertake deep-dive studies to investigate the impact of implementation with and through partners. While Pratham is working on classroom-level data collection with expansion partners in new countries in Africa, it is concurrently considering pursuing discreet studies to better understand the impact of variations on key program elements, such as mentoring. Pershad explains, “While we work to ensure that simple data systems are built into expanded programs, we also need to ask—and collect data to answer—context-specific questions.”³¹



AS YOU DRIVE TO SYSTEMS CHANGE: Shift Your Conceptions of Data

Data to track, measure, and identify systems change³² often requires a different approach than that to collect data from more discrete programs. So what frameworks and methods can social ventures use to collect and communicate data as they steer toward systems change? The organizations we interviewed shared advice on two important questions:

1. **WHAT should social enterprises be measuring in their systems change efforts?**
2. **HOW should social enterprises go about measuring that data?**

WHAT SHOULD WE MEASURE?

Bart Houlihan, Co-Founder of B Lab, spoke about his organization's theory of change: "We're trying to build a movement of people using business as a force for good to create a more inclusive economy. The approach is simple: we certify leaders as certified B Corporations (B Corps), shine a light on those leaders, and then we encourage others to follow." Although the approach may be "simple" in Houlihan's words, the vision is grand. Houlihan and his team at B Lab are trying to fundamentally change the system of capitalism, shifting the focus from maximizing returns for shareholders to serving all stakeholders. B Lab has defined the systems change it seeks to achieve, but what data should it collect to assess progress toward those goals?

In reviewing B Lab's key systems change metrics, we found that they largely aligned with a framework we find useful for informing and mapping such metrics: FSG's Systems Change Conditions. These six conditions of systems change (see box below) can serve as a prompt for organizations to identify the critical levers that will lead to systemic change and therefore the data points that must be tracked and measured.

FSG's Systems Change Conditions—Definitions

Excerpted from [The Water of Systems Change](#)³³

- **Policies:** Government, institutional and organizational rules, regulations, and priorities that guide the entity's own and others' actions.
- **Practices:** Espoused activities of institutions, coalitions, networks, and other entities targeted to improving social and environmental progress. Also, within the entity, the procedures, guidelines, or informal shared habits that comprise their work.
- **Resource Flows:** How money, people, knowledge, information, and other assets such as infrastructure are allocated and distributed.
- **Relationships & Connections:** Quality of connections and communication occurring among actors in the system, especially among those with differing histories and viewpoints.
- **Power Dynamics:** The distribution of decision-making power, authority, and both formal and informal influence among individuals and organizations.
- **Mental Models:** Habits of thought—deeply held beliefs and assumptions and taken-for-granted ways of operating that influence how we think, what we do, and how we talk.

Power Scale: As You Drive to Systems Change, Shift Your Conceptions of Data

See below how a sample of B Lab’s systems change metrics fall within FSG’s framework:

FSG Systems Change Condition	B Lab’s Activities	Sample B Lab Data Points
Policies	B Lab has helped create a new corporate legal form, the “benefit corporation,” and collaborates with businesses, the capital markets, and policy makers to drive adoption around the world.	<ul style="list-style-type: none"> • # states and countries in which “benefit corporation” legal form has been adopted. • # policy changes that create incentives related to tax and procurement preferences.
Practices	Through its B Impact Assessment (BIA), B Lab provides tools and tracks data to measure companies’ impact on workers, community, environment, and customers.	<ul style="list-style-type: none"> • Business-level changes in BIA metrics, e.g., through its “Inclusive Economy Challenge”; 173 companies have reported substantive changes to their inclusion-related practices, working to narrow racial and gender gaps.
Resource Flows	B Lab tracks the flow of both financial and human capital to understand how B Corps and benefit corporations attract these resources, compared to traditional businesses.	<ul style="list-style-type: none"> • Amount of funding raised by B Corps and benefit corporations. • Revenue growth rates, e.g., 3x faster growth rate than traditional businesses • Measures of employee attraction and retention, e.g., B Corps have more than double the employee engagement rate.
Relationships & Connections	Through the certification network, convenings, and recognition of industry leaders, B Lab empowers and inspires companies to form coalitions and pursue collective action around areas of common interest.	<ul style="list-style-type: none"> • Actions taken, e.g., 500+ B Corps committing to net zero greenhouse gas emissions by 2030³⁴ and coalitions formed around racial equity,³⁵ women’s empowerment,³⁶ and salary disparity.³⁷
Power Dynamics	B Lab’s mission and activities are fundamentally focused on changing the power dynamics in business—shifting from shareholder primacy to a balance of the interests of all stakeholders, including employees, underrepresented populations, and other community members.	<ul style="list-style-type: none"> • # of traditionally underrepresented populations in ownership, board, and management roles. • Data on CEO pay ratios, e.g., B Lab data shows pay ratios of 6:1 versus 271:1 for traditional companies. • # of companies adopting the benefit corporation legal form and therefore committing to balancing interests of all stakeholders.
Mental Models	B Lab recognizes that the most meaningful systemic changes come from shifting cultural narratives and underlying mental models, and so it partners with and empowers actors at key leverage points in the narrative on capitalism—universities training next generation consumers and business leaders, media, and key corporate influencers.	<ul style="list-style-type: none"> • # universities engaged in teaching about B Corps to train the next generation of leaders, employers, and consumers. • Use of the term “stakeholder capitalism” in media. • Earned media for B Corps. • Engagement of influencers—including multinational corporations—which are pursuing B Corp certification.

Power Scale: As You Drive to Systems Change, Shift Your Conceptions of Data

In another example, Jordan Kassalow, Co-Founder of VisionSpring and EYelliance, spoke about how the approach to data he had used at VisionSpring would not work at the systems change-focused EYelliance. “The idea behind EYelliance was to take a multi-sector stakeholder kind of approach, try to elevate the issue area onto the global development agenda, get governments and private sector players to come along and partner with the NGOs, and create a larger group of organizations who were caring about this issue area. That’s what the EYelliance is. It really is looking at how you solve the problem in its entirety.” In order to measure that work, EYelliance selected the most critical levers needed to shift the system (notably, also aligned with FSG’s framework) and identified data points around those:

FSG Systems Change Condition	EYelliance’s Activities	Sample EYelliance Data Points
Relationships & Connections	Building relationships, engaging the public and private sector.	<ul style="list-style-type: none"> • # of proven solutions integrated into government systems. • # of investable vision activities that can attract new sources of capital.
Resource Flows	Advocating for funds to be appropriated to eyeglasses in key donor country foreign assistance budgets.	<ul style="list-style-type: none"> • Amount of new funds raised. (Example: after four years of EYelliance advocacy, the US Government State and Foreign Appropriations Bills included a line item for eyeglasses, allocating \$2.5 million in 2019 and \$3.5 million in 2020.)
Mental Models	Changing the mental model around vision (i.e., vision is not just about sight but also critically about economic development, road safety, and education) and raising the profile in the development community.	<ul style="list-style-type: none"> • # new partners and sectors engaged. (Example: International Labour Organization thinking about vision as an input.)

Figuring out what to measure is just the start, though.

HOW SHOULD WE MEASURE IT?

By its nature, data for systems change is hard to measure systematically, and the outcomes are rarely controlled by one organization alone. Interviewees shared how this reality requires a shift in the way they approach data, and provided advice for others pursuing systems change.

ADVICE FROM THE FIELD

Shift expectations for data and evidence.

Given the nature of systems change and the challenges that it poses for data collection, interviewees spoke about having to shift their—and their funders’ and partners’—expectations around what data and evidence is reasonable and meaningful when it comes to systems change. Some factors to consider:

- *Contribution over attribution.* No single organization can be in control of all the factors contributing to systems change outcomes. Therefore, organizations stressed needing to value contribution, not attribution, and be transparent with funders about what is in and out of an organization’s sole control. Organizations also stressed the importance of selecting partners thoughtfully and agreeing to data collection and reporting goals upfront. For more on data collection with partners, see page 30.
- *Qualitative measures.* “Not everything that counts can be counted, and not everything that can be counted counts.” This quote, of disputed authorship,³⁸ applies well to systems change—as not all aspects can be assessed with quantitative data, nor should that be the ideal. Embracing high quality qualitative data (e.g., anecdotes, observations,

etc.) is equally important and might often be better able to capture key milestones. There are formal qualitative methods for evaluation that organizations can pursue (e.g., outcome harvesting, comparative case studies, and developmental evaluation [see a list and links on page 11 of Rockefeller Philanthropy Advisors’ “Assessing Systems Change: A Funders’ Workshop Report”]),³⁹ but many approaches may be less formal, requiring less intense resources.

- **Slower data cadence.** We usually think about data as something that is collected regularly—e.g., enrollment numbers weekly, monthly, and quarterly—and for which trends can be easily tracked over time. But measuring systems change takes longer and is less regular: new relationships that may take months or years to formalize, policies that have changed over the course of time spent advocating and educating, etc. To track these changes, organizations must identify shorter-term progress indicators and be prepared to stick around for the long term.

In Action: What if the data simply doesn’t exist?

International Bridges to Justice, a global human rights organization focused on protecting individuals’ legal rights and reforming the legal system in developing countries, faced a difficult challenge when it came to data: the fact that data often did not exist or if it did exist, it was wildly unreliable (sometimes intentionally). For example, many countries in which it works report that there is no torture taking place. Yet, in speaking with lawyers in the country, IBJ found that nearly all reported that individuals they defended had in fact been tortured. IBJ was not deterred by the lack of data and shared several steps it took to address this gap:

1. **Using anecdotal data to identify patterns when more comprehensive data is unavailable or unreliable.** IBJ works in a field riddled with corrupt data, which founder and CEO Karen Tse knows cannot be taken at face-value. “When we talk to five people and all five say their clients have been tortured, we see that as a sample. A small sample, but at least we can start to analyze. We shouldn’t be intimidated by a lack of data—there is a lot we can do with what we know.” IBJ pieces together what data is available to identify patterns and also trusts its intuition based on many years working in this field.
2. **Developing proxies for data that doesn’t exist or is difficult to collect.** IBJ has identified what it calls a litmus test to determine the health of a country’s justice system, given the dearth of reliable data to assess the system in many of the places it works. The organization understands that, of all of the interdependent elements of a justice system, the weakest link is usually the criminal justice system—and thus criminal defenders, the group actively responsible for enforcing due process, who usually have a finger on the pulse of the whole system. By looking at data around criminal defenders specifically, IBJ can assess the relative health of human rights in that country.
3. **Recognizing that just having a metric is an important step.** IBJ has worked with some Ministries of Justice to establish new thinking, access to detention facilities and formal systems around torture, representation, etc.—knowing that even if the data is largely inaccurate today, acknowledgement of the importance of the issue is an important first step. IBJ knows that what gets measured gets done and, with collaborative system-wide collection, it can work to bolster the accuracy and comprehensiveness of the data collected.
4. **Moving to a bottoms-up approach to data collection.** Recognizing that the data and experience exists—and has tremendous value—at the community level and that a top-down approach has yielded little data, IBJ has moved to developing a digital platform to crowdsource data. JusticeHub will not only provide critical services to stakeholders in the legal system but will also rely upon them to regularly report what they are seeing and experiencing—thus building a significant base of data to inform action at the local, country, and global levels.

Be bold in selecting long-term measures that may not be in your control.

The path to systems change unfolds over long timelines and in a complex web of stakeholders where no single organization controls all the inputs to that change. It can be scary for an organization to set goals that are bold and truly drive to the systems change needed, but interviewees stressed that setting bold goals with partners is the only way that systems change will be achieved.

According to Sara Standish, VP for Strategy, Learning, and Impact for Health Leads, “As we have shifted to be more focused on systems change, we’ve had to name population level changes we want to see. These changes are challenging to achieve but we must be courageous to pick those hard things, understand our role in driving progress, and pair that with a results-based learning approach. It is scary because we know it’s not dependent on us alone, but it forces us to think carefully and intentionally about who we bring into partnership to have the change we want to see, and to be open to measuring that in new ways.”

As Standish references, Health Leads moved from collecting only data around patient outcomes to adding longer-term, population-level health changes that all partners agree to and play a role in achieving. For example, Health Leads tracks the number of states where community health workers are making an hourly living wage, which it sees as an indicator of the sustainability of caregiving—an essential input for accessible health resources.

Go beyond counting, and prioritize critical leverage points.

Some aspects of systems change can be counted, such as the number of states enacting Benefit Corporation legislation as tracked by B Lab, or the number of dollars allocated to eyeglasses in the US Federal Appropriations Bill, as tracked by EYelliance. These quantifiable data points are often related to changes in policies, practices, and resource flows and are important to track.

But measuring progress against system change conditions, such as relationships, power dynamics, or mental models, is often much more than a numbers game. Simply counting the number of relationships built provides some notion of progress but has little meaning; rather, these data points are better qualitatively described to capture the magnitude and significance of each relationship. Is there one relationship that is critical to achieve due to the brand and power that organization holds? Is there one shift in power dynamics that will serve as a tipping point to the rest of the system?

For example, a core focus of education non-profit Pratham’s work is shifting the mental model of educational programming from a focus on increasing enrollment to a focus on achieving learning outcomes—which it knows requires buy-in from the highest levels of government. In part due to Pratham’s work, in 2012 India’s erstwhile Planning Commission came out with a learning outcome policy for the country’s 12th Five-Year Plan. The following year, the central government’s education department issued guidelines reflecting many of the core principles of Pratham’s “teaching at the right level” methodology. These guidelines have since been actioned by several state governments in the country—a major step forward in changing the education system and likely a more significant relationship than others counted by Pratham.

Incentivize other systems actors to collect and use data.

In order to collect data at the systems level, interviewees reflected that it is often necessary to collect that data from partners across the system. Since you do not control those organizations, understanding how to incentivize and motivate them is critical—and challenging. Some tips from interviewees include the following:

- *Co-creating to achieve buy-in:* Incentivizing others to collect and share data begins with the essential, yet often lengthy, process of co-creating data strategies together to ensure buy-in. When creating its data platform, MiracleFeet was clear that if it wanted to create systemic change, it needed data that didn't just work for MiracleFeet, but for anyone working in clubfoot. Working together with partners, MiracleFeet and others were forced to find common ground, distilling for each partner what data points were “need to haves” versus “nice to haves,” and coming to a consensus about the data that would be collected by all.
- *Articulating value for data providers:* International Bridges to Justice recognized a desire within the global justice community to “use technology and data as a force multiplier.” In response, IBJ created JusticeHub as a digital home for the global community to share data related to rule of law. The data collected allows the system to detect spikes and “early warning signals” while also providing value to the global community through matching of lawyers with defendants, sharing of universal standards and training materials, and providing a platform to connect community members to each other.
- *Benchmarking and the power of community:* Some organizations spoke about using the data to build community or allowing for benchmarking by increasing transparency of data collected. For example, according to Houlahan, “At B Lab, we learned that benchmarking really works. Entrepreneurs are naturally competitive—they want to see how they stack up against others.” So B Lab makes its B Impact Assessment scores transparent on the website to allow for benchmarking, publicly celebrates and awards top scoring companies each year, and ensures that the impact data that B Lab is asking to collect is helpful for the organizations themselves so that they have a personal incentive to collect and share that data.



AS YOU HAVE MORE DATA AVAILABLE: Consider Advanced Data Methods Like Machine Learning

Machine learning and artificial intelligence are all around us: for example, Amazon recommending the next product that you “must” have and ride-share apps telling you how long until your driver arrives. And as we witness an explosion of data in the impact sector—resulting from an increased focus on data collection and the huge quantities of data becoming available through mobile phones, satellites, social media, and more—we are seeing more opportunities for machine learning in service of social impact.

The positive implications are enormous: early prediction of earthquakes, more accessible and higher quality health diagnoses, real-time deforestation monitoring, and so much more. However, the negative impacts can be major as well, not only taking precious time and resources to deploy, but amplifying biases and leading to investments in misguided interventions (often resulting from lack of context and poor-quality data). The positives and negatives of machine learning warrant careful consideration; our interviewees who have ventured into this area provided advice as shared below.

Is machine learning right for you?

What is the problem that you are trying to solve and is machine learning—or something simpler—the tool to solve it? As Sharmi Surianarain, Chief Impact Officer of Harambee warns, “Organizations and funders always want to go for the new shiny tools. But my advice would be: don’t start the conversation focused on machine learning and artificial intelligence. Instead, start by asking ‘what is the question that we want to answer?’ Then determine what data you have and what the right methodology and tools are to move us toward an answer.” It could be machine learning, but it is more likely to be something far simpler.

Surianarain’s caution was echoed by multiple interviewees, including Crisis Text Line’s Bob Filbin. He gave the example of when, early in Crisis Text Line’s history, it wanted to understand how Crisis Counselors were using their time and ways in which they could be more efficient. While Crisis Text Line was developing capacity for highly complex analytic methods, it was able to identify a major opportunity by first using simple descriptive statistics. The discovery that approximately three percent of its texters were using up to 34 percent of its Crisis Counselors’ time allowed Crisis Text Line to develop processes to support those frequent texters while freeing up counselor time for other text conversations. With respect to machine learning, Crisis Text Line spent four years thinking about the problems it wanted to solve and building data systems and a strong data culture before beginning to employ machine learning.

So how do you know if you have a problem that is ripe for a machine learning methodology? And if you have the right problem, how do you know that you have the right data to enable machine learning to work? Follow the stories of Educate Girls and Crisis Text Line below to learn more.

“Data is the cake, machine learning and artificial intelligence are the frosting; 80 percent of data problems are simple business intelligence problems. Instead of asking ‘how do we use a solution like machine learning’ ask what is the problem we are trying to solve?”

BOB FILBIN
Crisis Text Line

STEP ONE: ASSESS YOUR PROBLEM'S FIT WITH MACHINE LEARNING

Machine learning is particularly good at taking large amounts of data and identifying relationships across many variables, recognizing patterns, and making predictions based on those patterns. For example, machine learning predictions can help with prevention (e.g., forecasting disease outbreaks), targeting of programmatic interventions (e.g., predicting yields of different types of crops or improving credit scoring algorithms to predict factors for repayment), and filling in data gaps.

In order to assess if your problem is right for machine learning, you must first identify a mission-critical problem with a need for predictive capacity to solve.

	Educate Girls, using machine learning to fill in data gaps	Crisis Text Line, using machine learning for targeted triage and intervention <i>Content Warning: The sections below include words and phrases associated with suicide.</i>
Step One: Identify a Mission Critical Problem	<p>After a three-year randomized control trial demonstrated Educate Girls' impressive impact on enrollment and learning outcomes, Educate Girls was selected by the Audacious Project to receive funding for an ambitious scale-up of its program, from approximately 10,000 villages to 35,000 villages over the next five years.</p> <p>Educate Girls had a problem, though. It did not want to spend its precious resources scouting tens of thousands of villages, and it did not possess village-level data for the entire country. Thus, it posed the question:</p> <p><i>Given the huge number of potential villages to serve, how could Educate Girls efficiently and effectively determine which ones to target so as to reach the most out-of-school girls possible?</i></p>	<p>Crisis Text Line sees surges in the number of text messages it receives during certain times of day and knows that some of the texters are at imminent risk of harming themselves. During periods when there is a surge in texter volume, Crisis Counselors may not be immediately available—but would want to prioritize responding more quickly to those who are in imminent danger.</p> <p>The problem was that Crisis Text Line did not know which texters were in imminent danger. Thus, it posed the question:</p> <p><i>Given the limits of available Crisis Counselors during surges in volume, how can Crisis Text Line determine which texters are at most risk of harming themselves so they can be reached in the shortest time possible?</i></p>
Need for Predictive Capability	<p><i>Challenge: Incomplete data</i></p> <p>Educate Girls did not have all the data it needed to be able to identify villages with the most out-of-school girls, and therefore had a prediction problem. Educate Girls had collected household data in 8,000 villages where it had previously operated, but it did not have data on villages in potential areas of expansion. Government census data included counts of out-of-school girls, but it was outdated, of questionable accuracy, and did not go below the district level to the village level.</p>	<p><i>Challenge: Limitations of human analysis</i></p> <p>Crisis Text Line initially addressed this prioritization issue by creating a list of words based on clinical research that it believed were likely to be associated with imminent risk, such as "die," "cut," and "suicide." A text parser would monitor incoming texts and flag those including words from this list as "imminent risk" to have them moved to the front of the queue. Yet Crisis Text Line recognized the limitation of this approach, which missed the nuance of context and thus generated a level of performance that formed a baseline on which to improve the organization's high-risk detection.</p>

Want to learn more about Machine Learning for Impact?

Content in this section was based on interviews and also draws from the following excellent guides for learning about machine learning for impact:

- USAID, Reflecting the Past, Shaping the Future: Making AI Work for International Development, <https://www.usaid.gov/sites/default/files/documents/15396/AI-ML-in-Development.pdf>;
- Ben Brockman, Andrew Fraker, Jeff McManus and Neil Buddy Shah, "Can Machine Learning Double Your Social Impact?" Stanford Social Innovation Review 17, no. 1 (2019): https://ssir.org/articles/entry/can_machine_learning_double_your_social_impact;
- *Using AI for Social Good*, accessed May 15, 2020, <https://ai.google/education/social-good-guide>; and
- Serna K. Sgaier, "Demystifying Machine Learning for Global Development, Stanford Social Innovation Review 17, no. 3 (2019): https://ssir.org/articles/entry/demystifying_machine_learning_for_global_development.

STEP TWO: MAKE SURE YOU HAVE THE RIGHT OUTCOME AND PREDICTOR DATA

Once you determine that you have a problem that is a good fit for machine learning, your next step is to ensure that you have the data necessary to make machine learning successful. There are two types of data that are important for effective predictions: a subset of high-quality outcome data and reliable predictor data.⁴⁰

	Educate Girls	Crisis Text Line
A subset of high quality⁴¹ outcome data	<p>Jeff McManus, Senior Economist at IDinsight, spoke about IDinsight’s work with Educate Girls and how important it was that the organization possessed reliable data in some regions on the outcome that it was seeking to change: the number of in-school versus out-of-school girls. Through a door-to-door census, Educate Girls had already collected real household data for households in the 8,000 villages where it already operated, which included a count of out-of-school girls. This data was reliable, complete, and was critical for building and training the predictive model.</p>	<p>Once a text is received by Crisis Text Line, the texter is connected as quickly as possible to a live Crisis Counselor. Crisis Text Line has access to a growing corpus of over 150 million text messages—and, most importantly, has tagged these text messages with information about the outcome that Crisis Text Line is seeking to understand: the texter’s level of risk as assessed by a trained Crisis Counselor.</p> <p>The data is plentiful and high quality; the Crisis Counselors are trained on standards for assessing risk in addition to being overseen by Crisis Text Line supervisors who monitor and audit conversations.⁴²</p>
Reliable predictor data	<p>IDinsight used the government’s education dataset (published annually for every school in the country and including data on student/teacher ratio, gender, school infrastructure, etc.) as well as census data to find 300 indicators that could correlate with, and therefore predict, whether girls are in-school or out-of-school in various villages.</p> <p>McManus noted that “secondary data, such as that collected and published by government ministries, is often underutilized” but also warned that it often requires a lot of effort to get this publicly available data in the right format. “We had to scrape data from different websites, clean it, and match the government data across sources and with Educate Girls’ data—including identifying where the same village names had been spelled differently. This took a significant amount of time.”</p>	<p>Crisis Text Line was already in possession of a large quantity of predictor data embedded in the text messages it received and archived: the words used by the texter in their messages.</p> <p>By parsing the words in text messages that have been flagged by Crisis Counselors as imminent risk, the model is able to learn the words associated with this status and attach a relative determination of risk.</p>
Putting machine learning to use	<p>The factors above led Educate Girls and data partner IDinsight to use machine learning to build on the available data to predict where out-of-school girls were clustered. IDinsight estimates that over a five-year timespan Educate Girls would have been able to reach around 1 million out-of-school girls with the old approach. Using machine learning, it estimates Educate Girls will be able to reach over 50% more, a total of about 1.6 million, for roughly the same cost.</p>	<p>By using machine learning, Crisis Text Line is able to analyze large swaths of data and better assess the relationship between the words used and risk levels assigned by crisis counselors. Surprisingly, the machine learning model determined that words most predictive of a high risk of suicide are not the actual word “suicide” but words like “EMS” (five times more likely) or over-the-counter drug names such as “Advil” or “Ibuprofen.” By using the predictive power of machine learning, the team can identify 86 percent of people at severe imminent risk for suicide in their first conversations, allowing for those incoming messages to be prioritized immediately.</p>

READY TO TAKE THE NEXT STEP? Additional Advice from the Field:

Beware how machine learning can perpetuate inequity.

Jake Porway, Founder and Executive Director of DataKind stated, “If not designed properly, data science interventions can be ineffective or, worse, harmful. That goes doubly when it’s being used to support nonprofits and governments charged with caring for the most vulnerable populations.”⁴³ If bias exists in the way a question is posed or in the data that a machine learning model is based upon, the model and predictions will dangerously replicate and amplify that bias.

Our interviewees discussed grappling with these issues and addressing them through adjusted modeling and through programming. For Educate Girls, as it employed machine learning to find villages with the highest number of out-of-school girls, it recognized the possibility of the algorithm favoring certain types of villages, such as those with the largest number of people, which could disadvantage smaller, more rural, lower caste areas. Educate Girls and partner IDinsight built checks into the algorithm to ensure that this type of bias was not occurring. Whenever any especially vulnerable villages were given lower scores by the algorithm, Educate Girls and IDinsight manually overrode the model to ensure those villages were not excluded from the program. In Harambee’s case, it realized (through more basic data analysis) that young people are less likely to stay in a job if they are more than one to two taxi rides away from the work location. Using this insight, Harambee engaged machine learning to predict taxi routes from candidate home addresses to work locations, and thus match people against cost/time/distance parameters to increase the likelihood of job retention. It used this data, along with market intelligence on hiring patterns and employment opportunities, to provide suggested work opportunities to the youth it served. However, Harambee realized that youth living further away from economic centers would be left out using this data, as they were further from potential jobs. To address the bias against more rural youth with this particular intervention, Harambee is piloting a targeted program to enable rural youth to increase their income mobility—instead of focusing on job matching alone.

Social enterprises can take steps to avoid the dangerous consequences of bias in a machine learning effort, including (as a start):

- 1. Clearly articulating all of the objectives of particular effort with the data analysis team at the start (e.g., in Educate Girls’ case, the objective of reaching the most vulnerable out-of-school girls), and establishing processes to test the model against these objectives;**
- 2. Including multiple stakeholders and checkpoints in the development and validation of the model; and**
- 3. Using open source tools, such as the Aequitas project at University of Chicago or Deon to help assess data sets for potential bias.**⁴⁴

AI Principles

The OECD has developed a set of shared principles for Artificial Intelligence which have also served to inform the G20’s principles for responsible AI.⁴⁵ Both frameworks set out key standards and highlight human centeredness, fairness, transparency, security, and safety, among others, as central principles for all actors designing or leveraging AI solutions.

Build your team for machine learning.

Supporting machine learning projects requires a broad range of skills and experiences. The most obvious are the data scientists that understand how to build, test, and maintain the models. For Crisis Text Line, whose business centers on data and technology, having a variety of data scientists (including machine learning experts) on staff is natural. But for many other organizations, machine learning is best supported through external experts. Harambee partnered with a local Google cloud partner for its machine learning projects. Educate Girls has been working closely with IDinsight, a research and consulting organization. Several social enterprises focusing on community health workers are partnering with DataKind, a nonprofit that connects social impact organizations with data scientists. Many other organizations also exist that can provide external technical expertise.

Just as important as engaging the technical data science expertise for a machine learning project is bringing in subject matter experts (e.g., mental health clinicians), programmatic experts (e.g., those heavily engaged in the organization and implementation), and those who are closest to the ultimate client or community (e.g., field supervisor, community health worker). The model must match the realities of the ecosystem in which it operates, the organization in which it lives, and the people or other entities it serves.

Machine learning is not a replacement for humans.

Jackie Weiser, Lead Data Scientist at Crisis Text Line cautioned, “Think of machine learning as an assistant to humans ... and always have a healthy skepticism about the algorithms. Even two years into our machine learning effort, we question it regularly.” Weiser makes an important point that, although artificial intelligence and machine learning can help increase efficiency and provide more complex analytics, regular review and validation by humans remains critical—especially when working with vulnerable populations. Only humans can contextualize and identify extenuating circumstances, such as factors not captured in the data on which the algorithm was trained or nuances that exist in some contexts and not others. Two years into Crisis Text Line’s machine learning work, humans are still in the loop on all text conversations; supervisors have access to all messages, scanning and bumping up messages in the queue when they are deemed higher risk than the algorithm produces. In another example, Educate Girls continues to validate its model as new data becomes available. For example, as it moves into new states in India, it spends time collecting primary data—going door to door in several hundred villages to feed that data into the model and ensure that different cultural dynamics or other factors are considered appropriately.

FUNDER IMPLICATIONS

How can funders accelerate impact at scale with a different approach to data?

For many social enterprises, funders are one of the earliest, and most prioritized, stakeholders in their data strategy. As a result, funders are uniquely positioned to influence social enterprises to use data to power scale. On the flip side, funders also have the power to derail ventures through processes or structures that discourage learning from data, distract from the data that matters most, or disincentivize pursuing higher level goals like systems change. To effectively support a data strategy that powers scale, funders must adapt their approaches; the advice below reflects the experience of the social enterprises we interviewed and surveyed.

1

Ensure Data Strategy—Yours and Your Portfolio’s—is Grounded in Equity and Inclusion.

Scale introduces a new pace of growth and reach that has the potential to exacerbate bias. While equity considerations are critical throughout all points of development, scale demands particularly close consideration.

- **Align KPIs and reporting requirements to the metrics that hold social enterprises accountable to their clients/beneficiaries.**

- Many SEs prioritize data that holds them accountable to you, as a funder. Work with social enterprises to understand what KPIs would best hold them accountable to their clients—ensuring that client voice is captured in that process—and hold them accountable to those metrics.

EXAMPLE: Aligning KPIs with what matters to the client. Health Leads initially collected data on the wrap-around resources that its patients accessed (e.g., food pantry) as a measure of wellness. But in listening more closely to how patients themselves defined wellness, Health Leads recognized that, from its patients’ point of view, “wellness” is not so narrowly defined. Health Leads has begun to experiment with measurement approaches that are embedded in a more holistic inquiry into whether patients had the essential resources they needed to be healthy, thus holding itself accountable to wellness as its clients defined it.

- **Reconsider the types of data you value and require.**

- The Equitable Evaluation Initiative describes this challenge best: “Certain kinds of data and evidence have come to be viewed with value and legitimacy in philanthropy. Many foundation boards have come to expect simple quantitative dashboards, and those with particular academic backgrounds often value experimental research designs regardless of their fit to the situation. The field has come to treat with suspicion what is often called ‘self-reported data’ and to dismiss even systematically collected and analyzed qualitative data as merely ‘stories.’”⁴⁶ Consider the types of data you value from the enterprises in your portfolio—why it is so—and how you can prioritize both quantitative and qualitative data that allow for different voices to be represented.

- **Adopt or develop a set of principles on equitable evaluation.**

- There are several emerging standards that can be adopted by funders as a baseline to ensure evaluation efforts—by you and your portfolio enterprises—are equitable. The Skoll Foundation is exploring the equitable evaluation principles to inform how measuring and evaluating impact can “advance progress towards equity; answer critical questions about how historical and structural decisions have contributed to the social change being addressed; and foster community participation in defining success and shaping how evaluation happens.”⁴⁷ Other standards include the Responsible Data framework.

2

Approach Social Enterprise Data with a Learning Mindset, Not Just an Audit Mindset.

Just as it is critical for social enterprises to approach their own data with a learning mindset, it is critical for funders to model a similar approach to support the effective and sustainable scale of impact.

Model a culture of learning.

Funders can lead a move away from an accountability-only approach by modeling a culture of learning. Last Mile Health, the Liberia Ministry of Health, and funding partners model a learning approach by asking constructive questions and participating in problem solving, according to Last Mile Health President Lisha McCormick. This provides a “green light” to county health teams, helping to ensure that all data is used to promote program improvements and collective knowledge sharing.

Seek out—and be open to—multiple storylines to explain data.

Funders can play a role in encouraging organizations to seek out storylines to explain the data that challenge existing organizational narratives about what works and why. If you and your portfolio organizations are open to such inquiry, it could result in significant improvements—but also means that you must be open to such pivots as needed. This also means ensuring that your funding will allow for such data-driven changes.

3

Reconsider Your Data Expectations from Social Enterprises.

Many social enterprises find themselves creating data strategies that prioritize answering questions for funders over a more strategic approach focusing on data that can be collected in scalable ways and that will drive scaling decisions.

Help organizations keep their data lean and simple so that it can be scalable.

The collection and analysis of an organization’s core data must be simple and repeatable to allow for scale. The addition of too many additional reporting metrics from external stakeholders, including funders, can burden the data collection process and undermine the ability for strategic data to help power scale. Before asking for additional data, ask yourself the three questions that social enterprise Harambee uses before adding new metrics:

1. **Why are we tracking that?**
2. **What behavior will it drive internally?**
3. **Does it keep the client/beneficiary at the center?**

Some funders, including the Skoll Foundation, do not require grantees to report on a set of common measures because that would not truly capture the impact of each unique organization. Instead, the Skoll Foundation allows grantees to define their own metrics of success which they regularly report to the foundation. This approach is similar to the “bring your own lunch” approach described within MIT D-Lab’s “The Metrics Café.”⁴⁸

Be okay with “enough.”

Organizations must find a balance between feeling that they have too little data to act and collecting too much data which can lead to distraction and wasted time on collection and analysis. As a funder, come to agreement with your organizations on what is “just enough” and “good enough” data.

“Data is incredibly powerful to a point. It can also paralyze an organization,” says Houlahan of B Lab who encourages enterprises to “get good enough and go.”

Support Strategic Data Infrastructure Needs.

Substance must come before beauty when it comes to data systems and platforms—social enterprises must first establish their needs before landing on a solution. Once organizations are ready to invest in a data system that will bolster impact at scale, consider how to support them financially and with connections/expertise.

Question the timing of investment in a data platform.

Many organizations reported using basic data collection and analysis tools (often Excel) or inherited legacy systems while they build a more nuanced understanding of data requirements. Probe organizations on their readiness to invest in a new data system by asking if they have sufficient experience implementing in different geographies, with different populations, and in different logistics scenarios; have the people in place to use and manage a system; and have a sense of future data needs.

Provide a dedicated source of funding for data platforms.

Without a dedicated infusion of capital for data and technology, many of the organizations reported that they would not have taken the resources away from programs to invest in a data platform in the same way. When asked how much they would reasonably spend on a data system to meet their needs, if given a dedicated infusion of capital, 67 percent of survey respondents would spend USD\$250,000 or less.

Consider incentivizing the creation of systems that serve greater field needs.

With its Google Impact Challenge Award, MiracleFeet invested time working with other entities focused on clubfoot treatment to develop a data system that worked for the global challenge of clubfoot, rather than just for its own organization. If you want an organization to take into account the needs of others working in a particular sector when developing a data system, you should provide an incentive (e.g., funding) for them to take the time to do so.

Ensure organizations have people in place to actively use and manage the system.

Sixty-four percent of our survey respondents said that if they could have done something different to make their current data platforms more useful, they would have ensured they had the right skills and bandwidth on their team to use and manage the system. Ensure that as organizations are planning for an investment in data infrastructure, they are looking beyond development and are planning how they will use and maintain over the long-term.

“If we had developed our data system even a year or two earlier, it would have been a disaster. We needed to have enough diversity of experience in implementing our program—in different country contexts and facing different logistical challenges—to be able to design a system that would meet our needs.”

CHESCA COLLOREDO-MANSFELD
Executive Director of MiracleFeet

“If the social sector is going to benefit from the digital data revolution, funders will have to pay for it. For this shift to become reality, investments in improving measurement need to be seen as essential program costs. And that will require a mindset shift.”⁴⁹

JIM FRUCHTERM
Founder, Benetech

Shift Conceptions of Systems Change Data.

Data to track, measure, and identify systems change often requires a different approach than that to collect data from more discrete programs. As the Skoll Foundation's Liz Diebold and Anna Zimmerman Jin shared in "Measuring What Matters: Supporting Systems Level Change," "When looking across our portfolio, it can be tempting to anchor on countable impact metrics tied to direct service programming, e.g., the number of students reached, or patients treated. But even for organizations with highly countable interventions, the collaboration and influence that resist quantification matter too."⁵⁰

- **Align critical levers for systems change with data points to track and measure them.**

- FSG's six conditions of systems changes, as described in "The Water of Systems Change," provide a helpful framework for organizations and funders to map metrics. Work with organizations to identify the metrics that best align with the short-term and long-term change they are seeking.⁵¹

- **Look for shorter-term signals of systems change.**

- As the Skoll Foundation's Diebold and Jin shared in their article, "Systems change takes many years, but we look for shorter-term signals to spot it. For example, successful replication of a solution by government, the private sector, or civil society; leveraging a network of actors to achieve common impact goals; or contributing to changes in policies governing an ecosystem."⁵²

- **Value contribution over attribution.**

- Most organizations will have to deeply engage many stakeholders over time to achieve systems change and will often have to minimize the credit they seek in their work in order to empower others to sustain the change. Thus, funders should seek to understand how organizations are contributing to the change, as success will be beyond the contributions of one entity alone.

- **Be open to different types of evidence and methods.**

- While many funders and organizations are most comfortable with quantitative data and experimental evaluation methods, systems change requires a broader set of approaches to track progress and understand impact. Work to embrace different types of qualitative data, to be flexible when there is a dearth of data, and to appreciate the difficult-to-quantify magnitude or importance of milestones (such as a policy change, new partnership, or series of anecdotes).

KEY TAKEAWAYS FOR SOCIAL ENTREPRENEURS

Are you part of the 94% of leaders who feel you are underutilizing your data?⁵³ *Scaling Pathways' "Using Data to Power Scale"* draws on the experience of successful social enterprises to offer a high-level roadmap for designing a data strategy that can power impact at scale.

Data has never been easier or cheaper to collect, analyze, and mobilize—presenting an enormous opportunity for social enterprises to make informed pivots, find efficiencies, prioritize investments, and continuously improve. Yet there are also new pressures—the push to collect more, the process of democratizing data use, and the lure of the “shiny new object”—in this case, fancy dashboards and advanced analytics.

Is your data strategy ready for scale? Use this quick assessment with your team as a tool to take stock of your data strategy and identify where you need additional work and how you could benefit from the tips and insights within *Scaling Pathways' "Using Data to Power Scale."* Check all statements that apply, and consider additional exploration of those that do not yet apply.



LAY THE FOUNDATION

Ensure data efforts drive toward equity and inclusion.

To unlock the potential of data to power scale, the approach must be grounded in equity—ensuring that you are giving voice to all stakeholders involved in the scaling process.

We...

- Track and report on metrics that hold us accountable to our clients and what matters to them—not just to ourselves and our funders. For examples, see Harambee’s and Health Lead’s stories on page 5.
- Check for biases in our data collection by testing questions and collection methods with different constituent groups to account for different life experiences, interpretations, and cultural nuances. This is particularly important when implementing more advanced data-based solutions such as Machine Learning (see page 42).
- Mitigate against power imbalances in data collection—recognizing that when individuals with real or perceived power (such as people from outside the community) collect data, it can impact responses.
- Actively explore whose voices might be missing and review our data to ensure target stakeholders’ voices are included. If they are not, we question why and take action to address it—including considering different data collection methods (i.e., qualitative vs. quantitative).
- Have a data privacy policy and disclose this policy to stakeholders. We also obtain informed consent by providing information to those from whom we are collecting data on how we intend to use, store, and share that data.
- Provide our stakeholders—including our clients—access to the data we collect in a timely, consumable, and accessible manner.
- Engage diverse stakeholders, internal and external to our organization, in interpreting and analyzing our data. Divergent interpretations of data are unpacked to deepen understanding.
- Take action on client insights, measuring and tracking the integration of client input into decision-making.

For examples & tactics on embedding equity and inclusion in data strategy, see pages 4-7.



SET THE BUILDING BLOCKS

Answering the how, what, and who of data for scale.

Once the equitable foundation is laid, there are critical building blocks—the how, what, and who of data—that will power your drive to scale. A two-page memo for each building block will give you the top tips to get it right—but start here to see where you are already on the right path:

Creating a learning culture as you scale: *The road to scale is rarely linear; therefore, social enterprises must create a learning culture that uses data to improve, adapt, and pivot. Without this culture, organizations will likely default to accountability mode—driving toward targets but missing key insights and opportunities to achieve greater depth of impact.*

We...

- Make improvement an explicit purpose within our data strategy.
- Have routine processes for digging deeper into our data to challenge existing narratives and improve practice.
- Model a learning—not a punitive—mindset when discussing challenging data results with internal and external stakeholders.
- Embrace continued experimentation that offers new data to inform pivots and innovations.

For examples & tactics on creating a learning culture as you scale, see pages 9-11.

Decide what data to collect to drive action: *Too much or the wrong kind of data can stifle a leader's abilities to make decisions—or worse, lead to poor decisions. The best teams collect just enough of the right kind of data to support scaling decisions and drive action.*

We...

- Regularly track a limited number of KPIs (between 4 and 15) that will drive scaling decisions.
- Track data points that help manage the tension among financial sustainability, impact, and reach as we scale.
- Value both quantitative and qualitative data and collect both types regularly.
- Proactively advocate with external stakeholders (e.g., funders, partners) for the data that we know is critical to our unique scaling plans.

For examples & tactics on how to decide what data to collect to drive action, see pages 12-14.

Determine how to collect data with scale in mind: *More manual or resource-intensive data collection may work well in early stages but will be inefficient as the effort scales.*

We...

- Have data collection processes that are simple and repeatable and can be implemented by people with limited training.
- Use technology, where appropriate, to standardize, simplify, and streamline the data collection process.
- Identify opportunities for passive data collection by integrating it into actions our staff or clients are already taking.
- Leverage internal resources—integrating collection into existing roles, including volunteers, when feasible.
- Balance efforts to simplify data collection with equity considerations by ensuring that simplification does not reduce opportunities for diverse voices, perspectives, and nuance to be captured.

For examples & tactics on how to collect data with scale in mind, see pages 15-16.

Build data infrastructure to bolster scale: *The right data infrastructure can help manage and maximize data for scale—particularly the systems that store and communicate data—which can range from excel files to cloud-based, enterprise-wide data management systems.*

We...

- Started simple by leveraging lightweight tools (like spreadsheets) to meet our immediate needs while deepening our understanding of our longer term needs.
- Are seeking out, or have secured, a dedicated source of capital to support investment in our data infrastructure.
- Have considered how our data infrastructure can contribute to or support other actors in our ecosystem. For an example, see Miracle Feet's story on page 17.
- Ensure our data infrastructure meets or exceeds the legal privacy and security standards for our jurisdiction.

For examples & tactics on building data infrastructure to bolster scale, see page 17-20.

Engage the right people to support data goals: *As demands for and uses of data continue to evolve, new people strategies may be required to maximize use.*

We...

- Utilize a “deep and broad” people strategy; we have technical experts on our team while also ensuring the rest of the organization can use basic data and draw insights for their work.
- Have a team that brings strong local context and expertise in diverse data collection methods.
- Regularly train staff across the organization to use data and data infrastructure effectively.
- Have built up internal talent to maintain data infrastructure in-house (not relying entirely on external consultants).

For examples & tactics on engaging the right people to support data goals, see pages 21-24.



POWER SCALE:

Using data to drive specific scaling strategies.

With the foundation and building blocks in place, more advanced data approaches can be layered on. Depending on the scaling strategy you are pursuing and stage of scaling that you are at, consider the following approaches.

If you...are scaling through growth

Activate data use at all levels: *As organizations grow, the number of data users will increase and these users will need to be empowered to collect, analyze, and use data to drive scale.*

We...

- Have mapped who within our organization needs what type of information and when, so they can complete their work most effectively.
- Use technology, where appropriate, to ensure consistent and cost-effective access to data.
- Ensure team members and partners are incentivized to collect data by designing processes that save time, increase efficiency, and provide useful and timely information.
- Create tailored dashboards to make data easier to interpret and to drive action across all levels of the organization.
- Invest in training staff at all levels on how to draw actionable insights from data.
- Have built-in triggers and alerts that help keep data on the radar of busy team members.

For examples & tactics on activating data use at all levels, see pages 26-29.

If you...are scaling through partners

Adapt your approach to data: *Achieving impact at scale often involves various types of partnership, all of which will inevitably diminish the level of control you have over what is collected and how it is collected. Therefore, your approach to data must adapt for partnerships.*

We...

- Align with partners on a shared intention to use data for learning and improvement.
- Work with partners to align on metrics that matter to them—even if it means analyzing our existing data in different ways.
- Ensure our partners have trust in the data we're collecting and sharing.
- When possible, share data infrastructure with our partners—either building off of their systems or plugging them into ours.
- Keep it simple—and advocate for simplicity when partners tend toward more complex data requirements.

For examples & tactics on adapting your approach to data for partnerships, see pages 30-32.

If you...are driving systems change

Shift your conceptions of data: *Data to track, measure, and identify systems change often requires a different approach than that to collect data from more discrete programs.*

We...

- Have mapped data points that help us measure progress across different levers of systems change. For example, FSG's Six Conditions of Systems Change provides a framework for identifying levers and mapping possible data points. See page 33 for more.
- Have expectations of data that are reasonable and meaningful for systems change, including valuing contribution over attribution, embracing qualitative measures, and accepting slower data cadence.
- Are not afraid to identify ambitious and long-term measures we're working toward even as we recognize that many factors in achieving them are not within our control.
- Recognize the importance of gathering data from across the system and have found ways to incentivize others to collect and share their data.

For examples & tactics on shifting conceptions of data to support systems change, see pages 33-38.

If you...are driving systems change

Consider advanced data methods like machine learning: *Machine learning can be a powerful tool, but you must first determine if it is the right tool for the problem you are trying to solve.*

We...

- Have identified a mission-critical problem that requires predictive capacity to solve.
- Validated that the problem cannot be addressed through simpler data approaches.
- Have some good quality outcome data available and reliable predictor data upon which to build a model.
- Know machine learning can perpetuate inequities and have strategies in place to mitigate against this potential for bias.
- Have a team of technical and subject-matter experts—either internally or through trusted consultants—to help implement the approach.
- Don't expect machine learning to replace humans, so we have processes in place to ensure people regularly review our model to assess its continued fit and effectiveness.

For examples & tactics on using machine learning, see pages 39-43.

APPENDIX A:

Project Overview & Methodology

PROJECT: The Scaling Pathways project brings together the Innovation Investment Alliance (IIA) (a funding and learning partnership between the Skoll Foundation and USAID’s Global Development Lab, with support from Mercy Corps Ventures) and the Center for the Advancement of Social Entrepreneurship (CASE) at Duke University to study organizations that are attempting to scale impact and draw out lessons that are applicable to the social enterprise community at large. In Phase 1 of Scaling Pathways, we shared high-level lessons about scale in *Pivoting to Impact* and profiled the scaling journeys of three organizations —VisionSpring, Imazon, and Evidence Action—in in-depth Case Studies. In Phase 2, we are creating Theme Studies that distill advice from a variety of social enterprises related to financing, government partnerships, talent, and data. Find the full series at www.scalingpathways.com.

PROCESS: The Scaling Pathways partners surveyed social enterprises from across the Innovation Investment Alliance, USAID’s Development Innovation Ventures (DIV), and the Skoll Foundation portfolios to understand the challenges that they face on the road to scale. From this initial set of 100+ leading social enterprises, we conducted in-depth conversations with funders and reviewed literature and background materials to identify enterprises that we believed had interesting stories and lessons to share about each theme. We then conducted interviews, literature reviews (by theme and by organization), and conducted analyses for each of the enterprises interviewed. We also distributed a follow-up survey to 100+ social enterprises in the Skoll Foundation and USAID/DIV portfolios to gather additional insights specific to data (see more information below).

As part of the interview process, we gathered insights from the following individuals whose organizations are highlighted throughout the paper:

Wendy Chamberlin, Global Program Director, The BOMA Project; Bart Houlahan, Co-Founder, B Lab; Jackie Weiser, Lead Data Scientist, Crisis Text Line; Bob Filbin, Co-Founder and Chief Data Scientist, Crisis Text Line; Jordan Kassalow, Co-Founder EYElliance (and Co-Founder, VisionSpring); Marzanne Collins, Chief Information Officer, Harambee Youth Employment Accelerator; Sharmi Surianarain, Chief Impact Officer, Harambee Youth Employment Accelerator; Damon Francis, Chief Medical Officer, Health Leads; Sara Standish, VP for Strategy, Learning and Impact, Health Leads; Jeff McManus, Senior Economist, IDinsight; Karen Tse, Founder and CEO, International Bridges to Justice; Andrew Ozanian, Program Management and Development Associate, International Bridges to Justice; Sanjeewa Liyanage, International Program Director, International Bridges to Justice; Lisha McCormick, President and COO, Last Mile Health; Asif Akram, Chief Technology Officer, Living Goods; Chesca Colloredo-Mansfeld, Co-Founder and CEO, MiracleFeet; Amanda Springer, Director of Program Operations, MiracleFeet; Christie Pettitt-Schieber, Deputy Director of Programs, MiracleFeet; Hailey Leon, Application Systems Analyst, MiracleFeet; Katy Falletta, Monitoring, Evaluation, and Learning Manager, MiracleFeet; Colin Christensen, Global Policy Director, One Acre Fund; Devyani Pershad, Head of International Collaborations, Pratham; Karthik Menon, Co-Lead of Measurement, Monitoring, and Evaluation Team, Pratham; Siddesh Mhatre, Co-Lead of Measurement, Monitoring, and Evaluation Team, Pratham; and Ella Gudwin, CEO, VisionSpring. We also relied upon previous interviews with Maharshi Vaishnav, Chief of Staff, Educate Girls; Safeena Husain, Co-Founder and CEO, Educate Girls; Thea Aguilar, Deputy Managing Director, One Acre Fund; Lisa McCandless, Director of Business Development, Living Goods; and Emily Bancroft, President, VillageReach.

We distributed a second brief survey to the 100+ social enterprises in the Skoll Foundation portfolio. We received responses from 23 enterprises; 22 of those were complete submissions. We did not collect demographic information in the survey to save time for respondents; the Skoll Foundation awardee cohort is made up of high impact, scaling social enterprises in many sectors around the globe. Findings are included and noted throughout the paper.

APPENDIX B

Organizations featured in Scaling Pathways' Using Data to Power Scale

THE BOMA PROJECT

bomaproject.org

The BOMA Project is a nonprofit organization that implements a high-impact poverty graduation program for ultra-poor women in the drought-threatened arid lands of Africa—the “last mile” of economic and social isolation. Operating at the nexus of four critical United Nations Sustainable Development Goals, the BOMA Project empowers women, builds resilient families, instills hope, and transforms the conversation about what is possible.

Regions Served	Legal Structure	Impact Area	Year Founded
Kenya, Uganda	Nonprofit	Economic opportunity; Livelihoods; Women's Empowerment	2005

B LAB

bcorporation.net

B Lab supports a global network of people using business as a force for good. B Lab is the organization behind the global certified B Corp movement and they also promote mission alignment in businesses through innovative corporate forms and measuring what matters through the B Impact Assessment and B Analytics.

Regions Served	Legal Structure	Impact Area	Year Founded
Global	Nonprofit with For Profit Subsidiary	Sustainable Markets, Responsible Supply Chains	2006

CRISIS TEXT LINE

crisistextline.org

Crisis Text Line saw the need to create rapid-response crisis counseling by leveraging big data and the dominant communication medium of today—text messaging. Crisis Text Line supports people in imminent danger and uses its real-time data to continuously improve its intervention services. Crisis Text Line hosts the largest public dataset on mental health and shares aggregated and anonymized data via CrisisTrends.org with the goal of strengthening the broader mental health system.

Regions Served	Legal Structure	Impact Area	Year Founded
USA, Canada, United Kingdom	Nonprofit	Mental Health, Health Delivery	2013

EYELLIANCE

eyelliance.org

EYelliance is a multi-sector coalition that drives the global strategy to increase access to eyeglasses at scale so that those in less developed countries can fully avail themselves of vital educational and economic opportunities. EYelliance identifies highly effective, proven models and integrates those solutions into broader, pre-existing public and private systems to increase access to glasses at scale.

Regions Served	Legal Structure	Impact Area	Year Founded
Africa, Asia, Latin America	Nonprofit	Education; Economic opportunity	2014

HARAMBEE YOUTH EMPLOYMENT ACCELERATOR

harambee.co.za

Harambee Youth Employment Accelerator builds scalable solutions for the youth labor market across the formal and informal economy. Its matching tools and real-world training methodology help employers quickly and reliably gauge work-readiness and increase retention. Harambee is scaling its impact through expanding government and business partnerships and through knowledge sharing for practitioners to adopt and learn from the model.

Regions Served	Legal Structure	Impact Area	Year Founded
Rwanda, South Africa	Nonprofit	Economic Opportunity, Youth Job Skills	2011

HEALTH LEADS

healthleadsusa.org

Health Leads partners with communities and health systems to address systemic causes of inequity and disease by removing barriers that keep people from identifying, accessing, and choosing the resources everyone needs to be healthy. Health Leads focuses on addressing health-related social needs (e.g. food, housing, utilities) through a multi-pronged approach, including partnering with (and supporting the goals of) local communities, accelerating learning on a national level through networks and collaboratives, driving adoption of open-source data strategies, and targeting advocacy at all levels.

Regions Served	Legal Structure	Impact Area	Year Founded
United States	Nonprofit	Health, Health Delivery	1996

INTERNATIONAL BRIDGES TO JUSTICE

ibj.org

Bridges to Justice (IBJ) is a nonprofit organization that is dedicated to protecting the basic legal rights of ordinary citizens in developing countries. Specifically, IBJ works to guarantee all citizens the right to competent legal representation, the right to be protected from cruel and unusual punishment, and the right to a fair trial. It does this by providing direct technical support and training to emerging legal aid organizations; building international communities of conscience in support of these organizations; and advocating and supporting the prioritization of just and effective criminal justice systems on the agenda of organizations involved with international human rights and legal development.

Regions Served	Legal Structure	Impact Area	Year Founded
Africa, Asia, Europe and Latin America	Nonprofit	Human Rights, International Justice	2000

LAST MILE HEALTH

lastmilehealth.org

Last Mile Health's mission is to save lives in the world's most remote communities. The organization has worked for more than a decade to support governments to build national community health systems.

Regions Served	Legal Structure	Impact Area	Year Founded
Africa	Nonprofit	Health	2007

LIVING GOODS

livinggoods.org

Living Goods saves and improves lives by working in partnership with local governments to transform community health. Facilitated by cutting-edge mobile technology, results-based performance management systems, and a cadre of motivated and supervised community health workers, Living Goods supports the delivery of cost-effective basic health care to the doorsteps of people in resource-constrained settings.

Regions Served	Legal Structure	Impact Area	Year Founded
Africa	Nonprofit	Health, Livelihoods	2008

The Scaling Pathways series also includes a Scaling Snapshot of Living Goods, which you can find at www.ScalingPathways.com.

MIRACLEFEET

miraclefeet.org

MiracleFeet provides organizational, technical, and financial support to clinics throughout the world in order to provide treatment to children born with clubfoot. Using a bottom-up approach, MiracleFeet offers a continuous cycle of support for the treatment of children by supporting local practitioners who are trained in the Ponseti method, a non-surgical treatment that involves a series of plaster casts.

Regions Served	Legal Structure	Impact Area	Year Founded
Africa, Asia and Latin America	Nonprofit	Health and Disability	2010

ONE ACRE FUND

oneacrefund.org

One Acre Fund provides smallholder farmers in the most vulnerable regions a complete bundle of services focused on helping them increase their yields and farm profits, improve resilience, eliminate chronic hunger, and contribute to health. This bundle includes financing for the purchase of inputs required at the beginning of the season (e.g., seeds and fertilizer), delivery of farm inputs, training on modern agricultural techniques, and market facilitation.

Regions Served	Legal Structure	Impact Area	Year Founded
Africa	Nonprofit	Economic Opportunity, Agriculture, Financing	2006

The Scaling Pathways series also includes a Scaling Snapshot of One Acre Fund, which you can find at www.ScalingPathways.com.

PRATHAM

pratham.org

Pratham is on a mission to improve the quality of education in India and worldwide. Pratham focuses on developing and delivering solutions that are low-cost and replicable and can be implemented at a large-scale to deliver improved educational outcomes for children and youth.

Regions Served	Legal Structure	Impact Area	Year Founded
India	Nonprofit	Education	1995

VISION SPRING

visionspring.org

VisionSpring provides affordable, high quality eyeglasses to people living on less than \$4 a day. The organization accomplishes its work through a network of distributors and "micro-franchises," also providing livelihoods for community-based entrepreneurs.

Regions Served	Legal Structure	Impact Area	Year Founded
Asia, Africa, Latin America	Nonprofit	Health, Livelihoods	2001

The Scaling Pathways series also includes a Case Study of VisionSpring, which you can find at www.ScalingPathways.com.

REFERENCES

1. Unless otherwise noted, all quotations in this theme study are from interviews conducted by Kimberly Langsam, Ellen Martin, and Erin Worsham between December 2018 and April 2020.
2. Throughout the Scaling Pathways series, we use the term “social enterprise” interchangeably with the terms “social venture” and “impact enterprise” to mean a nonprofit or for-profit organization that aims to achieve social and/or environmental impact. We use “social entrepreneurs” to indicate the leaders of these organizations.
3. Kathleen Kelly Janus, “Creating a Data Culture,” Stanford Social Innovation Review 16, no. 2 (2018): https://ssir.org/articles/entry/creating_a_data_culture.
4. “Data as a Democratic Process,” Camfed, March 5, 2015, <https://camfed.org/latest-news/sxswedu-data-as-a-democratic-process/>.
5. Monisha Kapila, Ericka Hines, and Martha Searby, “Why Diversity, Equity, and Inclusion Matter,” Independent Sector, October 6, 2016, <https://independentsector.org/resource/why-diversity-equity-and-inclusion-matter/>.
6. Client refers to the individuals that social enterprises aim to serve whether they be paying customers, product or service users.
7. Rhonda Evans, Tony Siesfeld, and Gabriel Kasper, “Re-Imagining Measurement: A Better Future for Monitoring, Evaluation, and Learning in the Social Sector,” Deloitte, 2020, <https://www2.deloitte.com/us/en/pages/monitor-institute/articles/re-imagining-measurement-strategic-learning-toolkit.html>.
8. “Why am I Always Being Researched? A Guidebook for Community Organizations, Researchers, and Funders to Help us Get from Insufficient Understanding to More Authentic Truth,” Chicago Beyond, accessed May 1, 2020, <https://chicagobeyond.org/researchequity/>.
9. All mentions of a Scaling Pathways survey throughout the paper refer to the survey delivered by the Scaling Pathways team to social enterprises in the Skoll Foundation portfolio. For more information on the survey respondents, see Appendix A: Project Overview.
10. “Our Data, Our Destiny: Young Women Deliver Community Accountability to Girls,” Camfed, May 18, 2016, <https://camfed.org/latest-news/young-women-deliver-data-accountability/>.
11. Evans, Siesfeld, and Kasper.
12. Erin Worsham, Catherine Clark, and Robyn Fehrman, “Pivoting to Impact: Navigating the Road to Scale,” Innovation Investment Alliance and CASE at Duke, Scaling Pathways, May 2017, www.scalingpathways.com.
13. Erin Worsham, Robyn Fehrman, and Cathy Clark, “VisionSpring: Business Model Iteration in Pursuit of Vision For All,” Innovation Investment Alliance and CASE at Duke, Scaling Pathways, May 2017, [https://static.globalinnovationexchange.org/s3fs-public/asset/document/FINAL%20Scaling%20Pathways%20-%20VisionSpring%20Case%20Study%20\(5.16.17\).pdf?30S338DAAAnU9Wr8xS28IRnhr4oAlk9AL](https://static.globalinnovationexchange.org/s3fs-public/asset/document/FINAL%20Scaling%20Pathways%20-%20VisionSpring%20Case%20Study%20(5.16.17).pdf?30S338DAAAnU9Wr8xS28IRnhr4oAlk9AL).
14. One Acre Fund, Scale Innovations: Farmers First (December, 2014), https://oneacrefund.org/documents/122/Scale_Innovations_One_Acre_Fund_Social_Enterprise.pdf.
15. For a list and description of qualitative methods, see Rockefeller Philanthropy Advisors, “Assessing Systems Change: A Funder’s Workshop Report,” Scaling Solutions Toward Shifting Systems (June 29–August 1, 2019), <https://www.rockpa.org/wp-content/uploads/2019/10/Assessing-Systems-Change-A-Funders-Workshop-Report-Rockefeller-Philanthropy-Advisors-August-2019.pdf>, 11 or MarketShare Associates, https://usaidlearninglab.org/sites/default/files/resource/files/leo_selected_methods_and_tools_for_measuring_systemic_change_20150224.pdf.
16. Sasha Dichter, Tom Adams, and Alnoor Ebrahim, “The Power of Lean Data,” Stanford Social Innovation Review 15, no. 1 (2016): https://ssir.org/articles/entry/the_power_of_lean_data.
17. USAID, Considerations for Using Data Responsibly at USAID, <https://www.usaid.gov/sites/default/files/documents/15396/USAID-UsingDataResponsibly.pdf>, 4.
18. “Responsible Data Resources,” Responsible Data, March 28, 2019, <https://responsibledata.io/2019/03/28/responsible-data-resource-list/>.
19. The Data Lab, From Storming to Norming: Building High Impact Data Teams (2020), <https://www.thedatalab.com/wp-content/uploads/2020/02/Storming-a-Data-Team-Brochure-screen.pdf>, 6.
20. Asha Curran and Julia Rhodes Davis, “Code Switching Across the Social/Data Divide,” Stanford Social Innovation Review 16, no. 2 (2018): https://ssir.org/articles/entry/code_switching_across_the_digital_data_divide.
21. Jim Fruchterman, “Using Data for Action and for Impact,” Stanford Social Innovation Review 14, no. 3 (2016): https://ssir.org/articles/entry/using_data_for_action_and_for_impact.
22. “Our 2019 Results,” LivingGoods, accessed May 1, 2020, <https://livinggoods.org/what-we-do/results-evidence-and-research/our-results/>.
23. “Disruptive Digital Tools for CHWs,” LivingGoods, accessed May 1, 2020, <https://livinggoods.org/what-we-do/mhealth/>.
24. Fruchterman.
25. “Disruptive Digital Tools for CHWs.”

26. Elaine Chang, "Poverty Graduation Model: Mobile Tech Can Enable Success," Taro Works, September 27, 2019, <https://taroworks.org/poverty-graduation-model/>.
27. For more tips on creating management and program level dashboards, see Andrew Wolk, Anand Dholakia, and Kelley Kreitz, Building a Performance Measurement System: Using Data to Accelerate Social Impact (Root Cause, 2019), <https://www.issuelab.org/resources/24133/24133.pdf>.
28. Jake Porway, "Three Things Great Data Storytellers Do Differently," Stanford Social Innovation Review 14, no. 3 (2016): https://ssir.org/articles/entry/three_things_great_data_storytellers_do_differently.
29. "The Trials and Tribulations of Data Visualization for Good," Digital Impact, March 15, 2016, <https://digitalimpact.io/the-trials-and-tribulations-of-data-visualization-for-good/>.
30. Scaling Up Community of Practice's Education Working Group Webinar, October 10, 2019.
31. Resources for "lean" and social-enterprise friendly deep dive studies include 60 Decibels, 2020, <https://www.60decibels.com/>, formerly Acumen Lean Data, <https://acumen.org/lean-data/> and Evidence Lab at the Duke Global Health Institute, "Evaluation Toolkit," Duke Global Health Institute, 2018. <https://sites.globalhealth.duke.edu/evidencelab/resources/tools/>.
32. Systems change is defined as "the fundamental change in policies, processes, relationships, and power structures, as well as deeply held values and norms" in Srik Gopal and John Kania, "Fostering Systems Change," Stanford Social Innovation Review 13, no. 4 (2015): https://ssir.org/articles/entry/fostering_systems_change.
33. John Kania, Mark Kramer, and Peter Senge, "The Water of Systems Change" FSG, May 2018, https://www.fsg.org/publications/water_of_systems_change.
34. "500+ B Corps Commit to Net Zero by 2030," 11 December 2019, <https://bcorporation.net/news/500-b-corps-commit-net-zero-2030>.
35. Read more about the Dismantle Collective at <https://www.dismantlecollective.org/>.
36. Read more about We The Change at <https://www.dismantlecollective.org/>.
37. Read more about Desafío 10X at <https://www.desafio10x.cl/#lp-pom-block-9>.
38. This quote is often attributed to Albert Einstein, but many say it is from a 1963 paper by sociologist William Bruce Cameron.
39. Rockefeller Philanthropy Advisors.
40. Ben Brockman, Andrew Fraker, Jeff McManus and Neil Buddy Shah, "Can Machine Learning Double Your Social Impact?" Stanford Social Innovation Review 17, no. 1 (2019): https://ssir.org/articles/entry/can_machine_learning_double_your_social_impact.
41. "Subset" and "high quality" are the important words here. If you had all the data, you would not have a prediction problem. But it is critical that there is some good quality outcome data upon which the model can be developed.
42. For more information on Crisis Text Line's training processes, see Erin Worsham, Kimberly Langsam, and Ellen Martin, "People Matter: Evolving Talent to Drive Impact at Scale," Innovation Investment Alliance and CASE at Duke, Scaling Pathways Series, July 2019, <http://bit.ly/sp2talent>.
43. Datakind (blog), <https://www.datakind.org/blog/creating-a-systems-change-approach-for-data-science-ai-solutions>.
44. Center for Data Science and Public Policy, Aequitas, The University of Chicago, 2018, <http://www.datasciencepublicpolicy.org/projects/aequitas/>. Driven Data, Deon: An Ethics Checklist for Data Scientists, accessed May 15, 2020, <https://deon.drivendata.org/>.
45. "Artificial Intelligence: OECD Principles on AI," OECD, accessed May 15, 2020, <https://www.oecd.org/going-digital/ai/principles/>.
46. "The Equitable Evaluation Framework," Equitable Evaluation Initiative, 2020, <https://www.equitableeval.org/ee-framework>.
47. Liz Diebold and Anna Zimmermann Jin, "Opinion: Measuring What Matters: Supporting Systems Level Change," Devex, December 20, 2019, <https://www.devex.com/news/sponsored/opinion-measuring-what-matters-supporting-systems-level-change-96215>.
48. Laura Dudzyna, The Metrics Café: Who Decides What Impact Data Gets Measured? A Guide to Bring Funders and Grantees to the Table (MIT Impact Alliance), <https://d-lab.mit.edu/sites/default/files/inline-files/Metrics%20Cafe%20full%20final%20fixes%20Nov%202017.pdf>.
49. Fruchterman.
50. Diebold and Jin.
51. See more on page 4 of John Kania, Mark Kramer, and Peter Senge, "The Water of Systems Change" FSG, May 2018, https://www.fsg.org/publications/water_of_systems_change.
52. Diebold and Jin.
53. In "Creating a Data Culture," Kathleen Kelly Janus reports that only 6% of social impact organizations surveyed reported that they believed they were using the data they collect effectively (Stanford Social Innovation Review 16, no. 2 (2018): https://ssir.org/articles/entry/creating_a_data_culture).
54. The organization descriptions are largely drawn from the Skoll Foundation (<http://skoll.org/community/awardees/>) and individual organization websites, with some modifications.



Scaling Pathways

Insights from the field on unlocking impact at scale

The Innovation Investment Alliance (IIA):

The Innovation Investment Alliance (IIA) is a funding and learning partnership between the Skoll Foundation and USAID's Global Development Lab, with support from Mercy Corps, that has invested over \$50 million in eight proven, transformative social enterprises to scale their impact. In 2017, with all its funding committed, the IIA is focusing on drawing out lessons on scaling that are applicable to the social enterprise community with the aim to inform the ongoing conversation on how to create systems-level change and sustainable impact at scale.

The IIA's partners include:

- **The Skoll Foundation** drives large scale change by investing in, connecting, and celebrating social entrepreneurs and the innovators who help them solve the world's most pressing problems. Skoll brings an expertise in identifying and cultivating social entrepreneurs. Learn more at www.skoll.org. The U.S. Global Development Lab (The Lab) increases the application of science, technology, innovation, and partnerships to achieve, sustain, and extend USAID's development impact to help hundreds of millions of people lift themselves out of extreme poverty. The Innovation Investment Alliance is supported by The Lab's Center for Transformational Partnership. Learn more at www.usaid.gov/globaldevlab.
- **The U.S. Global Development Lab (The Lab)** The U.S. Global Development Lab (The Lab) serves as an innovation hub. It takes smart risks to test new ideas, and partners within USAID and across other actors to harness the power of innovative tools and approaches that accelerate development impact. The Lab brings together diverse partners to catalyze the next generation of breakthrough innovations to advance USAID's mission to save lives, reduce poverty, strengthen democratic governance, and help people emerge from humanitarian crises and progress beyond assistance. Learn more at www.USAID.gov/GlobalDevLab
- **Mercy Corps** empowers people to survive through crisis, build better lives and transform their communities for good. Mercy Corps brings its experience in developing field-based programming in over 40 countries and investing in disruptive start-ups to the selection, evaluation and management of organizations selected for funding. Learn more at www.mercycorps.org.

The Center for the Advancement of Social Entrepreneurship (CASE) at Duke University:

CASE is an award-winning research and education center based at Duke University's Fuqua School of Business. Since 2002, CASE has prepared leaders and organizations with the business skills needed to achieve lasting social change. Through our research, teaching, and practitioner engagement, CASE is working toward the day when social entrepreneurs will have the skills, networks, and funding needed to scale their impact and solve the world's most pressing social challenges. Learn more at www.caseatduke.org.

Part of the
Scaling Pathway series

Find the full series at

www.scalingpathways.com