

# Measuring systems impact with a social network analysis

Understanding the multiplier effect of the  
Data Collaboratives for Local Impact program

Samhir Vasdev  
IREX

March 2020





#### ABOUT THE DATA COLLABORATIVES FOR LOCAL IMPACT (DCLI) PROGRAM

DCLI is a unique partnership between the Millennium Challenge Corporation ([MCC](#)) and the President's Emergency Plan for AIDS Relief ([PEPFAR](#)) to strengthen the use of data for more effective decision-making, particularly in areas that impact health, gender equality, and economic growth. Learn more at [www.dcli.co](http://www.dcli.co).

#### ABOUT THIS PUBLICATION

This report was prepared by Samhir Vasdev, Senior Technical Advisor for Digital Development at [IREX](#), a global development and education organization. It can be cited as follows:

**Vasdev, Samhir. "Measuring systems impact through a social network analysis."  
Prepared by IREX for the Millennium Challenge Corporation (March 2020).**

#### ACKNOWLEDGMENTS

The partnership of **Frederick Ruta**, **Omar Bakari**, and other staff at the Tanzania Data Lab ([dLab](#)) was instrumental to the design and implementation of data collection for this analysis. The data collection survey was shaped with the wise guidance of **Diana Harper** of [Linc Local](#). **Agnieszka Rawa**, Managing Director for DCLI, and **Mitch Blaser**, DCLI Regional Lead, provided invaluable direction and feedback on this report, as did **Dara Lipton**, Senior Program Officer at IREX.

# Measuring systems impact with a social network analysis

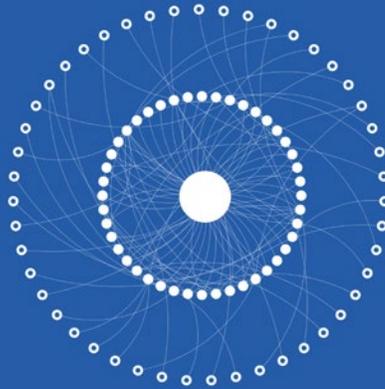
Understanding the multiplier effect of the Data Collaboratives for Local Impact program

## Summary

Since 2016, a unique partnership between MCC and PEPFAR adopted a systems thinking approach to strengthen Tanzania's community of actors using data to advance health, gender, and economic growth outcomes. This program, called Data Collaboratives for Local Impact (DCLI), was intentionally designed to strategically inject key resources into that system, with the hope that they would have sustaining impact that outlasts the program's implementation period.

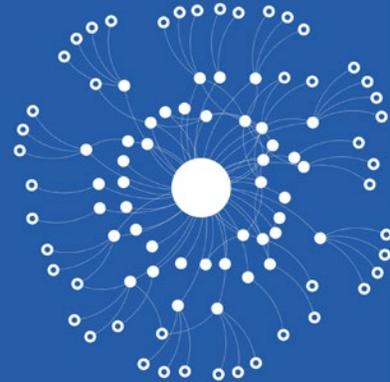
After concluding most of its activities, DCLI commissioned a social network analysis to understand the impact that the program had on Tanzania's system of data and health actors. This summary highlights key findings and implications about DCLI's systemic impact around **one year after the program concluded.**

42 direct beneficiary **institutions** in turn engaged 47 other indirect organizations.



DCLI's multiplier effect was between **1.1x** and **1.8x**.

DCLI catalyzed 70 new, unanticipated **relationships** without any DCLI involvement.



50 of these reached organizations with whom DCLI had never directly engaged.



Months, or even years, after originally engaging with DCLI, direct beneficiary organizations were **sharing data skills, exchanging financial resources, and partnering** on projects and proposals.



Two-thirds of survey respondents who directly interacted with DCLI proceeded to engage at least one other institution as a result of their interaction with DCLI. **Almost one-third of them engaged four others.**

These insights paint a picture of DCLI's lasting impact long after its programming concludes, and calls attention to the value and efficacy of **skills-focused, data-driven investments to elicit systemic change.**

As the interest of PEPFAR and MCC to strengthen in-country systems grows in order to build sustainable change—and as other funders also turn their attention to this type of foundational investment—the DCLI program demonstrates methods and tools which larger investments can use to take advantage of similar multipliers.

Follow the DCLI story at [www.dcli.co](http://www.dcli.co).



# Contents

|   |           |
|---|-----------|
| <b>Introduction</b>                       | <b>5</b>  |
| <b>Methodology overview</b>               | <b>8</b>  |
| <b>Findings</b>                           | <b>10</b> |
| <b>Conclusions</b>                        | <b>22</b> |
| <b>Annex A: Survey tool</b>               | <b>25</b> |
| <b>Annex B: Anonymized survey results</b> | <b>31</b> |
| <b>Annex C: Methodology</b>               | <b>32</b> |
| <b>Annex D: Terminology</b>               | <b>38</b> |

# 1 | Introduction

## CONTEXT

Responsive and impactful development programming relies on the effective and strategic use of data to inform decisions. Between 2016 and 2018, the Data Collaboratives for Local Impact (DCLI) program—a unique partnership between MCC and PEPFAR to strengthen evidence-based decision-making for improved HIV/AIDS and health, gender, and economic outcomes—implemented a constellation of activities across Tanzania to help institutions, entrepreneurs, and government actors use data more effectively and systematically (see [Box 1](#)). DCLI’s theory of change hypothesized that improved capabilities and resources to use data for effective decision-making would, over time, lead to improved decisions related to policies and programs including service provision and budget decisions related to HIV/AIDS and health, gender equality and economic empowerment in local communities.

Due in part to its relatively short implementation period (its projects ran between two and four years),<sup>1</sup> DCLI was intentionally designed from the outset to lay the foundation for sustained and locally owned impact that outlasts the program’s lifecycle. This approach was rooted in systems thinking, which affords deep attention to local contexts, relationships, roles, and rules that dictate how actors engage with each other to sustain development outcomes.<sup>2</sup> In practice, this meant focusing on local actors, including subnational ones who are part of weaker health data systems, embedding flexibility into DCLI’s implementation ethos, trusting local actors to facilitate dialog and relationships, and putting a premium on strengthening local institutions who can sustain impact after DCLI’s investments.<sup>3</sup>

These and other ingredients of systems thinking play a growing role in MCC’s approach to the design of individual activities and components of MCC’s program, but MCC has long championed investments whose benefits outlast and outgrow

---

<sup>1</sup> [Data Zetu](#)’s programming concluded in December 2018, while [DLI](#)’s concluded in June 2018. Data collection for this analysis began November 2019. The [dLab](#) project transitioned into a local NGO after its conclusion, with DCLI support continuing through November 2020. Learn more about these projects in [Box 1](#).

<sup>2</sup> Also referred to as systems change, systems strengthening, or systems design.

<sup>3</sup> See “[Ten Principles for Engaging Local Systems](#)” (USAID, 2014)

individual programs (beginning with MCC’s data-driven country selection methodology that seeks evidence of local ownership for partner countries).<sup>4</sup> Indeed, over the past decade, MCC has placed even more emphasis on fostering “durable systems” and approaches that enabled system-level impact.<sup>5</sup> How can we best capture these longer-term impacts?

## EVALUATING SYSTEMS CHANGE: SOCIAL NETWORK ANALYSES

The growth of MCC programs that are designed using systems approaches requires commensurate attention to evolving methods to evaluate change in those systems. MCC’s existing monitoring, evaluation, and learning (MEL) methods are robust and rigorous, but systems thinking challenges us to develop complementary and creative approaches that can enrich our understanding of how and why systems evolve as a result of MCC’s investments.

System mapping is an effective way to evaluate systems change, because they clarify how actors in a system engage with each other, why they do or do not, and what levers are primed to have the most impact. Although many methods to understand systems exist,<sup>6</sup> **social network analyses** (SNAs) are particularly relevant to DCLI, and indeed many other MCC programs. SNAs expose actors and their relationships with each other, providing useful insights into how a system operates and—if the same or similar SNA is issued sequentially over time—how that system is changing.

### WHAT IS A SOCIAL NETWORK ANALYSIS?

A social network analysis refers in general to an

<sup>4</sup> See “The Effects of the MCC Effect” (MCC, 2013)

<sup>5</sup> See “New Strategic Directions” (MCC, 2016)

<sup>6</sup> Other methods include political economy analyses, causal loop diagrams, performance improvement indices, and market studies. See Connecting the Dots (USAID, 2020), slide 7 for details.

### Box 1: What is the Data Collaboratives for Local Impact program?

DCLI ([www.dcli.co](http://www.dcli.co)) is a unique partnership between the Millennium Challenge Corporation (MCC) and the President’s Emergency Plan for AIDS Relief (PEPFAR) to strengthen the use of data for more effective decision-making, particularly in areas that impact health, gender equality, and economic growth.

DCLI operates in two countries in sub-Saharan Africa that have among the highest opportunity for impact in reducing HIV: Cote d’Ivoire and Tanzania. This study explores the impact of DCLI’s three projects in Tanzania, implemented between 2016 and 2019:

- The Tanzania Data Lab ([dLab](#)) is a premier center for excellence and innovation in data use. Located at the University of Dar Es Salaam’s College of ICT, the dLab converted into a locally owned and operated NGO in 2018.
- The Data for Local Impact Innovation Challenge ([DLI](#)) was an innovation competition that challenged Tanzanian youth to develop data and digital solutions to health, gender equality, and economic growth problems.
- [Data Zetu](#) operated in fourteen wards across Tanzania, supporting local institutions to build and adopt data use capabilities to strengthen evidence-based decision-making.

This study was conducted by [IREX](#), the lead implementing partner of the Data Zetu project.

investigation to understand the *relationships within a community*. An SNA often draws upon different tools, ranging from conventional survey analysis to applied data science, to understand how relationships exist between organizations, how information flows between those organizations, and how those elements changed as a result of an intervention.

SNAs are growing in popularity in mixed-method evaluative research, but they aren’t without limitations. For example, depending on how its data is collected, SNAs often identify areas

## Box 2: Strengthening systems with improved data use

DCLI is underpinned by the belief that improved use of data—that is, accessing, manipulating, producing, or sharing data more effectively and efficiently—fosters more frequent evidence-based decisions that improve service delivery, accountability, and governance.

This belief is harbored by the many institutions who ascribe to the “data revolution,” which is growing in momentum. The World Bank’s 2021 World Development Report will be [about data for development](#). The [Global Partnership for Sustainable Development Data](#) stewards greater availability of that and cross sectoral coordination in order to track progress towards sustainable Development Goals. [USAID](#) is in the process of rolling out a Digital Strategy; while the OECD Development Assistance Committee ([DAC](#)) launched a working group focused on data for development. The [Center for Global Development](#) is researching how best to balance the needs for good data governance, particularly in emerging economies so that they may balance the opportunities afforded by private sector digital innovation with the need for better service delivery as well as privacy, security and interoperability. Finally, leading technology firms also [have a role](#) in enabling the data revolution.

Investments into improved data use are sector-agnostic. That is, the same skills and assets used to improve education service delivery could be used to inform the design of a dam or health care facility. For this reason, data use interventions are particularly relevant for any program seeking systems-level change. Improved production and use of statistics can serve all aspects of a project’s life cycle (from program design and implementation to evaluation) and is a natural enabler of any systems change effort.

qualitative investigations. They might also be limited by the tools available to analyze and visualize the data.<sup>7</sup>

To mitigate these limitations, foreign aid agencies recommend complementing SNAs in concert with other evaluative approaches, which is why this SNA was designed in tandem with other

M&E framework (such as the program’s indicator tracking table) and an independent systems-level evaluation prepared by SPACES.<sup>8</sup> SNAs should be perceived as a complementary data collection instrument that can feed into and complement traditional M&E frameworks, to offer additional depth to indicators and targets.

<sup>7</sup> Matt Baker, [Demystifying Social Network Analysis in Development: Five Key Design Considerations](#) (USAID, 2019)

<sup>8</sup> SPACES is the [Strategic Program for Analyzing Complexity and Evaluating Systems](#), a consortium hosted by USAID’s Global Development Lab. To ensure coordination, this SNA was designed with input from the SPACES team.

## 2 | Methodology

### SUMMARY

Although substantial details of the methodology are provided in [Annex C](#), this brief summary of the SNA approach will help readers interpret the findings accurately and responsibly.

The SNA was designed to investigate four research questions:

- **R1:** How did DCLI investments in Tanzania create or strengthen linkages between data ecosystem stakeholders?
- **R2:** Which organizations produced the greatest number of new or strengthened relationships?
- **R3:** What was the modality of these new or strengthened relationships? (New partnerships? Shared experiences or knowledge? etc.)
- **R4:** What secondary relationships (that is, those beyond the scope of DCLI's interactions with primary beneficiaries) were cultivated as a result of the DCLI investments?

Collectively, these questions can help illuminate how programmatic inputs and short-term outcomes cascade within the ecosystem—thereby offering meaningful insight into the medium- and long-term impact of the change process catalyzed by DCLI's investments in Tanzania.

To answer these questions, an online, 5-10 minute English-language survey was shared via email to over 3,500 DCLI beneficiaries between October and November 2019. It invited respondents to identify up to four organizations they engaged as a result of their own interaction with any DCLI project, and then answer questions about the nature of those engagements. The full survey is available for reference in [Annex A](#).

To incentivize responses, recipients of the survey link were entered into a lottery for 100,000 shillings (roughly \$50) of mobile data credit. 63 people completed the responses, of which 47 passed the researchers' validation and cleaning process. This response rate of 1.3% is on the low end but generally consistent with what is expected from mass emailed surveys (see [Box 3](#)).<sup>9</sup> A link to anonymized raw survey results is available for reference in [Annex B](#).

---

<sup>9</sup> Acceptable response rates vary according to survey type and context. To illustrate the diversity: UN DESA generally seeks a minimum response rate of 60% for household surveys ([source](#), p. 223), while e-mail campaigns generally [achieve](#) between 1 and 10%.

## KEY BIASES AND LIMITATIONS

The entire process of survey design, survey distribution, data collection, and cleaning undoubtedly introduced biases that impact the final result of this SNA. A more comprehensive list of biases are in [Annex C](#) (Methodology), but some key biases and limitations to consider when interpreting the results are:

- Respondents who are not comfortable reading or comprehending English, who are unwilling or unable to use the internet to take the survey, who did not check their email, are excluded from this dataset.
- Respondents skew towards Data Zetu beneficiaries because Data Zetu had an outsized representation on the roster of potential respondents.
- Potential respondents who did not experience any strengthened relationships might not have responded at all.<sup>10</sup>
- Since our method (like other SNA approaches) did not have a control group, we don't know how relationships would have been established or strengthened without DCLI's engagement.

These constraints do not mean we can't draw meaningful conclusions about DCLI's impact. For instance, let's imagine that one finding is that a connection between organizations A and B was created as a result of DCLI. This fact is noteworthy and valuable, regardless of whether Organizations A and B might have created a relationship without DCLI, because the survey questions are designed such that respondents attribute their changes

to DCLI (i.e., the survey questions themselves included the language "...as a result of your engagement with DCLI"?). Non-experimental designs that rely on self-reported change are not the most robust scientific efforts, but as a complementary evaluation approach they are still considered valid by global development actors and funders including USAID.<sup>11</sup>

### Box 3: How should I interpret a response rate?

By itself, a response rate doesn't explain much about the validity of results. What matters more is how representative those respondents are of the population in question, and the impact of the non-response bias incurred by under-representing populations. For example, 44.5% of all DCLI participants based their interactions with DCLI in Dar Es Salaam, while around 57.4% of SNA survey respondents were from Dar. While this means we are under-representing non-urban DCLI participants, our findings suggest that the multiplier effect is larger in non-urban settings than in urban ones. So, the impact of the non-response bias is actually that we are probably **underestimating our multiplier effect** and offering more conservative impact estimates than the reality.

## PARTNERS

This SNA was conducted in collaboration with, and with the help of, several institutions. The Tanzania [dLab](#) (the NGO that arose from one of the three DCLI projects in Tanzania) tested the survey, prepared rosters of potential respondents, and helped administer the survey electronically to potential respondents. The USAID-led consortium [SPACES](#) provided external feedback and expertise to shape the survey, both strategically and tactically. MCC provided overall guidance and oversight on the process.

<sup>10</sup> This specific bias means that our analysis is more about how strengthened relationships were strengthened, instead of which relationships were or were not strengthened.

<sup>11</sup> See "[Introduction to Impact Evaluation Design](#)" (USAID)

# 3 | Findings

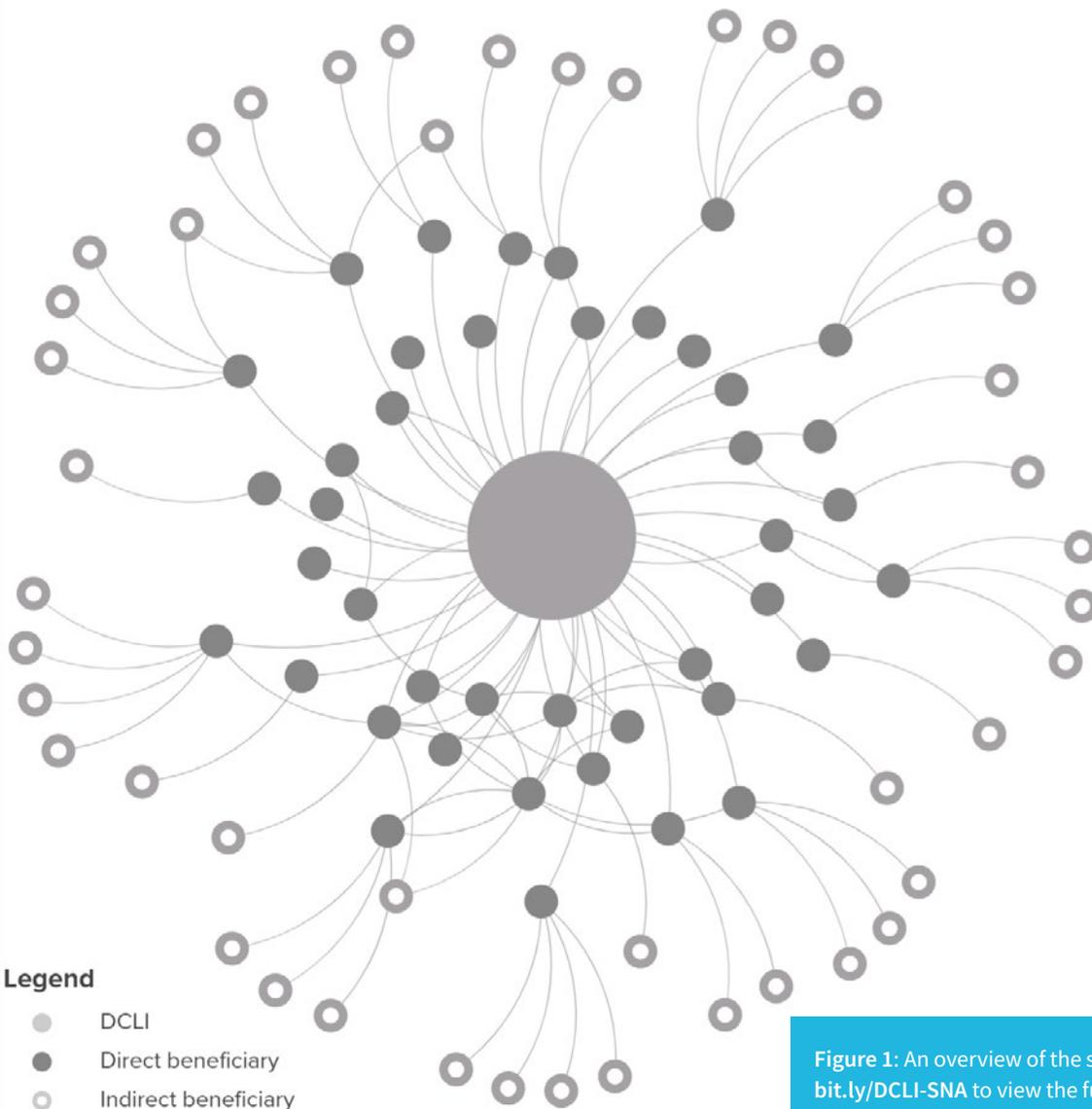


Figure 1: An overview of the system map. Visit [www.bit.ly/DCLI-SNA](http://www.bit.ly/DCLI-SNA) to view the full interactive version.

This analysis references some terms—like “indirect beneficiary” or “connection”—that warrant definition. Readers are encouraged to scan [Annex D: Terminology](#) before continuing, or reference it to aid with interpretation of these findings.

Our data collection, validation, and cleaning process enabled us to visualize the state of the DCLI network in Tanzania using Kumu, a free online tool for network mapping. Images from this visualization are interspersed throughout this section. **Readers are invited to view the interactive network maps here: [www.bit.ly/DCLI-SNA](http://www.bit.ly/DCLI-SNA).**

## INSIGHTS

Our analysis of the SNA survey response is structured according to the four research questions.

### R1: How did DCLI investments in Tanzania create or strengthen linkages between ecosystem stakeholders?

Our map of the ecosystem that resulted from DCLI’s investments shows **118 connections between 90 organizations**. Since our research question is about linkages between organizations, we will focus primarily on the nature of those 118 connections. These connections paint an exciting picture about the breadth and depth of DCLI’s secondary impacts and the relationships it catalyzed (see [Figure 1](#)).

70 of these connections—that’s 59% of all relationships in the ecosystem—resulted from DCLI *without DCLI’s direct involvement* (i.e. as systems effects). Twenty of those represent engagements between organizations whom DCLI directly reached in some way (such as through

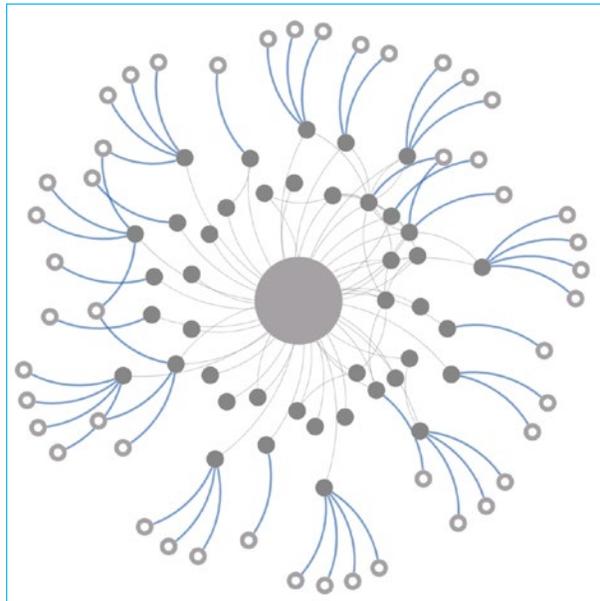


Figure 2: 42% of all connections reached organizations whom DCLI never directly engaged (highlighted in blue).

partnership, capacity building, etc.). **Fifty of them, or 42% of all connections in the system, reached organizations whom DCLI never directly engaged** (see [Figure 2](#)). The remaining 48 relationships were directly with DCLI.

Even at this early stage of analysis, this data paints a remarkable picture about the nature of DCLI’s systemic impact. DCLI, and systems approaches more generally, look towards multiplier and network effects to validate whether skills and assets transfer organically between local actors, which is an indicator of meaningful systems change.<sup>12</sup> Looking at the DCLI system about a year after its programming concluded, we see that **over half—59%—of all connections in the system were forged between organizations without DCLI’s direct involvement**. In other words, on average, each relationship sparked by DCLI resulted in at least one other relationship being formed as a result. This is a sign of organic strengthening of Tanzania’s data-enabled ecosystem.<sup>13</sup>

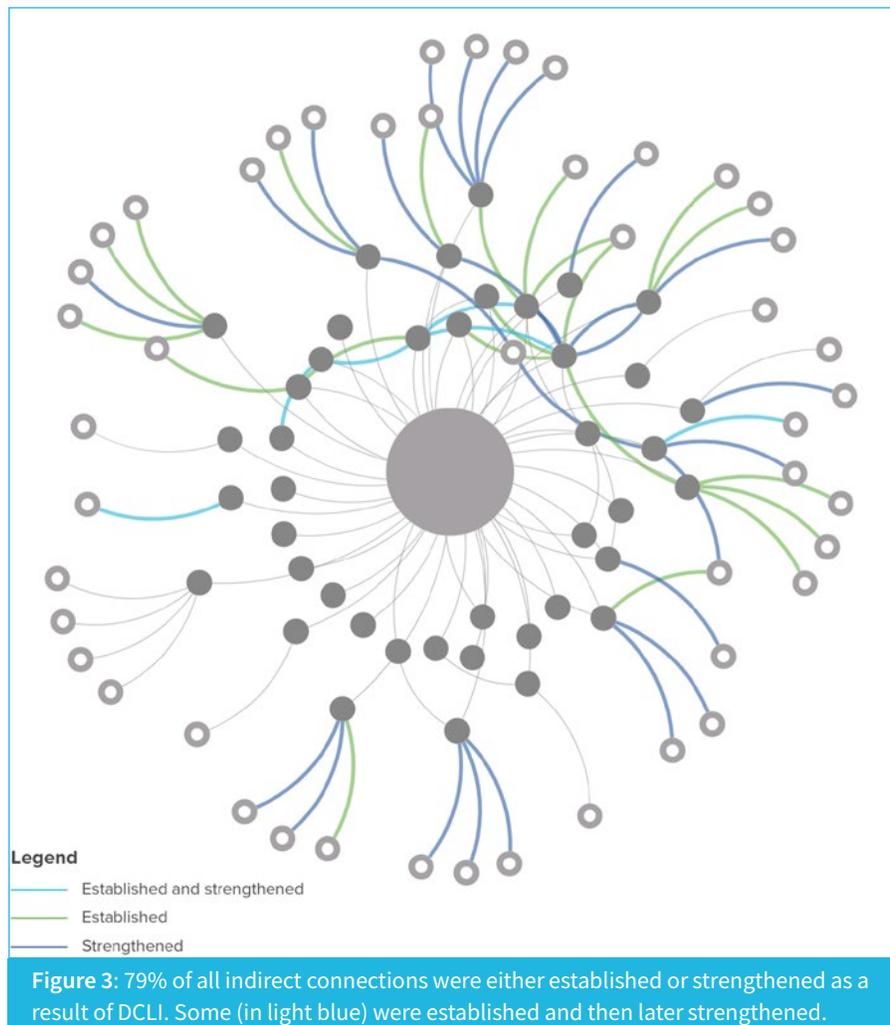
<sup>12</sup> See “Systems Thinking: An Introduction for Oxfam Staff” (Oxfam, 2015)

<sup>13</sup> The SNA survey did not collect information about whether the nature of these new relationships, formed between primary and secondary beneficiaries, was itself related to data use. More importantly, this distinction doesn’t matter to DCLI. The fact that this relationship, in whatever sector it may be, was reported by respondents to be a result of their engagement with DCLI is sufficient.

Our SNA survey dataset allows us to dive deeper into the relationships of these indirect connections (for the remainder of this analysis “indirect” connections refer to those 70 relationships which did not directly involve any DCLI project).

**79% of these indirect connections were either strengthened or established as a result of DCLI’s activities** (see Figure 3). Respondents reported *strengthened* relationships with nearly half of these connections (34). Perhaps more strikingly,

DCLI’s investments led to the *establishment* of 27 new relationships, meaning that **more than one in five relationships in the entire map was created organically as a result of DCLI’s investments**. Six out of all relationships were both established and strengthened—but only two such relationships exist among indirect connections.<sup>14</sup> Future similar efforts could try to more intentionally foster new indirect relationships that strengthen over time; for example, skills-building efforts conducted over time could request or reward direct beneficiaries who invite indirect ones to subsequent trainings.



<sup>14</sup> An example of this would be IREX, who both established a new relationship with an implementing partner while also, later on and via another touchpoint with DCLI, strengthened that relationship.

These findings are important because DCLI was intentionally designed to inject resources into an existing system, in order to foster organic growth between active and latent actors after the program ends. DCLI's theory of change was aligned with the systems change concept of network effects: as more people join and contribute to the same network, they gain more value from it, making that system as a whole more productive and sustainable.<sup>15</sup>

In this case, seeing how DCLI's investments are leading to or strengthening relationships between system actors suggests that the foundations are being laid for more value to be drawn from that system, as a result of more actors taking part in that system. These network effects might be particularly relevant to the context of a data use ecosystem, which relies on the availability of data to generate value by and for data users (commonly referred to as increased "demand" for data, which was an intention of the DCLI program). So, as more actors are organically brought into that system, more data is made available (through collaborative data collection efforts, data sharing agreements, etc.) that in turn contribute new value to that growing system. Although this SNA doesn't answer the extent to which that value is being realized, it does lay the foundation for a similar exercise in the future that focuses on that research question specifically. DCLI, and MCC programming in general, could therefore leverage the SNA approach as both an assessment and evaluation tool to understand how network effects are boosting the impact of their programs.

To conclude this initial discussion, let's revisit our research question: How did DCLI investments in Tanzania create or strengthen linkages between

#### Box 4: Managing duplicates in SNAs

Although our analysis reports 118 connections, a close count of the network visual will reveal fewer than 118 connections. Why is that?

Because SNA surveys are naturally about relationships between people, it's common for different people to report the same connection. This can pose challenges for the researchers, especially when—as in this case—questions were asked about those relationships to which respondents might answer differently. For example, respondents from two organizations might report differently about how their relationship evolved over time.

This can also be challenging when it comes to attributing specific traits to different organizations. For instance, the survey asked respondents to identify which DCLI project they engaged most closely. If two respondents, each identifying a different DCLI project, introduce the same secondary beneficiary organization, with what DCLI project should that beneficiary organization be tagged? They could be combined, but this could lead to disaggregated data about the three projects theoretically adding up to over 100% (since the same organization might be tagged to multiple DCLI projects).

Annex C (Methodology) explains more details about how the research team approached these potential overlaps and redundancies.

ecosystem stakeholders? Put simply, DCLI's investments in Tanzania led to the establishment or strengthening of 55 organizational relationships. To put that number in context, that's more than the number of relationships that DCLI itself established through its actual investments (48). In other words, **DCLI's direct beneficiaries created or strengthened more relationships (as a result of DCLI) than DCLI itself did.**

#### R2: Which organizations produced the most new or strengthened relationships?

Aggregate data, like the numbers shown thus

<sup>15</sup> Although the term originated in the technology sector, over the past decade there have been more efforts to amplify network effects in global development contexts. See for example "Network Effects of the Productivity of Infrastructure in Developing Countries" (World Bank, 2006).

far, are useful to understand at a macro level the evolution of Tanzania’s data-enabled system. But systems are composed of individual actors, so we can gain an additional depth of understanding by looking at some of them in detail. Specifically, a few outliers stand out for the breadth and depth of their interactions with other organizations that resulted from DCLI.

Three organizations—Pact Tanzania, Twende, and Mbeya Highlands FM (see Figure 4)—*strengthened* relationships with four others (the maximum allowed by the survey).<sup>16</sup> Mbeya Highlands FM, a community radio station operating in Mbeya Region, western Tanzania (home to two of the PEPFAR priority districts targeted in DCLI’s Data Zetu project), is a notable example, because

Mbeya Highlands also happens to be among the highest-ranked organizations on numerous social network analysis metrics, including betweenness centrality (which measures how many times an element lies on the shortest path between two other elements). Actors with high betweenness can play a role in controlling or sharing information flows in a system, although, given the low complexity of this SNA, we recommend caution before reading too heavily into this and other metrics (see Box 5).

Five organizations *established* relationships with at least three organizations as a result of DCLI: IREX, Sahara Ventures, Tanzania Bora Initiative, Tanzania Passion for the Needy, and UNIQUE. The first three were DCLI implementing partners

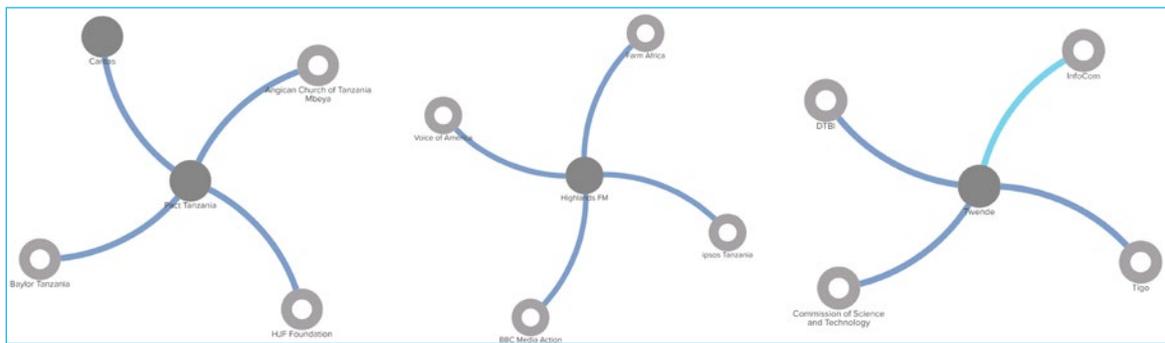


Figure 4: Three organizations strengthened relationships with four other institutions as a result of DCLI.

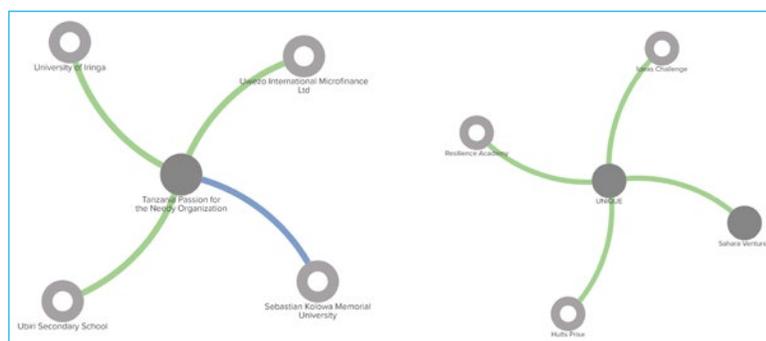


Figure 5: UNIQUE and Tanzania Passion for the Needy established relationships with at least three institutions.

16 See Annex C - Methodology (“survey instrument design”) to learn more about this limitation.

who worked together for the first time as a result of DCLI, so we would naturally expect them to report more established relationships. UNIQUE stands out because they reported establishing relationships with four other organizations (the maximum permitted in the survey) as a result of DCLI.

Although the content or purpose of these new relationships isn't clear based on the survey design and data, looking at UNIQUE's four new organizational connections (see [Figure 5](#)) does provide some hint as to the nature of those new interactions. Hults Prize, Resilience Academy, Ideas Challenge, and Sahara Ventures are organizations or initiatives that support technology innovation and social impact (in Tanzania and globally). UNIQUE engaged directly with DCLI's DLI project, which itself is a data innovation challenge. DLI's purpose was to strengthen Tanzania's innovation ecosystem by providing (among other assets) access to networks and other resources that can support innovation through financing, mentorship, etc. UNIQUE's snapshot, where they not only report these new relationships but also attribute them to DCLI, suggests that DLI's purpose was achieved.

Zooming in on a few organizations within the SNA system map offers additional depth of insights. The SNA visualization and metrics identify positive deviants, or outliers who (relative to others in the system) play an outsized role in amplifying DCLI's initial investments into Tanzania's data system. This has implications for MCC's broader approach to fostering systems change and embedding systems thinking in their programs. As an assessment tool, SNAs could help MCC to identify outliers—such as influential policy advocates or well-networked sector-specific experts—to engage early on in due diligence or program design. As an evaluative tool, SNAs could help those same programs

to understand whether more of these positive outliers exist after MCC's interventions compared to before, which would help MCC understand the extent to which its desired systems-level outcomes are achieved.

Highlighting the reported experiences of these selected organizations helps to answer our second research question (Which organizations produced the most new or strengthened relationships?). The third research question dives deeper into the nature of those relationships.

### Box 5: Making sense of SNA metrics

Digital tools to map and visualize networks offer researchers easy ways to access sophisticated metrics about the strength and diversity of those networks. Many SNAs lean on these metrics to inform interpretation and decisions.

These metrics can play a key role in supporting systems change evaluations or in designing systems change programs. For instance, **closeness centrality** is a measure of how close each element is to all other elements. Elements with higher closeness values are likely to be key linchpins in a system, perhaps as a broker of information or a network hub primed to spread information across a community. A program to support policies for improved energy access might use an SNA during the design or due diligence phase, in order to identify those actors which are best positioned to spread new regulations to key actors like energy providers or consumers.

Different SNA metrics are useful in different cases and depend on the research question. This report seeks to understand multiplier effects and secondary impact of DCLI's investments, so conventional SNA metrics like closeness and betweenness are less relevant, compared to other metrics like **outdegree**, which measures the number of connections an organization has to others. Sahara Ventures and Tanzania Bora Initiative—two local implementing partners of DCLI's Data Zetu project—have the highest outdegree values (11 and 8 respectively).

Depending on the focus of a program, MCC, PEPFAR, and others could use SNAs to identify institutions with high betweenness, centrality, or other metrics, and then intentionally engage them to spread change across a system.

### R3: What was the modality of these new or strengthened relationships? (New partnerships? Shared experiences or knowledge? etc.)

The DCLI social network analysis survey was distributed virtually using a brief online survey, so there was limited opportunity to investigate the specific nature of collaborations reported by respondents. That said, one set of questions invited respondents to clarify whether their engagement with another organization, resulting from DCLI's investments, were about:

- Partnering on new projects or proposals
- Sharing data sets or skills
- Exchanging financial resources
- Attending or organizing events together
- Staying in touch

These five options were not mutually exclusive (respondents could select multiple for each relationship), nor are they exhaustive. They do, however, represent key actions that the DCLI program believed would be indicative of desired change occurring within Tanzania's data use

system. For instance, the organic development of new partnerships between DCLI's direct and indirect beneficiaries would suggest *evolving business models*, or adaptations of DCLI's programming models. Sharing skills between organizations suggests indigenous investments into the *operational capacity* of Tanzania's data actors. Exchanging financial resources is indicative of increased *financial flows*. These indicators—improved or changing business models, operational capacity, and financial flows—individual- and collective-level indicators of systemic change,<sup>17</sup> so understanding whether DCLI's investments catalyzed these actions will help illuminate how, if at all, DCLI is imparting meaningful system change.

Of the 70 indirect connections in our SNA system map, respondents provided data about the nature of their connection for 60 of those relationships.

**60%, or 36, of these 60 indirect relationships were in the form of new partnerships or proposals** (see Figure 6). Additionally, two-thirds of those 36 were between DCLI's direct beneficiaries and indirect beneficiaries, while the other third of those new partnerships or proposals

<sup>17</sup> See "A Framework for Understanding Systemic Change" (USAID, 2016), pp.12-13.



A training at the dLab on data literacy skills. K15 Photos.

was between DCLI's direct beneficiaries. In other words, DCLI's investments spawned three dozen new collaborations between its direct and indirect beneficiaries.

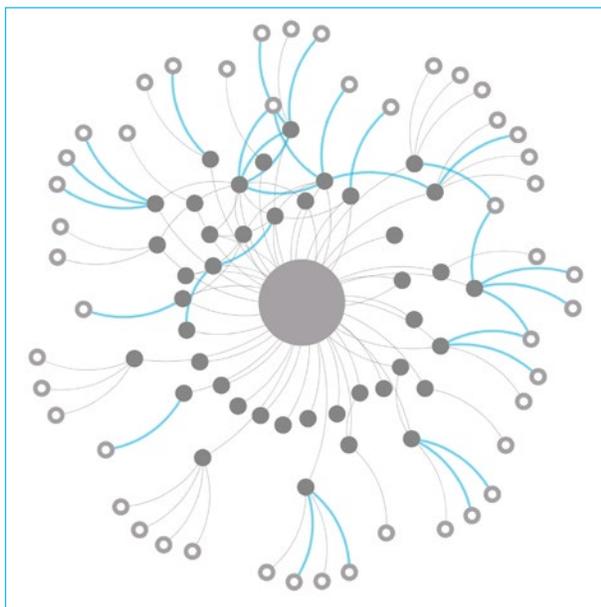


Figure 6: 60% of all indirect relationships were in the form of new partnerships or proposals.

**33 of those 60 relationships, or 55% of them, were about sharing data skills or datasets** (see Figure 8). About 30% of those dataset- or skill-sharing relationships were between DCLI's direct beneficiaries, while the remaining 70% occurred between DCLI's direct and indirect beneficiaries.

This indicator is particularly relevant to DCLI because of its commitment to building data use capacity and data availability as necessary preconditions to achieving the desired change in Tanzania's data ecosystem. **Over half of the secondary connections catalyzed by DCLI extended data skills or shared datasets with organizations with whom DCLI had no direct contact.** For example, the President's Office for Local and Regional Government (PO-RALG) shared data skills or datasets with the Ministry of Land, with whom DCLI never had direct contact (see Figure 7)

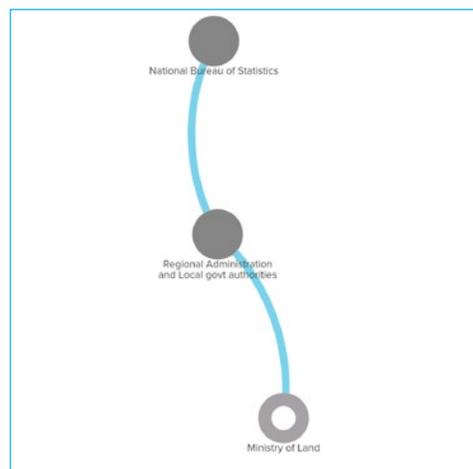


Figure 7: Tanzania's ministry for local government shared data skills or datasets with the Ministry of Land, with whom DCLI had no direct contact.

Moreover, **of the 42 direct beneficiaries of DCLI in this SNA map, one third shared data skills or datasets with other actors.** This is precisely the sort of multiplier effect that programs like DCLI, informed by systems thinking, seek to achieve (although there is room for improvement; DCLI could aspire to see more than one-in-three direct beneficiaries sharing their data skills or datasets with others).

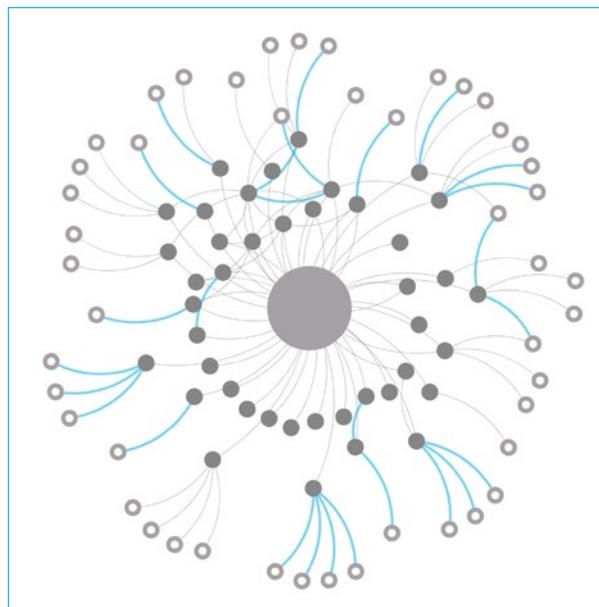


Figure 8: 55% of all indirect relationships involved sharing data skills or datasets.

Perhaps unsurprisingly, the portion of indirect connections that **exchanged financial resources** is low when compared to the other indicators discussed above. But it is not trivial; 14 of the 60 relationships—or almost one-quarter—involved the exchange of financial resources. As the visualization shows, the majority of these occurred between direct beneficiaries and indirect beneficiaries (see [Figure 9](#)). One organization, Data Zetu implementing partner Sahara Ventures, reported exchanging financial resources with three other organizations (see [Figure 10](#)). For donors like PEPFAR and MCC, this kind of information could be essential in identifying higher-capacity local organizations in a system that can be trusted to manage and disburse funds in-country.<sup>18</sup> This insight also speaks to the sustainability of the ecosystem; building initial connections is one thing, but being able to sustain and grow those relationships through continued financial partnerships

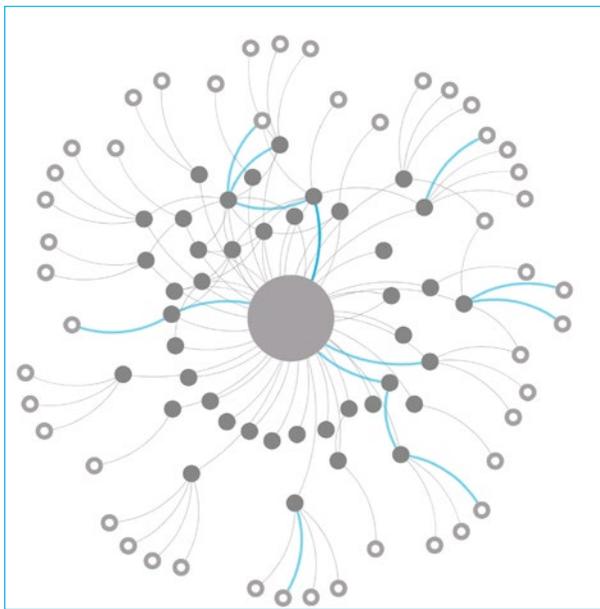


Figure 9: Almost one-quarter of all relationships involved the exchange of financial resources.

demonstrates a more sophisticated and self-reliant system.

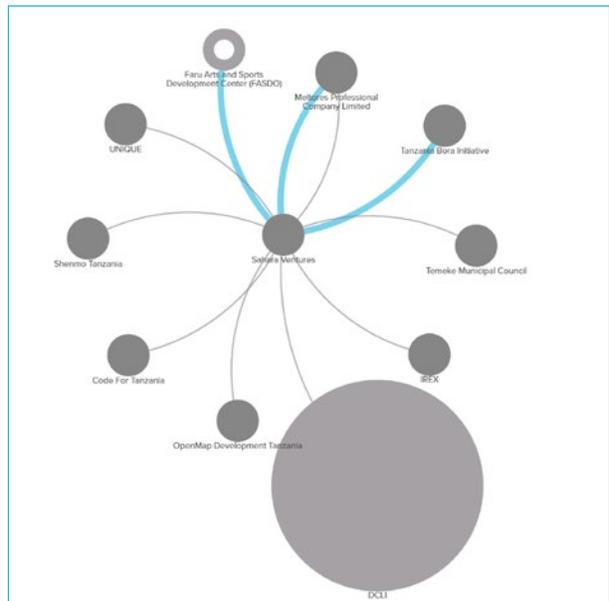


Figure 10: One organization exchanged finances with three others. This information could be useful for funders seeking high-capacity in-country partners.

The final two indicators—**staying in touch** and **attending or organizing events**—share similar patterns. Half of the 60 indirect connections attended or organized events together, while 34 stayed in touch. 57% of all relationships related to attending or organizing events took place between direct and indirect beneficiaries.

It's easy to get lost in the specific numbers, ratios, and indicators in this section, but what's most important is to understand how this analysis unveiled key activities taking place in the system—organically and without DCLI's direct involvement—that resulted from DCLI's investments. **Months, or even years, after originally engaging with DCLI, direct beneficiary organizations were sharing data skills, exchanging financial resources, and partnering on projects and proposals, as a result of DCLI.**

<sup>18</sup> This might be particularly relevant to organizations like PEPFAR, who want 70% of their funding to flow to local organizations by 2021.

Because those activities are indicative of systems change, this study demonstrates the utility of SNAs to help MCC programs capture the sustaining impact that their programs are designed to achieve. In some cases, the SNA can answer basic evaluative questions (for instance, we know that the data skills imparted by the DCLI to direct beneficiaries have transferred to indirect beneficiaries). In other cases, the SNA can help inform decisions in program design and partnerships. For instance, Tanzania Passion for the Needy (who support women, girls, and vulnerable communities) reported sharing skills/datasets and exchanging financial resources with four other organizations (see Figure 11); should they be more closely involved in designing programs meant to improve women’s economic empowerment?

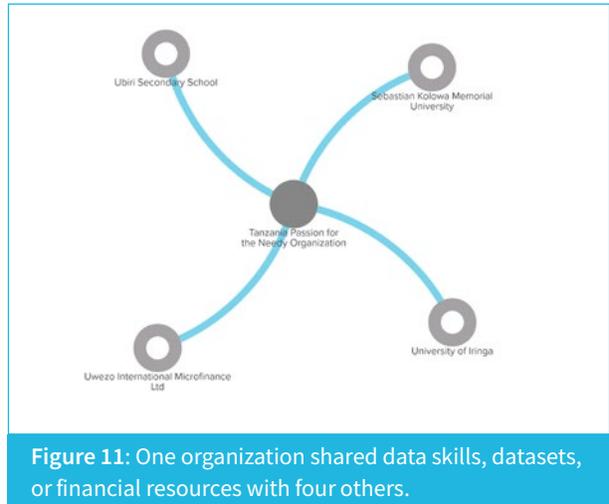


Figure 11: One organization shared data skills, datasets, or financial resources with four others.

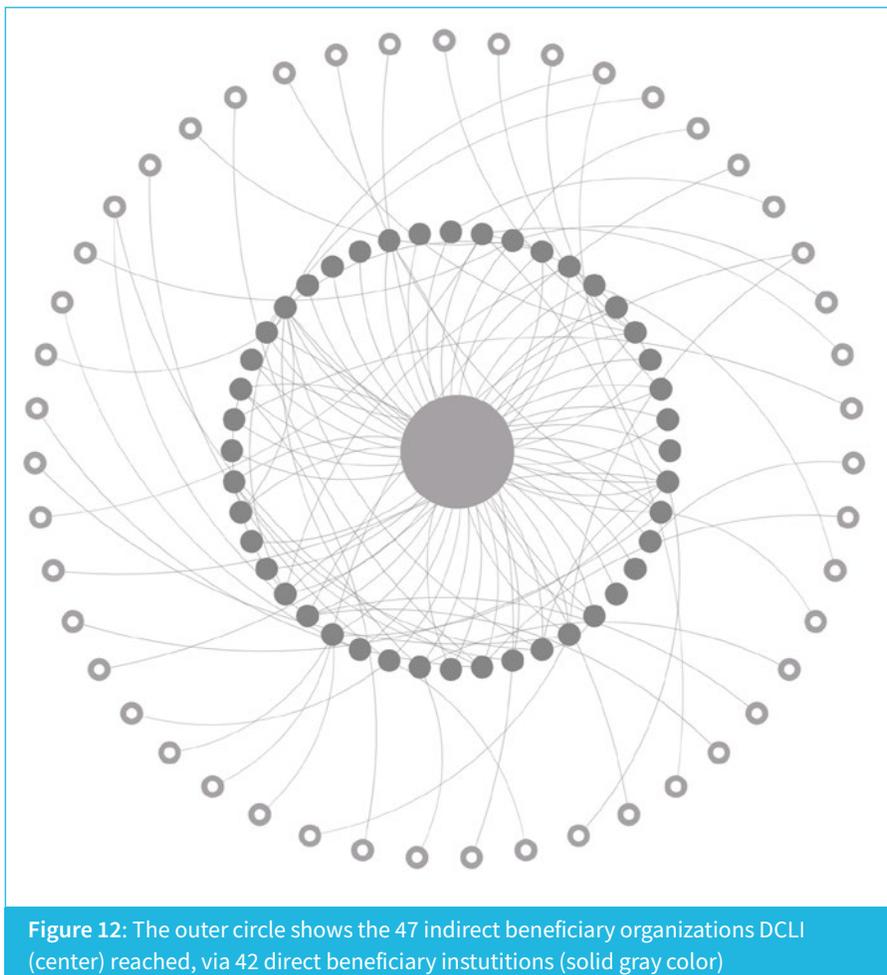


Figure 12: The outer circle shows the 47 indirect beneficiary organizations DCLI (center) reached, via 42 direct beneficiary institutions (solid gray color)

#### R4: What secondary relationships (that is, those beyond the scope of DCLI’s interactions with primary beneficiaries) were cultivated as a result of the DCLI investments?

Thus far our analysis has focused on the scope and nature of *connections* within Tanzania’s data ecosystem as reported by respondents of the SNA survey. But to understand our multiplier effect—a useful indicator from the perspective of investments into systems change—we can turn to look at the *organizations* themselves.

One way to understand transformative system change is to look for multiplier effects within these organizations.<sup>19</sup> Programs that result in multiplier effects can be considered more efficient than those which do not, because the return on investment expands as actors relay the outcomes of those investments (e.g. new skills) to others in their network.

In practice, multiplier effects can be defined in different ways. In most global development contexts, a multiplier effect exists when outcomes in one area manifest themselves or contribute to outcomes in another sector. For example, women who are healthier also tend to be better educated and participate more in formal economic activities.<sup>20</sup> For the purposes of the DCLI SNA, since no data was collected about the specific sectoral focus of indirect relationships, *a multiplier effect exists when a direct beneficiary engages another organization as a result of DCLI’s investments*. See **Box 6** for more details on how this was calculated.

This definition is sector-agnostic, because—from an impact evaluation perspective—we care less about *what* a collaboration is about and rather about *whether* DCLI’s investments contributed to that collaboration. This is aligned with DCLI’s

#### Box 6: 1.8 or 1.1? Calculating multiplier effects

How we choose to clean and analyze the data informs how we calculate the DCLI’s multiplier effect. The numbers offered in this research are based on survey response data, calculated by averaging the number of organizations that each respondent reports engaging as a result of DCLI.

But it’s possible that multiple respondents represent the same organization. It’s also possible that some of the organizations reported by a respondent might already be direct DCLI beneficiaries. Or, some of those organizations could be the same as those reported by another respondent.

For example, consider if two respondents from organization A report engaging organization B. Even if those were separate engagements, a social network map would treat this as one connection between two institutions. Should this be counted as one instance of the multiplier effect in action, or two?

There are pros and cons to adjusting for these potential duplicates or overlaps. On one hand, the numbers reported here represent productive engagements that were unanticipated by DCLI—even though they might be with organizations whom DCLI had already engaged, or separate engagements between the same two organizations.

An alternative method to calculate the multiplier effect would be to consider organization-to-organization engagements between indirect and direct beneficiary organizations. This is perhaps the most strict and conservative estimate of the multiplier effect in DCLI’s system based on this SNA. In this case, DCLI directly engaged 42 organizations, who in turn engaged 47 other organizations otherwise untouched by DCLI—**suggesting a multiplier effect of 1.1**.

Whether we use 1.1 or 1.8 as our multiplier effect, what’s clear is that DCLI’s reach extended beyond its immediate direct beneficiaries. A conservative calculation of the multiplier effect suggests that DCLI’s reach more than doubled, since its 42 direct beneficiaries engaged 47 other institutions as a result of DCLI.

approach to systems thinking in the context of data use ecosystems, recognizing that improved data use, experience, or perceptions can have ripple effects in sectors ranging from health and agriculture to education and governance.

<sup>19</sup> See “Supporting Transformational Change for Poverty Reduction and Shared Prosperity” (World Bank, 2016), p.70.

<sup>20</sup> See “To build a brighter future, invest in women and girls” (World Bank, 2016)

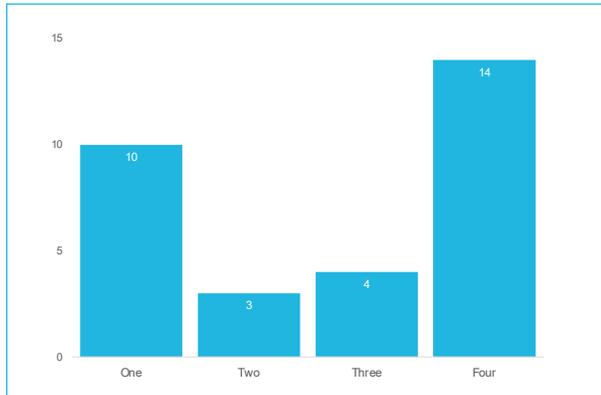


Figure 13: Fourteen respondents reported engaging at least four other organizations as a result of DCLI.

DCLI’s social network analysis reveals some fascinating insights about the program’s multiplier effect. Of the 47 respondents, **16 (34%) reported engaging with no other organization.** This means that it is likely that one in three organizations whom DCLI engaged did not extend the investments that DCLI made into them onwards to other organizations.

The story is far different for the other two-thirds of respondents. 31 respondents reported engaging at least one other organization, while **14 (30%) engaged with four other organizations as a result of DCLI’s investments** (see Figure 13).

A visual representation of these patterns demonstrates the expanding reach of DCLI via its direct beneficiaries to its indirect ones (see Figure 12).

On average, respondents reported engaging 1.8 other organizations as a result of their interactions

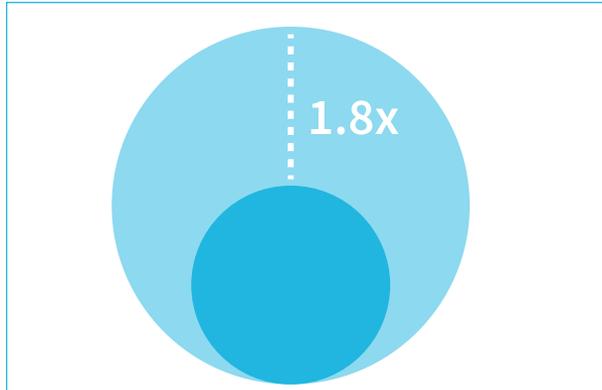


Figure 14: DCLI’s multiplier effect is up to 1.8x, meaning that every organization it directly reached in turn engaged 1.8 others, on average, as a result of DCLI.

with DCLI, **suggesting a multiplier effect of 1.8.** This multiplier effect varies among the individual DCLI projects. Respondents associated with the Data for Local Impact Innovation Challenge engaged 1.4 other organizations on average, dLab respondents engaged 1.7 others, and Data Zetu engaged 2 others.

The disaggregated data also reveals variations of the multiplier effect based on the location of the respondent. Respondents from organizations based in Dar es Salaam (a main urban center) engaged 1.6 other organizations as a result of DCLI, while respondents from outside of Dar es Salaam (rural and semi-urban areas) reported engaging 2.1 other organizations.<sup>21</sup> This means that **DCLI’s multiplier effect is larger outside Tanzania’s urban areas**—a particularly salient indicator given DCLI’s intentional commitment to strengthening data use systems in local and subnational communities.

<sup>21</sup> Dar es Salaam is not the only urban area with which Tanzania’s identified. The trend mentioned in this paragraph remains about the same when also including Dodoma respondents as “urban”.

## 4 | Conclusions

### UNDERSTANDING DCLI'S IMPACT: NEW TOOLS PROVIDE COMPLEMENTARY INSIGHT

This social network analysis unveiled some striking insights about DCLI's legacy in Tanzania's data and health ecosystem. It surfaced key insights about the system-level impact and reach of DCLI's investments which sustained or took place long after DCLI's programs concluded. Perhaps most salient is the fact that **over half—59%—of all connections in DCLI's system were forged between organizations without DCLI's direct involvement.**

From a systems thinking perspective, seeing so many relationships sprout as a result of DCLI's investments, suggest that at least some components of DCLI's programming remain driven and sustained locally. But this SNA alone is not sufficient to evaluate the success of DCLI's systems change strategy. As mentioned in the introduction, SNAs are just one of several tools to analyze how systems are evolving. DCLI has also commissioned an independent evaluation designed to map, visualize, and analyze the DCLI system based on qualitative secondary and first-hand research. This independent evaluation, combined with this SNA and other impact

measurement efforts (such as the qualitative [impact stories](#) available on the DCLI website) work in tandem with each other to help crystallize the picture of DCLI's impact.

These complementary evaluation methods can help to corroborate or challenge each other and inspire further investigations or M&E efforts. For example, through DCLI's traditional indicator tracking table (ITT), the DCLI team knows that its Data Zetu project trained 1,224 people on data literacy skills, who in turn reported training 874 people—that is, indirect beneficiaries with whom DCLI had no direct contact. This multiplier effect is 1.4, which rests right between two estimates offered in this SNA report (1.1 and 1.8). Although the multiplier effect based on self-reported sharing of skills in the ITT is not the same definition as the one used here (based on self-reported engagement with organizations), the fact that both numbers are roughly similar gives comfort to the veracity of these insights. More importantly, this stresses the importance of *intentionally deploying diverse and complementary evaluation tools to meaningfully capture systems change.*

## IMPLICATIONS OF SOCIAL NETWORK ANALYSES FOR FUTURE PROGRAMMING

As MCC, PEPFAR and others refine and expand use of systems thinking to design, implement, and evaluate programs, social network analyses are a useful tool to inform that process. The SNA methodology can:

- Enable more frequent and rapid assessments (for instance, using online surveys like this research did) to, throughout a program's multi-year life cycle, better engage and coordinate with its beneficiaries, as well as track how their lives are changing as a result of investments.
- Through lighter, more rapid assessments, build adaptive management and programming flexibility into development investments. DCLI worked to adapt to changing systems and activities during implementation, but may have benefited from additional information such as SNA outputs during these periods of adaptation.
- Inform how MCC and other agencies plan (and budget accordingly) for evaluation approaches to measuring the impact of programs, by including resources to and engaging actors skilled in conducting these types of evaluations.
- Provide a blueprint for measuring long-term impact in time-bound programs with relatively short implementation timelines, just as DCLI used an SNA to track the impact of its short performance period.
- Serve as a relatively lean baseline tool during compact development and due diligence processes, to map how relationships between actors change over time.
- Demonstrate how partnerships with

other funding agencies can rely on social network analyses or systems evaluations to capture complex impact in sectors that matters to both partners. For example, DCLI aimed for short-term impact around improved evidence-based decision-making and accountability (aligned with MCC and PEPFAR's goals), as well as longer-term impact on improved health services (which are the priority of PEPFAR, DCLI's funding partner). An SNA can capture these downstream impacts well.

To make these possibilities a reality, MCC, PEPFAR, and others can invest in building internal familiarity with and capacity to oversee or implement systems analysis methods like SNAs. These institutions could also plan and fund adaptive monitoring, evaluation, and learning (MEL) plans that include flexible budgets for SNAs, systems mapping, or other emerging evaluative approaches to help capture unanticipated systems impact.

## SYSTEMS THINKING IN MCC AND PEPFAR PROGRAMMING

This social network analysis unveiled specific examples of DCLI's investments in Tanzania being **locally driven** and **sustained**. These findings were made possible by the SNA, but it was DCLI's intentional design from the outset to embrace systems thinking that enabled these indigenous outcomes. DCLI's projects focused not only on individuals (such as building skills via data literacy trainings), but more importantly on local organizations (such as investing in data startups, or supporting hospitals and local government offices).

This type of human-centered and systems-minded design applies much more widely than to human or community development initiatives.<sup>22</sup>

<sup>22</sup> See "[Human-Centered, Systems-Minded Design](#)" (Stanford Social Innovation Review, 2018)

Infrastructure investments<sup>23</sup> and policymaking<sup>24</sup> are two areas where systems thinking can also play an important role. Donors could consider complementing or pre-empting infrastructure projects with programs like DCLI, which fuel in-country change and lay the foundations for those projects' successful adoption and local ownership.

As MCC, PEPFAR, and others continue to grapple with complex operating environments and

interrelated development challenges, designing programs with a systems lens is paramount. Efforts like those of DCLI, which targeted discrete investments into local systems and institutions for the explicit purpose of injecting assets and approaches that will be locally driven, can offer a blueprint for other programming effort—both in terms of embracing a systems design lens and adopting complementary evaluation approaches like SNAs to measure strengthened systems.

<sup>23</sup> See for example “Systems thinking for infrastructure professionals” (Institute of Civil Engineers, 2017)

<sup>24</sup> See “Systemic thinking for policy making” (OECD, 2019)



A community of Tanzanian and global data users has emerged at the dLab. *K15 Photos*

## Complete this 5-minute survey for a chance to **win Tsh 100,000/=** of phone credit!

If you complete this entire 5-minute survey, you will be entered into a lottery to win Tsh 100.000/= phone credit. Some terms and conditions apply. **Click Next to continue.**

### **Terms and conditions:**

- You must complete the entire survey to be eligible for the lottery.
- One respondent out of all respondents will be selected at random in the lottery.
- Individuals who respond multiple times to the survey will be counted as one single respondent during the lottery.
- The lottery winner must respond to a text message informing them of this prize within 24 hours, or the lottery will be re-run.
- Respondents who submit false or incorrect information, or information meant to deliberately skew results in any way, will not be eligible for the lottery.

# Annex A | Survey tool

**Please read this important information** and click Next when you are done reading it to begin the survey.

This survey will help the dLab team map relationships among organizations in Tanzania's data ecosystem. This data will be used to understand how relationships have changed since the start of the Data Zetu, dLab, and Data for Local Impact Innovation Challenge (DLI) projects.

The analysis based on this survey will be presented back to participants, and a summary will be posted online. This will include the names of specific organizations that you provide, so the survey should not be considered fully anonymous. However, we will always refer to organization names rather than specific respondent names; for example, we would refer to the answer from "IREX" rather than the person who responded on behalf of that organization. No personally identifying information will be shared.

**This survey should take no more than 5-10 minutes.**

*Please note:*

- Participation in this survey is completely voluntary. There will not be any negative consequences if you decide not to respond.
- If you decide not to participate, you may stop participating at any time.
- By submitting this form you are indicating that you have read and agree to the language on this page, and that you are at least 18 years old.

# Annex A | Survey tool

\* 1. What is your name?

\* 2. What is the name of your organization?

\* 3. What is your phone number?

4. What is your email address?

This is optional but must be provided if you would like to receive a summary of the analysis.

\* 5. Where are you residing?

\* 6. With which project did your organization **interact the most**?

- Tanzania Data Lab (dLab)
- Data Zetu
- Data for Local Impact Challenge (DLI)
- I'm not sure

# Annex A | Survey tool

\* 7. Did **{{Q2}}** engage with any other organizations, as a result of its interactions with DLI, Data Zetu, or dLab?

If so, please write the name of ONE organization in the space below. If not, you may close this survey.

*"Engage" can include hosting or attending events together, holding meetings together, partnering on funding opportunities, sharing data with each other, sharing data skills, or other meaningful interactions.*

In addition to that organization, **are there any other organizations that **{{Q2}}** engaged as a result** of its interactions with DLI, dLab, or Data Zetu?

If so, please provide their names in each space below. Although these questions are optional, we encourage you to add up to three more organizations. Click "Next" once you are ready to proceed.

8. Please enter the name of ONE OTHER organization **{{Q2}}** engaged as a result of its interactions with DLI, Data Zetu, or dLab.

9. Please enter the name of ONE OTHER organization **{{Q2}}** engaged as a result of its interactions with DLI, Data Zetu, or dLab.

10. Please enter the name of ONE OTHER organization **{{Q2}}** engaged as a result of its interactions with DLI, Data Zetu, or dLab.

# Annex A | Survey tool

\* 11. Which of these options best describes the **nature of the engagement that {{Q2}} had with these organization(s)?**

Please limit your response to any engagements that occurred as a result of {{Q2}}'s interactions with DLI, Data Zetu, or dLab.

*You may select multiple options for each organization.*

|         | Attended or organized events together | Exchanged financial resources | Partnered on projects or proposals | Shared datasets or data skills | Stayed in touch (in person, over the phone, etc.) |
|---------|---------------------------------------|-------------------------------|------------------------------------|--------------------------------|---|
| {{Q7}}  | <input type="checkbox"/>              | <input type="checkbox"/>      | <input type="checkbox"/>           | <input type="checkbox"/>       | <input type="checkbox"/>                          |
| {{Q8}}  | <input type="checkbox"/>              | <input type="checkbox"/>      | <input type="checkbox"/>           | <input type="checkbox"/>       | <input type="checkbox"/>                          |
| {{Q9}}  | <input type="checkbox"/>              | <input type="checkbox"/>      | <input type="checkbox"/>           | <input type="checkbox"/>       | <input type="checkbox"/>                          |
| {{Q10}} | <input type="checkbox"/>              | <input type="checkbox"/>      | <input type="checkbox"/>           | <input type="checkbox"/>       | <input type="checkbox"/>                          |

\* 12. Before interacting with DLI, dLab, or Data Zetu, **how frequently** did {{Q2}} engage with these organization(s)?

|         | Never                 | Once or twice a year  | Once or twice a month | Once or twice a week  | At least once or twice a day |
|---------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------------|
| {{Q7}}  | <input type="radio"/>        |
| {{Q8}}  | <input type="radio"/>        |
| {{Q9}}  | <input type="radio"/>        |
| {{Q10}} | <input type="radio"/>        |

\* 13. After interacting with DLI, dLab, or Data Zetu, **how frequently** did {{Q2}} engage with these organization(s)?

|         | Never                 | Once or twice a year  | Once or twice a month | Once or twice a week  | At least once or twice a day |
|---------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------------|
| {{Q7}}  | <input type="radio"/>        |
| {{Q8}}  | <input type="radio"/>        |
| {{Q9}}  | <input type="radio"/>        |
| {{Q10}} | <input type="radio"/>        |

# Annex A | Survey tool

\* 14. Were relationships between {{Q2}} and the organizations you identified **established or strengthened**, as a result of its interactions with DLI, Data Zetu, or dLab?

|         | Relationship was established | Relationship was strengthened | Relationship was neither established nor strengthened |
|---------|------------------------------|-------------------------------|---|
| {{Q7}}  | <input type="checkbox"/>     | <input type="checkbox"/>      | <input type="checkbox"/>                              |
| {{Q8}}  | <input type="checkbox"/>     | <input type="checkbox"/>      | <input type="checkbox"/>                              |
| {{Q9}}  | <input type="checkbox"/>     | <input type="checkbox"/>      | <input type="checkbox"/>                              |
| {{Q10}} | <input type="checkbox"/>     | <input type="checkbox"/>      | <input type="checkbox"/>                              |

# Annex B | Anonymized survey results

Anonymized survey results are available online as a CSV file at [www.bit.ly/dcli-sna-data](http://www.bit.ly/dcli-sna-data).

# Annex C | Methodology

## RESEARCH QUESTION

This effort began with a broad research question which shaped the data collection and analysis process:

**R1: How did DCLI investments in Tanzania create or strengthen linkages between data ecosystem stakeholders?**

Several follow-on inquiries naturally cascade from this research question, such as:

**R2: Which organizations produced the greatest number of new or strengthened relationships?**

**R3: What was the modality of these new or strengthened relationships? (New partnerships? Shared experiences or knowledge? etc.)**

**R4: What secondary relationships (that is, those beyond the scope of DCLI's interactions with primary beneficiaries) were cultivated as a result of the DCLI investments?**

Collectively, these questions can help illuminate how programmatic inputs and short-term outcomes cascade within the ecosystem—thereby offering meaningful insight into the medium- and long-term impact of the change process catalyzed by DCLI's investments in Tanzania.

## SURVEY INSTRUMENT DESIGN

An online survey was delivered to collect information that could answer the research questions. Although using an online survey can preclude respondents with low internet access, or who don't want to spend their available data

credit on a survey, the targeted respondents were already known by the DCLI program teams to have, in general, a sufficient rate of internet access and digital literacy to engage with an online survey. The survey was distributed as a web link that could be opened on any internet-connected device.<sup>1</sup> The online survey was developed and delivered in English, with local input by the dLab suggesting that the vast majority of potential respondents would have sufficient English literacy.

The survey was designed to take no more than 5-10 minutes to complete (this duration was tested and affirmed during survey pre-tests delivered to staff members at the dLab). The survey's preamble provided basic information about the survey, its purpose, the rights of the respondents (including their right to anonymity), and how the data would be used.

The data collection portion of the survey was structured as follows:

- Respondent provides demographic information
- Respondent identifies up to four organizations they engaged as a result of their own organization's interaction with any DCLI program.
- Respondent answers questions about the nature of those engagements. These questions are repeated up to four times, depending on how many organizations the respondent identified in the previous phase.

The second segment restricted respondents to

---

<sup>1</sup> The survey link shared to potential respondents was <http://bit.ly/tzdatanetwork>. It is no longer active since the survey is closed.

# Annex C | Methodology

submitting a maximum of four organizations. This has obvious implications for the analysis, because some respondents might have engaged more than four organizations as a result of DCLI. This tradeoff was deemed necessary for two reasons. First, the off-the-shelf survey tool had structural and design limitations, such that inviting more than four organizations would lead to visually confusing question formats later in the survey. Second, the survey was tested by stakeholders of the dLab, who had familiarity with the desired respondents and suggested that up to four organizations would be a reasonable limit, based on their experience of DCLI implementation.

The third segment investigated the nature of the engagements that DCLI beneficiaries had with other organizations. Specifically, the survey asked about:

- What kinds of activities did beneficiaries engage other organizations in, as a result of beneficiaries' engagement with DCLI? (such as writing proposals together, or sharing skills)
- How frequently did beneficiaries engage those organizations, before and after engaging with DCLI?
- Were relationships between beneficiaries and other organizations strengthened or established, as a result of beneficiaries' engagement with DCLI?

A limitation to a remote survey, intentionally designed to take only a few minutes to respond in order to increase response rates, is that the researchers were not able to collect qualitative insights about the nature of the activities between

organizations. For instance, we might know that two organizations shared data skills with each other, but with what impact? Through what modality? These limitations are part of this design but could easily be addressed in future iterations which could include follow-on interviews with specific survey respondents.

Finally, given the fundamental research question's interest in how relationships evolved between organizations, individual respondents were asked to respond on behalf of their organization, not on behalf of their personal experiences.

The full survey is available for reference in [Annex A](#).

## RESPONDENTS AND INCENTIVIZING RESPONSES

This survey had a specific, targeted audience: individuals at institutions who directly engaged with any DCLI project at any point throughout the program's life span. Individuals were not required to respond and opted-in only if they wanted to, and if they received the link. Thus, the sampling method used was a combination of a purposive sample (we directed the sample towards individuals who shared the common trait of having engaged with a DCLI program) and a voluntary sample (of those who share that common trait, only those who opted-in to the survey were included in the final sample). Notably, this approach meant that our results are not statistically representative but are roughly proportionally representative to the total number of people engaged by each of the three DCLI programs (see table below).

# Annex C | Methodology

The survey link was sent via email to 3,615 individuals. This roster was configured by digitizing and combining all available participant sign-in sheets that were completed at all eligible activities of the three DCLI programs (dLab, DLI, and Data Zetu).

| Project                      | DLI   | dLab | DZ    |
|------------------------------|-------|------|-------|
| # of potential respondents   | 1,075 | 523  | 2,016 |
| Share of all potential resp. | 30%   | 14%  | 56%   |

It's noteworthy that this roster skews heavily towards the Data Zetu project, and somewhat towards the DLI project. Both of these projects had significant virtual or field-level engagements, making their reach more widespread than the dLab which is a physical space bound by geographical limitations of accessibility and reach.

Wary about survey fatigue in Tanzania's data ecosystem, an incentive structure was deployed to catalyze engagement and a higher response rate. All eligible survey respondents were entered into a lottery where one respondent, selected at random, would receive 100,000 Tanzanian Shillings (roughly \$50) of mobile data credit. This incentive was advertised in the email correspondence to the potential respondents, as well as in the survey's preamble.

The survey was issued and publicly available between late October and early November 2019. These dates are important, since most programmatic elements of the three DCLI programs had concluded nearly one full year prior.<sup>2</sup> As a result, this survey offered a rare

chance to capture long-term impact and changes.

## DATA SECURITY, PRIVACY, AND BIAS

Personally identifying information of potential respondents were digitized and shared in private Google spreadsheets. Data submitted by respondents was collected electronically and stored using SurveyMonkey's servers. Cleaned data, without personally identifying information but with non-PII responses, were uploaded to a private account on Kumu, the network visualization software used to analyze and visualize the network.

This cleaned data excluded 15 responses (23% of all responses) which were not able to demonstrate sufficient engagement with the DCLI program to warrant inclusion (these may have been individuals responding for the sole purpose of winning the incentive lottery). This cleaned data also aggregated responses by organization, not by individual respondent (see [Box 4](#)).

Both of these datasets were at times downloaded on secure internet connections to personal computers for analysis. Researchers were instructed to delete all copies of these datasets after their analysis.

The entire process of survey design, survey distribution, data collection, and cleaning undoubtedly introduced biases that impact the final result of this SNA. A summary of potential biases includes:

- Respondents who are not comfortable

<sup>2</sup> The exact duration between the program's conclusion and this SNA's data collection varies depending on the DCLI project. An earlier footnote offers more details.

# Annex C | Methodology

reading or comprehending English are excluded.

- Respondents unwilling or unable to use the internet to take the survey are excluded.
- Respondents skew towards Data Zetu beneficiaries because Data Zetu had an outsized representation on the roster of potential respondents.
- Respondents who did not check their emails during the survey delivery period are excluded.
- Respondents may have inflated or otherwise misrepresented their engagement with the DCLI program in order to be eligible for the lottery incentive.
- Respondents not comfortable sharing their personal information are excluded.
- Potential respondents who did not experience any strengthened relationships might not have responded at all.<sup>3</sup>

Other limitations warrant mention as well. Although every effort was made to provide clear and specific language, individuals may have employed different interpretations of key words when formulating their responses. For example, respondents were asked to report other organizations with whom they “engaged” as a result of their “interactions” with DCLI. Specific interpretations of “engagement” or “interaction” could vary by individual. To mitigate this concern, the survey forced respondents to categorize their

engagement with organizations, as well as their interaction with DCLI, to help the research team determine validity of their responses.

Another significant limitation of our method (and many other SNA approaches) is that there is no control group against which to compare the DCLI network (such as a similar network of actors in a different country which did not have any investments like DCLI), nor is there a baseline SNA against which to gauge change over time. These limitations mean that:

We don’t know how relationships would have been established or strengthened without DCLI’s engagement.

We don’t know if the changes we see in the data community are more or less than the changes we’d seen if there were no DCLI involvement.

These constraints do not mean we can’t draw meaningful conclusions about DCLI’s impact. For instance, let’s imagine that one finding is that a connection between organizations A and B was created as a result of DCLI. This fact is noteworthy and valuable, regardless of whether Organizations A and B might have created a relationship without DCLI, because the survey questions are designed such that respondents attribute their changes to DCLI (i.e., the survey questions themselves included the language “...as a result of your engagement with DCLI?”). Non-experimental designs that rely on self-reported change are not the most robust scientific efforts, but as a complementary evaluation approach they are still considered valid by global development actors

---

<sup>3</sup> This specific bias means that our analysis is more about *how* strengthened relationships were strengthened, instead of *which* relationships were or were not strengthened.

# Annex C | Methodology

and funders including USAID.<sup>4</sup>

## COLLECTED DATA

The survey had a low response rate (see [Box 3](#)). Of the 3,500+ potential respondents, 104—or just 2.9%—opened the survey link.<sup>5</sup> Of these, 63 people completed the survey, but 15 of those responses were deemed ineligible since they could not answer the validation question of “with which [DCLI] project did you interact the most?”.<sup>6</sup> This resulted in 47 valid responses—or 1.3% of all potential respondents. These 47 respondents represented 31 unique organizations.

Recall that the survey link was sent to over 3,500 potential respondents, but these were not evenly distributed across all the projects. We might expect the share of actual respondents’ affiliations with a DCLI project to roughly match the share of potential respondents from each project’s contact list. To test this, we calculated the share of actual respondents who are affiliated with each DCLI project.<sup>7</sup>

| Project                     | DLI | dLab | DZ  |
|-----------------------------|-----|------|-----|
| Share of potential resp.    | 30% | 14%  | 56% |
| Share of actual respondents | 21% | 23%  | 55% |

These calculations suggest that response rates from people affiliated with the dLab overperformed by 50% compared to what we

might have expected based on the share of potential respondents who came from dLab’s contact lists. In this respect, DLI is about 25% less represented than we would have expected, and respondents affiliated with Data Zetu roughly match what we might have expected.

## ANALYSIS, CLEANING, AND VALIDATION PROCESS

Duplicates exist in the response data. For instance, two different respondents might report a connection with each other’s organization, leading to two different connections in the dataset which are between the same two organizations. To manage this challenge, duplicates were addressed in the online visualization tool (Kumu), not in the raw or cleaned dataset uploaded to that tool. Kumu has the ability to recognize when the same connection is reported multiple times in the dataset, and Kumu only displays that connection as one connection. [Box 4](#) has more details.

This approach to managing duplicates poses challenges for presenting the data. For instance, the final system map might only display 50 connections, but the raw response data might have 60 connections, with 10 being duplicates.<sup>8</sup> Likewise, when respondents provided conflicted information about duplicate connections, the following steps were taken:

<sup>4</sup> See “Introduction to Impact Evaluation Design” (USAID)

<sup>5</sup> Acceptable response rates vary according to survey type and context. To illustrate the diversity: UN DESA generally seeks a minimum response rate of 60% for household surveys ([source](#), p. 223), while e-mail campaigns generally [achieve](#) between 1 and 10%.

<sup>6</sup> If a respondent is unable to identify any DCLI project with which they interacted the most, we assumed their DCLI interaction was of insufficient depth or rigor to validate for the purposes of this analysis.

<sup>7</sup> We use respondents’ response to the questions “with which [DCLI] project did you interact the most?” to determine their affiliation.

<sup>8</sup> A “duplicate” response is not necessarily a redundant response. If Org A reports an engagement with Org B, and vice versa, these are both valid responses and could be about separate engagements.

# Annex C | Methodology

- Responses were combined.
- If combining responses would lead to a logical contradiction, the more productive or positive response was retained.

For example, if Organization A reports that their connection with Organization B was “Neither strengthened nor established”, while Organization B says it was “Established”, the connection would be marked as “Established”.

For example, if Organization C reports that their frequency of interaction with Organization D increased, while Organization D said it did not, the connection would be marked as having increased in frequency.

Survey data analysis was conducted in December 2019 thru February 2020. This involved developing an initial version of early findings, to share with the SPACES team who were, concurrently, conducting an independent

evaluation of the DCLI program in Tanzania. This initial version was also shared as a prototype at a November 2019 event hosted by MCC in Washington DC.

These occasions served as early validation efforts of the analysis, but further validation will begin with the publication of this report and its findings.

## PARTNERS

This SNA was conducted in collaboration with, and with the help of, several institutions. The Tanzania dLab (the NGO that arose from one of the three DCLI projects in Tanzania) tested the survey, prepared rosters of potential respondents, and helped administer the survey electronically to potential respondents. SPACES provided external feedback and expertise to shape the survey, both strategically and tactically. MCC provided overall guidance and oversight on the process.

# Annex D | Terminology

Our methodology and analysis rely on some specific terms that warrant definition:

**Connections** (occasionally called “relationships”) link one organization to another. They derive from survey respondents who provide an organization name in response to the question “Did your organization engage with any other organizations, as a result of its interactions with DLI, Data Zetu, or dLab?”. If they provide a name, then a connection is deemed to exist between the respondent’s organization and the one they provided. These connections are attributed to DCLI’s investments, since the question asks respondents to only list those organizations with whom they engaged as a result of their interaction with DCLI. Duplicate connections are possible, such as when two different respondents both report connections to each other. Although these are technically two separate connections, this analysis treats them as one and the same connection.

**Organizations** represent the specific organization to which a respondent belongs. All organizations are connected by at least one connection, but some organizations can have multiple connections. For example, here is where two organizations (Tanzania Bora Initiative and Sahara Ventures) both independently reported engaging with FASDO. This distinction is important because it explains why the number of connections and the number of organizations aren’t the same. It also explains duplicates and some instances where percentages offered in the following pages don’t always add up to 100%; for instance, if TBI were associated with the Data Zetu project, and Sahara Ventures were associated with the DLI project, then FASDO would be counted as an

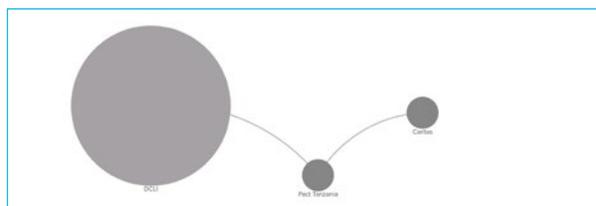
indirect beneficiary organization of both DLI and dLab:



**Engagement** with an organization was defined using the following options, which respondents could provide for each connection.

- Attending or organizing events together
- Exchanging financial resources
- Partnering on projects or proposals
- Sharing datasets or data skills
- Staying in touch (over the phone, in person, etc.)

**Direct beneficiaries** are those organizations who have a connection directly to DCLI, while indirect beneficiaries are those organizations who are connected only to direct beneficiaries. In some cases, an organization could simultaneously be an indirect and direct beneficiary. For example, here we can see that Pact Tanzania reported engaging with Caritas as a result of DCLI. Meanwhile, Caritas was also directly engaged by DCLI. For the purposes of this study, Caritas would be tagged as a direct beneficiary, not an indirect beneficiary, since they were directly engaged by DCLI:



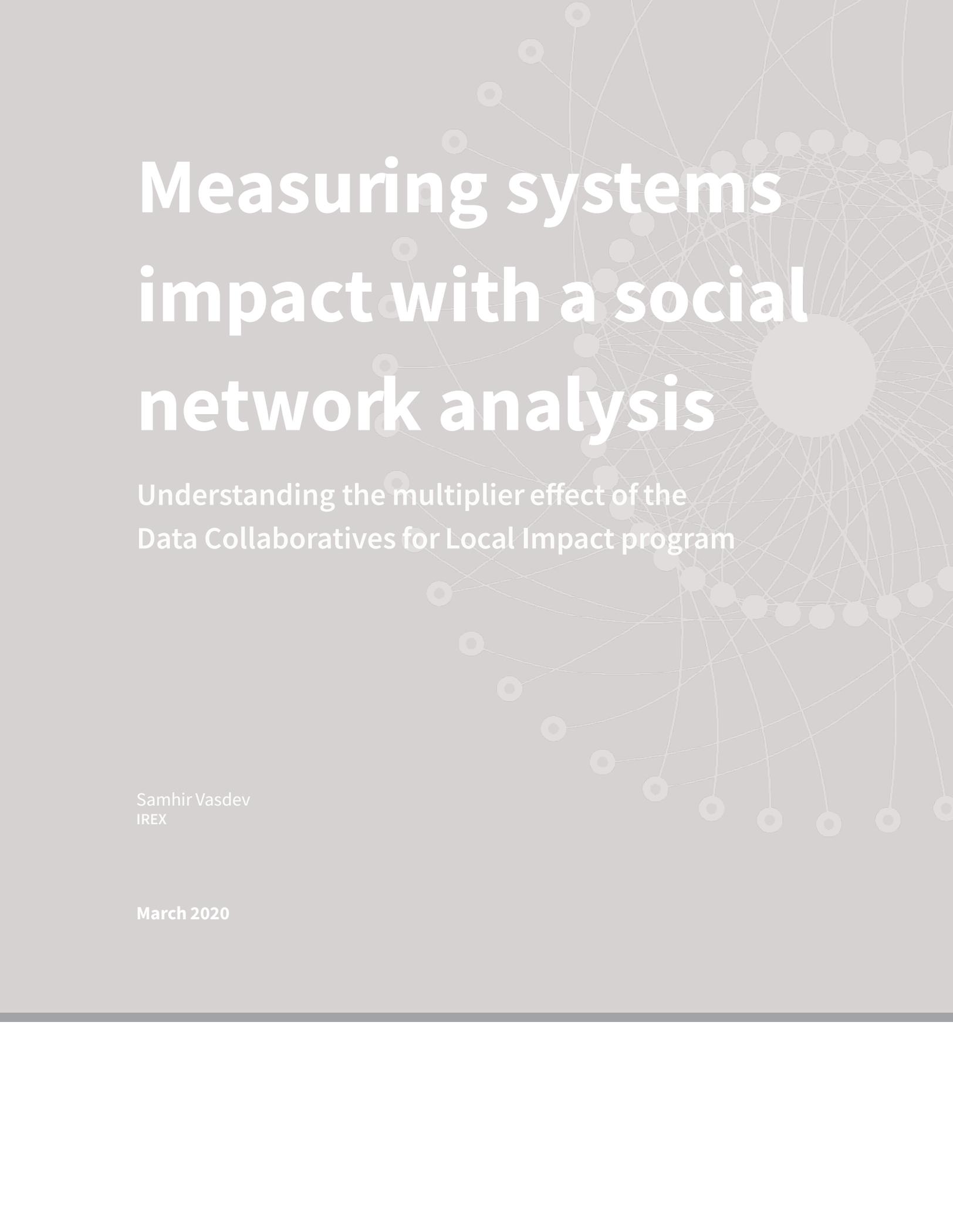
# Annex D | Terminology

Respondents were also asked to define whether their relationship with another organization was “**established**” or “**strengthened**” as a result of their interactions with DCLI. We did not provide definitions of these terms.

Finally, it’s worth a reminder that all findings presented in this report are **self-reported**.

Although we cannot guarantee that a connection that appears in our analysis really exists, we must

assume it does based on our chosen approaches to data collection and validation. For this reason, we omit terms like “reportedly” or “reported” in our findings and urge readers to remember that all findings are based on what beneficiaries reported to the research team.



# Measuring systems impact with a social network analysis

Understanding the multiplier effect of the  
Data Collaboratives for Local Impact program

Samhir Vasdev  
IREX

March 2020