2023 Data for Social Impact Report

Accelerate Aspirations: Moving Together to Achieve Systems Change

Danil Mikhailov, Ph.D.

Ginger Zielinskie

Perry Hewitt

Uyi Stewart, Ph.D.

data.org
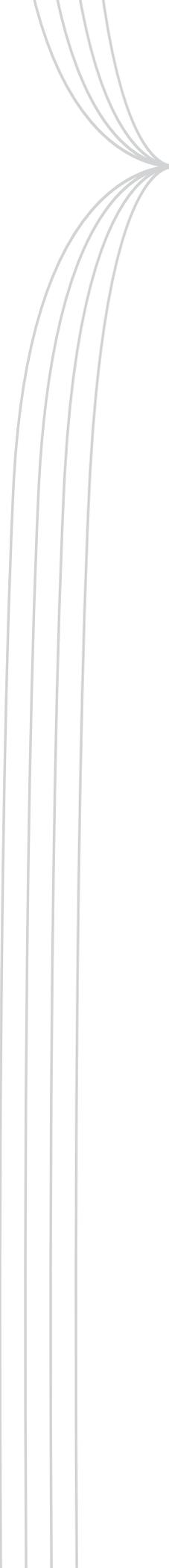
CONTENTS

ACKNOWLEDGMENTS

Cover image: Pratima Baral, researcher at the International Maize and Wheat Improvement Center (CIMMYT) leading a workshop with female farmers in Surkhet, Nepal. Photo by C. de Bode/CGIAR.
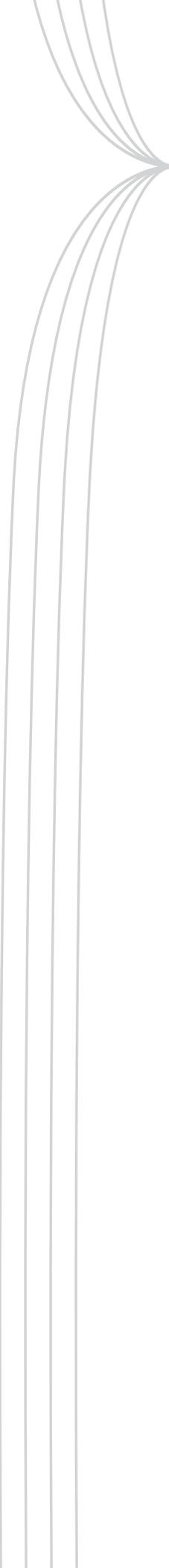
ABBREVIATIONS

AI   artificial intelligence

CBOs   community–based organizations

CEO   chief executive officer

CSO   civil society organization

DMA   data maturity assessment

DSI   data for social impact

IDEA   inclusion, diversity, equity, and accessibility

LMICs   low- and middle–income countries

MEL   monitoring, evaluation, and learning

NGOs   non-governmental organizations

SIOs   social impact organizations

STEM   science, technology, engineering, and mathematics
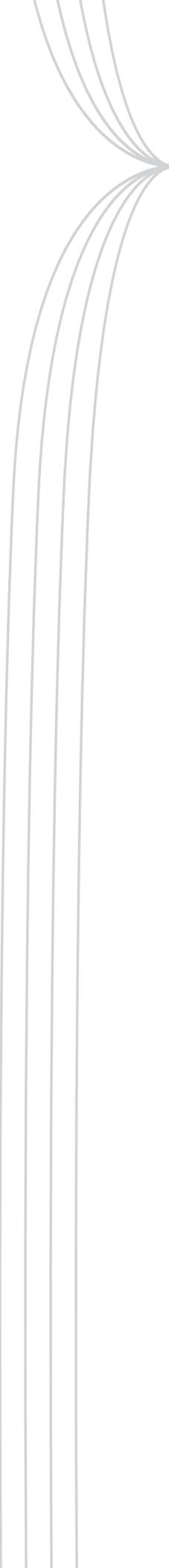
FOREWORD

Our world is at an inflection point. We face the convergence of historic challenges that demand historic solutions—shifting demographics, rising inequality, public health crises, widespread threats to democracy, and a climate emergency.

The good news? Everyday data and data-driven technologies are helping to accelerate positive change. They are creating high-impact, global interventions to help our planet and the humans that live here unlike ever before.

And as data collection technologies evolve, we continue to gather unprecedented volumes of information. Dramatic advances in artificial intelligence allow us to find meaning and patterns in vast data streams—data and information that's accessible to anyone with an internet connection at a fraction of the cost of past technical efforts.

Indeed, 2.5 quintillion megabytes of data are collected about people every day, yet a third of people worldwide live offline and are disconnected from the digital and data revolution. The speed of this change has raised urgent questions about who benefits, who regulates, and, ultimately, who decides how the data revolution reshapes our society.

Given the urgent challenges our world faces and the complicated questions on how to use data for good, we must work to ensure that this time of accelerated discovery and use serves more people, not fewer, and serves them with a commitment to equity and justice. It is this demand that drove us to study the progress, trends, risks, and opportunities in the nascent field of data for social impact (DSI).

We found encouraging signs: heightened interest in the value and opportunity that data can deliver as a tool to achieve social impact and widespread impatience to achieve that change. Yet aspiration has not yet become systemic, sector-wide action. We must do the hard work to change how we as individuals, organizations, funders, and ecosystems operate. This requires not just transformation in how we work, but how the sector functions overall.

We only have to look to other sectors to see how the use of data has revolutionized their practices and output. In the finance industry, for example, data capabilities have transformed both the customer experience and the way companies detect and prevent fraud. Across many industries, small and medium-sized businesses are able to collect data from a range of sources (e.g., loyalty programs, social media engagement, purchase history) to gain a 360-degree view of customers that enables significantly faster and more accurate decision making. And in healthcare, machine learning is driving rapid, low-cost, accurate diagnostics of diseases, and the use of data has been instrumental in fighting COVID-19.

We have the power to build—and fund—a data-driven social impact sector that drives affordable and innovative ways of addressing the multitude of challenges we face. But to do so, we must be thoughtful, open, and bold. As this emerging sector develops we must ensure that it is more coordinated, and that it builds on principles of diversity, inclusion, equity, and accessibility (IDEA). A more diverse, global workforce with interdisciplinary perspectives can provide a foundation for data for social impact work that is both effective and just.

We pose this question to all reading this report—from data practitioners to social impact leaders to funders to the frontline workers: As we move from aspiration to action to achieve the vision of a powerful DSI sector equipped to change entrenched systems, are we willing to take risks, align incentives, and partner differently? Are we willing to think and act boldly?

At data.org we are indebted to those pioneers—past, present, and future—who have started this work and together, we are willing to support new risk taking, new opportunities, and new ways of thinking. We hope you will join us.

January 2023

Danil Mikhailov, Ph.D.
Executive Director
data.org

Ginger Zielinskie
CEO
Evergreen Strategic Advisors

Perry Hewitt
Chief Marketing and Product Officer
data.org

Uyi Stewart, Ph.D.
Chief Data and Technology Officer
data.org
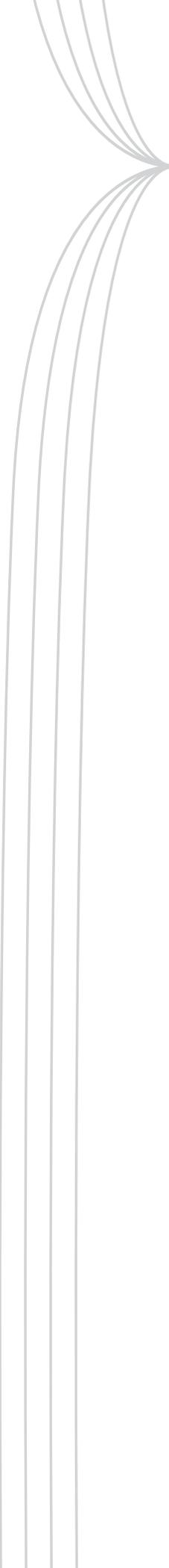
EXECUTIVE SUMMARY

To solve some of our greatest global challenges we need to accelerate how we use data for good. But to truly make data-driven tools that serve society, we must re-imagine data for social impact more broadly, more inclusively, and in a more interdisciplinary way.

This is easy to talk about, but harder to act on as we work to build a new field of data for social impact.

The field of "data for good" is not only overshadowed by the public conversations about the risks rampant data misuse can pose to civil society, it is also a fractured and disconnected space. There are multiple interpretations of what it means to "use data for good" or "use AI for good,"—creating duplicate efforts, nonstrategic initiatives, and confusion about what a successfully data-driven social sector could look like.

Add to that funding.

Resources are scarce for a field that requires expensive tools and skills to thrive. These enduring challenges result in work being done at an activity and project level, but do not create a coherent set of building blocks to constitute a strong and healthy field that is capable of solving a new class of systems-level problems.

The 2023 Data for Social Impact report strives to bring clarity to the conversation by synthesizing our collective progress to date, outlining key trends and tensions affecting social impact efforts, and exploring what comes next as our global community continues to unlock data to achieve positive social impact.

We believe in data for social impact and this report has only affirmed our optimism in the promise of a field in transition.[1]

We sought to evaluate the field across three categories of opportunity, challenge, and risk as: PURPOSE, PRACTICE, PEOPLE. This study aims to do three things:

1   Bring visibility to the nascent field of data for social impact and the ways in which it can transform global interventions and services and drive resilience.

2   Explore the potential to accelerate the strategic growth of this sector, particularly when it comes to increasing, sustaining, and nurturing the talent pool of interdisciplinary data professionals.

3   Offer recommendations for how to dramatically apply, govern, share, fund, and expand access to purpose-driven data around the world.

---

1   Porway, J. (2021). "Charting the 'Data for Good' Landscape." data.org.

To achieve these goals, this report focuses on a few key questions:

PURPOSE:

- How do organizations from small, local community-based organizations (CBOs) to large, international non-governmental organizations (NGOs) use data to drive their work?

- How does data inform theories of change and organizational and project goals?

PRACTICE:

- How do organizations broach questions regarding data access and infrastructure?

- What tools and governance structures are widely in use?

- What incentives do they use to partner across sectors?

PEOPLE:

- How can we support people across their careers to develop and strengthen the diverse skills necessary to further establish the field of data for social impact?

- How can we help leaders understand the potential they can unlock if their organizations use data better to achieve impact?

- How can we ensure that the communities we seek to empower are involved from the ground up?

What is Data for Social Impact (DSI)?

DSI is a nascent field that uses data, data science methods, and modern technologies to benefit people, communities, organizations, and the environment. DSI has already transformed and driven innovation across a wide range of industries, and delivered new ways to analyze giant datasets, advance predictive models, and harness machine learning for societal and environmental benefit.

Key findings

PURPOSE examines how the DSI sector uses data to achieve social and environmental change.

- **Opportunity**: We found that despite a significant gap in global connectivity—a third of people do not have access to the internet—digital connectivity in low- and middle-income countries (LMICs) is on the rise due to web-enabled mobile phones and smartphones. This increase is driving the use of more and more data and AI-based interventions and solutions. And social impact organizations (SIO) are finding innovative ways to drive transformation around the globe.

- **Challenges + Risks**: Our research also found that SIOs are focusing on project-based interventions instead of advanced data strategies that focus on systemic solutions, risking the long-term impact and sustainability of their efforts. In addition, nonprofits struggle with data management across their internal operations. According to Salesforce, only 22 percent of SIOs have achieved high data maturity today and fewer than half make decisions "based on data and evidence as well as designing programs and services using information about and engagement with recipients."

- **Path Forward:** To improve and enhance how the DSI sector uses data to achieve social and environmental change, we recommend greater access to digital public goods, such as open source tools and software for data analysis and forecasting; an increased focus on capacity building, particularly increased talent, within organizations and collaboratively across the sector to help deploy advanced data strategies; and persuade funders of the need for long-term, sustained investment to scale systems-wide data projects along with foundational investments in infrastructure, capacity, and talent.

PRACTICE explores how the DSI sector collects, accesses, and uses data.

- **Opportunity:** We found that there are increasing opportunities to extract extraordinary insights through data sharing—insights that can drive action at an unprecedented and scalable level, particularly cross-sector data sharing. The integration of new and different types of information can be transformed into actionable intelligence that can build stronger communities and stronger solutions. But we must begin to think differently about how we can share data, share infrastructure, and potentially share talent at scale. If we do so, we can create small and large-scale networks to drive systems-change around the globe with illuminating outcomes.

- **Challenges + Risks:** But while the promise of innovation ecosystems, enhanced coordination, and shared services models are clear, they can be hard to achieve without aggressive investment and incentivization. And in today's resource-constrained environment that's challenging. Private sector organizations who want NGOs to use their solutions are often incentivized to deploy a specific product, not necessarily to build a suite of tools suitable for

multiple NGO partners. NGOs who may want to share work and findings more frequently are constrained by the constant need to seek more resources. And research and academic institutions guard their data's intellectual property and publish papers off its fruit. Add to this the lack of coordinated global governance and regulation around data and AI and you face both ethical risks and very real risks for human rights, including the rights to privacy, personal data protection, equality, and more.

- **Path Forward**: To improve and enhance how the DSI sector collects, accesses, and uses data, we recommend increased investment in innovation networks and the co-creation of shared data ecosystems with community partners; a focus on innovative, diverse, collaborative partnership models to meet the demand for data for social impact vs. traditional funding methods that are project based; tailored incentives depending on the type of organization, the project, and even the individuals involved to increase data partnerships and sharing; and establishing a global framework to implement ethical AI technologies, starting with collaboration around best practices and common language.

PEOPLE examines how to grow and support the talent, leaders, and training propelling the DSI sector.

- **Opportunity:** Data capabilities are the most in-demand skills globally and there exists a significant opportunity to create and support a talent pool of over 3.5 million data for social impact jobs in developing countries over the next decade. Creating a workforce of highly-skilled, diverse, interdisciplinary data practitioners could advance our sector toward more effective and scalable solutions. Our research identified numerous small- and large-scale training hubs, sharing knowledge and resources with other hubs around the globe. However, nurturing the entire pipeline—beginning with education to new, existing and emerging talent to leadership—is critical to filling the growing demand.

- **Challenges + Risks:** It can be difficult for SIO leaders to understand the important role data can play in advancing their work. Even those that do understand the value of data struggle to justify allocating existing limited resources toward data professionals instead of those on the front lines. We also risk doing more harm than good if we don't involve the communities we seek to empower. When a solution is dropped onto a community, rather than built with them, data, AI, and other technologies we deploy may be creating new or exacerbating existing problems. Part of the solution lies in greater localization, particularly in low- and middle-income countries where solutions have been historically uninformed by local understanding and context.

- **Path Forward:** To improve and enhance how the DSI sector grows and supports talent, leaders, and training, we recommend sustained and intentional investment in the entire workforce pipeline, starting with education of new talent to existing and transitional talent to leadership; investment in constant training to get the best talent to do the best job; eincluding data

scientists in decision-making within social impact organizations to build knowledge; and data infrastructure and solutions informed by and supported by the local community. We know that we must prioritize IDEA, investing in the growth of diverse data practitioners with interdisciplinary skills and lived experience. This investment will enhance their creativity, capabilities, trust, and sustainability and strengthen the sector as a whole.

Recommendations for advancing the field

1. Improved data strategies through common governance and tools, data sharing, aligned incentives, and cross-sector coordination.

Our research found that so long as shared services and shared data collaboratives remain the exception, sector inefficiencies will remain. However, the very prevalence of this issue indicates an opportunity to evolve from project-based data projects to systems-level strategies. Improvement lies in better coordination between diverse partners and actors, and long-term, well-resourced multi-sector (public, private, academic, SIO, philanthropy) partnerships based on honest and clear conversations about our diverse incentives.

2. More diverse and interdisciplinary purpose-driven data practitioners who can drive change locally.

The Workforce Wanted report identified an opportunity for 3.5 million jobs in the data for social impact space in developing countries over the next 10 years.[2] To meet this opportunity head on, we need to train diverse, interdisciplinary data practitioners from the communities we seek to support. We must also inspire leaders, organizations, funders, and intermediaries to proactively invest in the time, tools, and efforts required to build a workforce that can learn from the local context and embed subject matter expertise into the data lifecycle.

3. Stronger funding models with longer time horizons, more flexible funding, and better coordination.

The DSI sector is largely supported by grant-based funding from philanthropic, government, and NGO organizations. As our research has shown, many grants are narrow in scope, focused on short term objectives, and constrained by only supporting what can be measured by sometimes outdated frameworks. We must move to a more collaborative funding model, increasing coordination both across verticals within individual funding organizations and across funding partners, more generally. To achieve this vision, we also must persuade funders of the need for long-term, sustained investment in talent and technology ecosystems.

---

2 2022. "Workforce Wanted: Data Talent for Social Impact." data.org

At data.org, we are committed to driving progress in the data for social impact sector across all of these areas. We are investing in coordination and supporting activity and energy in the sector to accelerate impact with alignment. We are partnering with innovation, education, and social impact organizations to increase talent across all four powerful pathways. And we are focused on funding, working with philanthropy, private sector, and other funding partners on practical approaches for longer horizon investment.

Join us.

Women entrepreneurs in Iringa, Tanzania. Photo by Solar Sister.

Purpose

Operator of a cold room at one of the pilot sites set up by Oorja in Muzzarfapur, Bihar, India. Photo by BASE/Empa.

PURPOSE

PURPOSE examines how the DSI sector uses data to achieve social and environmental change. How do organizations from small, local community-based organizations (CBOs) to large, international non-governmental organizations (NGOs) use data to drive their work? How does data inform goals and theories of change?

Opportunity:
Data can drive high-impact insights and interventions like never before.

Only bold, collective action can address the sheer number and the severity of crises and challenges that we face in the world today. Fortunately data-driven technologies have the potential to create high-impact interventions like never before—interventions that are affordable and most importantly, that drive resilience.

"We are observing a huge rise in digital connectivity in low- and middle-income countries (LMICs) over the last two decades, due to web-enabled mobile phones and smartphones," explains George Kibala Bauer, Director of Digital Utilities at GSMA, an industry organization that represents the interests of mobile network operators worldwide. "Usage keeps going up and up, which enables a whole new array of data for social impact projects and use cases."

Take agriculture in Sub-Saharan Africa. Many farmers around the world today have access to a variety of mobile applications that can identify crop yield, pests, disease, and more—dramatically stabilizing and increasing production in an area where food insecurity is at an all-time high. To help make these tools more relevant and useful for African farmers, the Lacuna Fund is supporting a dozen organizations to create or complete localized agricultural data sets.

And data itself is also driving action at increasing rates. Citizens Advice (CA) in the UK is an 80-year-old organization that provides free, confidential advice to assist people with legal, economic, housing, and more. At the beginning of the 2022 UK cost-of-living crisis, CA quickly realized they were sitting on a gold mine of data. What did they do with that data? They created a public dashboard to inform not only the communities they served, but to inform and influence local and central policy makers.

These are just two examples out of the hundreds we highlight as use cases at data.org to show the innovative ways that social impact organizations (SIOs) are using data to drive transformation around the globe. And it's our work as social impact leaders, data practitioners, funders, and governments to make sure that transformation is for the better.

As we illustrated in our Workforce Wanted report, there are three main ways in which the data field is helping SIOs to amplify their efforts:

1 gaining a better understanding of the broader context in which the organization works,

2 strengthening operational efficiency and effectiveness, and

3 designing better solutions to serve beneficiaries and address social issues.

Our research found that it was in the third category—designing solutions—where SIOs are gaining the most funding and taking action by utilizing data and data-based technology, while the first two categories are less understood and resourced, posing a risk to the sustainability of both project-based and systemic interventions.

Data Maturity Assessment

data.org is committed to capacity building of mission-driven organizations that seek to integrate data and data science into their work. As such, we recently launched the Data Maturity Assessment, a tool to help SIOs measure and understand their capabilities and connect them with the resources they need to move forward.

Take the Data Maturity Assessment at datajourney.data.org »

In late 2022, Salesforce published the fifth edition of their Nonprofit Trends report. They found that nonprofits are seeking to automate major parts of their business operations, including "donor management, gift processing, payroll, financial reporting, and more."[3] Furthermore, 79 percent of the nonprofits studied in the report indicated they "expect to implement widespread usage of artificial intelligence (AI) technology in the next three years."

But in concert with our own research, the report illustrated nonprofits' struggles with data management across their operations. According to Salesforce only 22 percent of social impact organizations have achieved high data maturity

---

3 Jensen, J. Rizek, J. 2022. "Nonprofit Trends Report." Salesforce.

today and fewer than half told them they make decisions "based on data and evidence as well as designing programs and services using information about and engagement with recipients." In conversations we held with sector leaders, they described silos and lack of knowledge. "We get a lot of requests from nonprofits to use their data for AI," said a senior expert, leading the social impact work at a global consulting firm. "But they're not understanding the difference between basic data analytics versus when you need data science versus when you can really look at predictive analysis. That distinction is missing completely, in the social sector."

Which leads to (drumroll please) SIOs' greatest challenge when it comes to utilizing data…

Challenge:
While data is abundant, advanced data strategies for sustained solutions are largely absent.

It's clear that data is everywhere. While 86 percent of respondents to data.org's Data Maturity Assessment (DMA) told us they collect data from clients or end users, just over half either sometimes or never/rarely use data to implement their programs or improve program efficiency. Add to that the nearly 40 percent who never, rarely, or only sometimes use data to understand the clients and the environments they serve.

In other words—SIOs are data rich but insight poor.

But when diving further into data.org's DMA results, it became clear just why this might be: quality, consistency, and accuracy. The DMA showed that fewer than a quarter of respondents believe that their data is mostly complete; fewer than half of respondents believe their data is mostly accurate; and fewer than a third of respondents believe that their data is mostly consistent.

That's a lot of incomplete, inaccurate, and inconsistent data.

But as Fraym Energy Sector Lead Jackie Mwaniki notes, "Even those that are not necessarily collecting and storing their data according to best practices, still have access to valuable information." That information is not always easily accessible, however, because the systems that transform data into insight and action are unknown to organizations, or are new and not widely available. Hence, the fact that only 65 percent of DMA respondents felt they even had the tools to conduct data analysis.

Mwaniki illustrates this further when explaining how Fraym, a software company that provides location–based data to governments, companies, and organizations works to help SIOs. "The problem isn't that they don't have data. It's that they don't have the tools to contextualize and understand the value of their data. The first step is helping them ask—what data do we have? What does that data mean?

What tools can we use to run analytics to see how to use the data for decision making?"

We have heard clearly that SIOs need better access to additional resources to meaningfully analyze their data. In fact, a survey by TechSoup has found that 82 percent of NGOs in low- and middle-income countries cite a lack of funding as their biggest barrier to adopting digital tools for data efficacy. "There needs to be more freely available, openly licensed material to support leaders in social impact organizations," noted Stuart Coleman, Director of the Open Data Institute (ODI).

In addition to the lack of access to analytic tools, many leaders and employees within SIOs lack the appropriate skills to steward data through the pipeline from collection to action. As a senior expert leading the social impact work at a global consulting firm explains "Workers grasp the importance of collecting data, but are at sea when called upon to inform their ongoing practices. They still struggle to understand what data is important and how to use it effectively and efficiently."

Deploying advanced data strategies requires a commitment to capacity building—and it must happen across the ecosystem, from the organizations to the funders. Investing in the capacity of humans is the most powerful thing that we can do to move along the transformation curve at this early stage of the DSI sector's growth.

Katherine Lucey, who leads Solar Sister, offers a great example of what you can unlock when you invest in the humans and internal capacity behind a solution. With data.org's support through the Inclusive Growth and Recovery Challenge, she invested in making sure she had data experts on her team and the budget to support them in the long term. As a result, her work in supporting local women entrepreneurs in Africa who work with clean energy has become a model for how data science can help steer social impact.

But to induce this level of capacity building within organizations and across the sector, we must work better together. "There is this total lack of coordination, resulting in broad fragmentation," says Stefaan Verhulst, Co-Founder of GovLab, which works to develop and test prototypes for new data collaboration platforms to enable open, collaborative problem solving inside and across organizations. "If we're going to solve problems in a way that extends beyond one community or geography or sector, entities and organizations must work together to establish different models for data collaboratives as opposed to everyone having to set up their own thing."

Capacity Assessment

As part of the Inclusive Growth and Recovery Challenge report, data.org and DataKind reviewed over 100 applicants that scored highest in their data science assessment to determine their capacity for data, talent, technology, and partnerships:

Research institutions and universities had the highest potential for data capacity; however, on partnerships they scored quite low. Furthermore, their interventions were focused primarily on insight that could potentially be applied, rather than a full plan to reach impact.

Think tanks showed better performance on partnership commitment. However, they tended to lack experienced data science or relevant technical talent to lead the technical portion of the research.

Multilateral organizations were most equipped for diverse, multi-sector partners, especially including the public sector, though formal confirmation of partner commitments varied greatly depending on the stage of the application: idea, pilot, or scaling. The majority of the applications relied on external partnership or contractors to lead and expand the technical portion of the project.

Service providers primarily used their own internal transaction data from technology systems. Internal tech teams may lack experienced data science talent to lead, develop, and implement data science and machine learning projects. Most project proposals were self contained solutions, so service providers also tended to have few or no partners.[4]

Risk:
Most organizations are funded on a project basis, risking the large-scale impact of system-wide coordination.

In our Rising Equitable Community Data Ecosystems (RECoDE) report from 2022, we found that when it comes to financial support, data systems are too often developed for what funders want, not what communities or organizations

---

4 2022. "Inclusive Growth and Recovery Challenge Report." data.org

necessarily need. And from India to Colombia, our research confirmed that funding often focuses on a specific project and not the broader, systemic needs.

In fact, 86 percent of survey respondents to the 2021 All In National Inventory—an overwhelming majority—agreed that their organizational leaders have a clear idea of how data can be used to drive decisions, but many—54 percent—indicated that funding requirements still define what data they choose to collect.[5] And according to Connect Humanity, recent digital equity funding has largely come in the form of relatively small grants. For example, in the US, 70 percent of grants for digital equity between 2018–2020 have been for amounts between $1,000–100,000. Only two percent were for over one million dollars.[6]

This is problematic, to say the least.

Durable, data-based social impact solutions must be system-wide interventions that invest up and down the data stack and within and across organizations. "A lot of the problems can be solved by educating the funders," says Donald Lobo, Executive Director of Chintu Gudiya Foundation, a private family foundation based in San Francisco that funds mostly India-based NGOs. "We need to ask them to not only fund the project, but the larger technology ecosystem."

Kriss Deiglmeier, Chief Global Impact Officer at Splunk, a leading security and observability platform that helps organizations use data at any scale, agrees. "For all those nonprofits that want to be data forward, they are stuck in a funding system that does not fund organizational effectiveness. Step one is that funders need to start investing in data and technology to drive mission results. Thinking in terms of overhead is for the past."

> *"If you're really trying to change people's way of thinking, if you're really trying to change the social context, if you're really trying to change entire paradigms, a two-year project is not going to do it."*
>
> — *RECoDE Interviewee*

Director at NESTA, a UK-based social impact innovation agency, Juan Mateos-Garcia agrees, "If you're going to be addressing systemic challenges," he says, "create an ecosystem of interventions that are well orchestrated and coordinated, not one-offs."

---

5 2022. "Rising Equitable Community Data Ecosystems (RECoDE)." data.org

6 Worsdman, C. 2022. "Funding to Bridge the Digital Divide." Connect Humanity.

Scaling systems–wide data projects requires both longer time horizons along with foundational investments in infrastructure, capacity, and talent. In other words— organizational and sectoral transformation requires unrestricted capital and longer time frames.

Take the example of early warning systems in South America. When trying to create early warning alerts for epidemics of infectious diseases that are exacerbated by climate change, the team at Innovolab at Universidad Peruana Cayetano Heredia knew what was missing—a digital infrastructure of harmonized databases across nations and sectors. "If we want to couple climate data with health data with economic data with demographic data, it becomes very complicated," says Director Gabriel Carrasco. "These data sets don't speak the same language, but we need greater insights for action, and that means there has to be a coordinated ecosystem of translation and connection."

The Harmonize project will bring together ministries, universities, private companies, SIOs, and more to create a complex data infrastructure. It will be hard. It will take years, it will take a lot of back and forth, and it will take a lot of investment. The end result will not be a shiny new AI application, but the harmony of various databases that, if put toward action, could save many lives.

Systems level data strategies like this, by nature, will take a longer to bear fruit. In crafting these strategies it is incumbent on us to understand the structural investments necessary to collect, clean, store, analyze and govern data across partners, and that's not always sexy to funders.

> *"We need to get serious about building the infrastructure to work on these issues over the long haul. Not sort of bat at them and then be like, 'Oh, that's hard.' Yes. It is hard— so is poverty. So is hunger. And we don't stop working on them."*
>
> *— Chris Worman, Co-Founder, Chief Partnership and Strategy Officer of Connect Humanity, Founder and Board President of TechSoup*

"A coherent, strategic approach to actively shaping technological development in ways that benefit society should pursue different streams of work in parallel: reorienting the direction of research and development, while also better-orchestrating data, evidence, AI, and experiments," argued four Carnegie Mellon researchers in a 2020 paper.[7]

But the report concludes that if civil society lacks the capacity to realize solutions requiring that scale of strategy, foresight and imagination, the gulf will only grow—military, surveillance, and corporate actors will define the data systems and priorities. Social impact organizations will be left behind in their ability to leverage data and technology, at best applying inappropriate solutions designed for different sectors. They add, "All of this will make it harder to anticipate, prepare, and respond to the big changes that will likely take place over the next decade."

7   Mulgan, G. 2021. "The Social Economy and the Fourth Industrial Revolution." Stanford Social Innovation Review.

Recommendations:

To improve and enhance how the DSI sector uses data to achieve social and environmental change, we recommend:

1 Greater access to digital public goods, such as open source tools and software for data analysis and forecasting, to help SIOs gain insight and take action on their data sets.

2 An increased focus on capacity building within organizations and collaboratively across the sector to help deploy advanced data strategies.

3 Persuasion of funders for long-term, sustained investment to scale systems-wide data projects along with foundational investments in infrastructure, capacity, and talent.

data.org resources and guides:

- Crafting Your Data Strategy »
- How to Make Your Data Actionable »
- Data Maturity Assessment »

Practice

PRACTICE

PRACTICE explores how the DSI sector collects, accesses, and uses data. How do organizations broach questions regarding data access and infrastructure? What incentives do they use to partner across sectors? What tools and governance structures are widely in use?

Opportunity:
Data sharing can unlock extraordinary insights and actions from the local to the global.

Sharing is hard, even for adults. It's messy, it's uncomfortable, it's vulnerable. But in the world of data—it's also essential. Without relating data sets to one another across nations, sectors, environments, and more we risk missing extraordinary connections and insights.

As the World Bank's World Development Report 2021: Data for Better Lives explains, "Data collected for one purpose has the potential to generate economic and social value in applications far beyond those originally anticipated."[8]

The bad news? Quality data is often siloed across the public, private, and social sectors, and not easily shared across sectoral boundaries. There are logistical and legal barriers, such as data governance and data sharing agreements, and important questions around equity and how policies impact marginalized communities.

These barriers are preventing SIOs from adequately accessing important data from governments and the private sector—data that can meaningfully improve their ability to succeed. As Stefaan Verhulst, Co-Founder of GovLab, explains, "The amount of data is piling up. But we haven't made progress in actually nudging the data holders to provide access to their data, so that it can be reused responsibly to solve societal problems."

The good news? There's a lot of room for improvement. As The Rockefeller Foundation's Vice President Asia Region Deepali Khanna said, "The opportunities to share are boundless. Given how new the DSI field is, now is the time to really push the importance of inclusion, transparency, and innovation when it comes to creating shared data, infrastructure, governance, and more."

And we must seize these opportunities. We must think differently about how we can share data, share infrastructure, and potentially share human capital. Shared services, infrastructures, and learning models offer the clearest alternatives. But this is a daunting undertaking that will require a lot of work and a host of actors to work in concert—funders, innovators, practitioners, and regulators (think back to

---

8 2021. "Data for Better Lives." World Bank Group Flagship Report.

the unsexy, laborious work of Innovolab's Harmonize project in PURPOSE).

Yet, not every organization has to build a complex data system and team, but all organizations should be able to derive value from data to meet their missions and achieve impact. Investing in innovation networks connected through a shared problem statement or geography holds promise. Around the world there are many individuals and organizations creating small and large-scale networks with illuminating outcomes.

Take the Development Data Partnership, a collaboration between international organizations and companies supporting the responsible use of alternative data in international development started by the World Bank's Holly Krambeck. "What was missing," Krambeck noted, "was an efficient market that could match the supply of private sector data and public-sector demand, grounded in a shared legal and IT framework that would protect rights and benefit both parties." In the past, she notes, they were approaching companies to participate, but now that model is changing. Companies are approaching the partnership because they see value in working with international organizations who share data governance principles and infrastructure.

While this and many other case studies support the positive outcomes that can come from sharing data and infrastructure, it's important to also keep in mind that sharing can have risks associated as well. In its 2020 Toolkit for Centering Racial Equity Throughout Data Integration, the Actionable Intelligence for Social Policy (AISP) reports, "cross-sector data sharing and integration enable the transformation of individual-level information into actionable intelligence that can… ultimately build stronger communities. Yet, [it] can also reinforce legacies of racist policies and produce inequitable resource allocation, access, and outcomes."[9]

As we highlighted in our RECoDE report, if marginalized communities and stakeholders are not a meaningful part of the initial data ecosystem design, their perspectives are too often overlooked. Data ecosystem leaders need to make a deliberate effort to ensure that these models are co-created with community partners and evolve in ways that are responsive to and aligned with the needs of the people they serve.

---

9 2022. "A Toolkit for Centering Racial Equity Throughout Data Integration." AISP

What are community data ecosystems?

Community data ecosystems are made up of the what, the who, and the how that enables data sharing and collaboration within a community. They include data infrastructure, tools, user capabilities, standards, and policies

For instance, in St. Louis, community leaders weren't just looking for an online portal or a signed agreement that would allow them to look under the hood at regional health data or other key social drivers. They wanted to go deeper—they wanted to understand the priorities, challenges, and opportunities faced by their SIO colleagues across the city and create a more holistic regional data sharing system that could improve outcomes for all. And so the St. Louis Regional Data Alliance (RDA) was born. It created a shared infrastructure to make data sharing easier and data more accessible and actionable in the process by engaging the communities it was serving.

Organizations like RDA exist around the United States and are models of community-based data ecosystems. The National Neighborhood Indicators Partnership connects residents, nonprofits, and the government to share and understand neighborhood level data; the Alamo Regional Data Alliance emphasizes the collaborative work of a community of practice; and the Western Pennsylvania Regional Data Center created an open data portal that sourced information from the city and county.

Of course, developing more mature infrastructure and ecosystems requires caution to ensure governance structures distribute power equitably and adhere to governance best practices,[10] which leads to the greatest PRACTICE challenge facing SIOs…

---

10 Novitske, L. (2018.) "The AI Invasion is Coming to Africa (and It's a Good Thing)." Stanford Social Innovation Review.

Challenge:
Sharing means aligning incentives amid competing priorities and resource-constrained environments.

The promise of innovation ecosystems, enhanced coordination, and shared services models are clear—but they can be hard to achieve without aggressive investment and incentivization. And in today's resource-constrained environment? That's tough.

Private sector organizations who want NGOs to use their solutions are often incentivized to deploy a specific product, not necessarily building a suite of tools suitable for multiple NGO partners. NGOs who may want to share work and findings more frequently are constrained by the constant need to seek more resources. And research and academic institutions guard their data's IP and publish papers off its fruit.

So how do we make progress with so many competing interests? As Frederic Pivetta, Co-Founder and CEO of Dalberg Data Insights, which builds inclusive data ecosystems for international development challenges, explains, "We need to have a kind of market design view of how to make this happen. It is not going to happen by magic. It is not going to happen only by philanthropy."

So we need to think more expansively than traditional funding and partnership structures and instead look toward innovative, diverse, collaborative partnerships to meet the demand for data for social impact.

Case studies suggest exciting examples of sharing data, infrastructure, and talent from the private sector, and then partnering with research organizations and leading NGOs to address a specific problem. Take the IDEA initiative from J-PAL Africa for example. After piloting innovative partnerships and data sharing agreements between researchers, data providers, and governments, they published the IDEA Handbook, which provides case studies on how data can be safely and ethically accessed, and how partnerships between researchers and data providers can be navigated.

But the challenge remains—how do we drive the fundamental shifts required to ensure this level of partnership at scale and over time? In other words, how do we incentivize this shift for all parties in a way that builds mutual trust and positive outcomes for all?

It's hard, to say the least. As Cornell University Assistant Professor Nate Matias explains, "Trailblazers working on technology and social good have faced the mismatch between the goals of tech company leaders and the public purpose that social good innovators have. Often, there's a lot of overlap and opportunity for collaboration. But sometimes there's real conflict…There are just certain kinds of public service that are easier to do well in a trustworthy way if there's no

question about the influence of industry on your work, especially where your work holds companies and industries accountable."

Working across the private, public and NGO sector, Bayo Adekanmbi, Founder of Data Science Nigeria, shared with us how he balanced distinct value propositions across stakeholders as he worked to unlock telecom data from the private sector, "When you engage the private sector, you always have to answer the question—what's in it for them? If companies make their data available to governments, then in turn they will want to generate some value from that exchange, a tax hold for example. Our experience has shown that there must be some incentive from governments who are often the biggest beneficiary, especially in low data resourced economies."

Despite the challenges, our research and the many case studies out there still hold promise that aligned, cross-sectoral incentives can drive greater collaboration, coordination, and data sharing.

But these incentives must include:

- revenue generation
- deeper insights
- better service
- organizational efficiencies
- cost savings
- credit sharing

These incentives must be tailored and distinct depending on the type of organization, the project, and even the individuals involved. It will never be a one-size fits all model. But if the sector as a whole can recognize the value of strong incentivization from the outset of a partnership, their approach and outcomes will dramatically improve, and the number of partnerships will scale exponentially.

It sounds challenging, but it is possible.

"I've seen academics and funders and practitioners walk into a room a bit wary," says Arturo Franco, Senior Vice President at Mastercard Center for Inclusive Growth. "But inevitably you come to a moment when you all realize that you can speak the same language and understand the different roles that each sector needs to play. From there, you can get to a very concrete and exciting idea that everyone can support. It's magic."

Risk:
Without strong data governance and regulation, we risk doing more harm than good.

Remember how we said sharing was messy? It can also be quite risky. Especially when there are not clear and present rules around the sharing. Think of one of the most infamous examples of misuse between two parties sharing data. In 2018, a UK political consulting firm used personal data from Facebook users that was originally collected from a third party for academic research. Cambridge Analytica misused the data of nearly 87 million Facebook users. Within two months of the breaking news, Cambridge Analytica was bankrupt and Facebook was fined $5 billion by the Federal Trade Commission.

That's why data governance is so important.

It encompasses the infrastructure policies, technical mechanisms, data laws and regulations, and institutions entrusted to effectively enable the safe, trustworthy use of public intent and private intent data.

By providing confidence that these rights are protected, an effective data governance framework can strengthen trust in the data system, and as a result, incentivize data sharing and the use of data-driven products and services.

In the words of the World Bank's Data for Better Lives Report, "Data governance enforces the social contract around data by applying the principles of trust, value, and equity."[11]

Pivetta explains, "We need to have proper data governance because it builds trust. And with trust comes sharing. And so that governance layer, that regulatory layer is so very important."

But the absence of data governance across the public, private, and social sectors is impeding the production of usable, data-based insights. "Often it's too late when organizations realize they need a data governance strategy because they're already collecting data. They're collecting it as we speak with no architecture or governance," says Rodrigo de Reyes Lanfranco, Country Managing Director, Fundación Capital, an organization that works with governments and the private sector to create digital technology solutions to social challenges.

---

11 2021. "Data for Better Lives." World Bank Group Flagship Report.

"What do we talk about when we're talking about governance, because I think there's a little bit of an assumption in a lot of places that if you just design the tools and the interfaces, right, and you make sure that everything's legal, then that's governance."

— RECoDE Workgroup Member

Pursuing stronger governance, enhanced regulation, and wider accountability is particularly relevant in the context of AI—technology that is increasingly running up against ethical risks and considerations as it becomes more widespread.[12] As one observer wrote in Tech Policy Press, "AI is not neutral and its implementation may imply potential risks for human rights, such as the right to privacy and to personal data protection, the right to equality and the right not to be discriminated against.[13]"

If we want AI and related technologies to be responsibly and ethically adopted globally, we must take a proactive (rather than a reactive) approach to regulating them.[14] And that proactiveness must include near-continuous assessments of not only the ecosystem but also their internal structures and approach to governance to keep pace with the fast-changing technologies.

Sampriti Saxena and Stefaan G. Verhulst, who teach a free online course on AI ethics[15], urge the importance of establishing a global framework to implement ethical AI technologies, but acknowledge that the "the borderless nature of technology and the diversity of the field pose challenges to building effective governance frameworks." Khanna agrees, "Regulation and requirements around our digital infrastructure can't be thought of as something with clear geographic boundaries or constrained by financial penalty. Our digital infrastructure will only become more global; the systems for designing and enforcing policies for that infrastructure needs to be global in outlook while grounded in regional realities, too."

A good first step? Collaboration around best practices and common language will go a long way toward helping to level the playing field.

---

12 Saxena, S. And Verhulst, S. 2022. "10 Learnings from Considering AI Ethics Through Global Perspectives." Data Stewards Network.
13 Levy Daniel, M. 2022. "How Does the Public Sector Identify Problems It Tries to Solve with AI?." Tech Policy Press.
14 Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). "Algorithmic Accountability for the Public Sector.
15 Saxena, S. And Verhulst, S. 2022. "10 Learnings from Considering AI Ethics Through Global Perspectives." Data Stewards Network.

As Tim Davies, Research Director at Connected by Data, an initiative working to put community at the center of data narratives, practices and policies, explains, "we can not underestimate the importance of engaging with open standardization in order to avoid investing in new data silos. Without agreement across different datasets of how to refer to a farm, factory, or other site, integrating data and data exchange can be complex. But maintaining these reference lists is also a big job, a potential point of centralisation and power, and a gap in data infrastructures."

And that's just one gap among many. Luckily, "AI Localism"—efforts to use AI on a local level—has emerged to fill gaps left by incomplete state, national, or global governance frameworks. AI Localism creates the mechanisms to strike a balanced approach to both the opportunities and risks of advanced data capabilities.[16] Examples include local bans on AI-powered facial recognition technology, local procurement rules pertaining to AI technology, public registries of AI systems used in local government, and public education programs on AI that are offered by local governments or institutions.[17, 18]

Further, as the World Bank's Data for Better Lives Report concluded, "Non-governmental institutions and mechanisms such as data intermediaries can help governments and other actors safely share and use data to capture greater value, while promoting equitable access to data and the value they create."[19]

Indeed data intermediaries in particular are crucial to protect data, gain meaningful consent and build trust in data companies. A data intermediary serves as a mediator between those who wish to make their data available, and those who seek to leverage that data. In any data exchange there is a creator, there may be an intermediary, and there is a data user. The real value of standards comes when you have multiple creators, multiple users, and potentially, multiple intermediaries.

Nongovernmental institutions and mechanisms such as data intermediaries can help governments and other actors safely share and use data to capture greater value, while promoting equitable access to data and the value they create.

Our research made clear the urgency of building an intermediary layer to enable sector-wide transformation of data analysis practices. But what does that layer look like? What skills do we need to deliver and then operate it? And what market incentives exist to build then? Who has the greatest incentive to fund and build such an entity?

Answers to some of these questions can be provided by intermediary organizations that support the sector as a whole or key functions within it—such

16 2022. "AI Localism." GovLab.
17 Verhulst, S. 2022. "The AI Localism Canvas: A Framework to Assess The Emergence of Governance of AI within Cities." Medium.
18 Saxena, S. And Verhulst, S. 2022. "10 Learnings from Considering AI Ethics Through Global Perspectives." Data Stewards Network.
19 2021. "Data for Better Lives." World Bank Group Flagship Report.

as data.org and others, like Global Partnership for Sustainable Development Data (GPSDD) and Academic Data Science Alliance (ADSA).

Well-positioned intermediaries have the potential to advance shared services for a whole range of actors in the DSI field and can broker more sustained and coordinated investment in underpinning infrastructure and capacity from funders.

Further, the Open Data Institute recently called for wider participation in data policymaking.[20] They argued that we are witnessing a Fourth Industrial Revolution—"an era of widespread automation of industry, made possible by the greater availability of data and innovations in the use of that data (such as new digital technologies such as AI)."

And as we've seen time and again in earlier industrial revolutions, without governance and regulation we risk widespread degradation of the environment and human rights—the exact things we are working to protect and enhance.

"Governance, assessments and classifications of data have to be implemented on the operations layer," says Abhijit Balakrishnan, ICT Business Analyst, Enterprise Architecture, SOS Children's Villages International. "Only then can you really say that yes, we are in a position to responsibly collect and use data points. That takes precedence for me, everything else comes second."

---

20   2022. "Experimentalism and the Fourth Industrial Revolution." Open Data Institute.

Recommendations

To improve and enhance how the DSI sector collects, accesses, and uses data, we recommend:

1. Increased investment in innovation networks connected through a shared problem statement and shared services, infrastructures, and learning models between organizations and sectors.

2. Co-creation of shared data ecosystems with community partners that ensure they are responsive to and aligned with the needs of the people they serve.

3. A focus on innovative, diverse, collaborative partnerships to meet the demand for data for social impact vs. traditional funding and partnership models.

4. Tailored incentives depending on the type of organization, the project, and even the individuals involved to increase data partnerships and sharing.

5. Long-term goal of establishing a global framework to implement ethical AI technologies, starting with collaboration around best practices and common language.

data.org resources and guides:

- 3 Key Steps to a Successful Data Commons **»**

- How to Build Your Data Stack **»**

- Data Security for Nonprofits **»**

- The Power of Partnerships **»**

People

PEOPLE

PEOPLE examines how to grow and support the talent, leaders, and training propelling the DSI sector. How can we support people across their careers to develop and strengthen the diverse skills necessary to further establish the field of data for social impact? How can we help leaders understand the potential they can unlock if their organizations use data better to achieve impact? How can we ensure that the communities we seek to empower are involved from the ground up?

Opportunity:
Investment in diverse, interdisciplinary talent can unlock the growth of the nascent DSI sector.

Global demand for data skills is fiercely competitive and changing rapidly. In fact, a World Economic Forum (WEF) Future of Jobs report[21] found that data skills were the most in-demand skills globally in 2020 and in our Workforce Wanted report, we found that there exists a significant opportunity to create and support a talent pool of over 3.5 million data for social impact jobs in developing countries over the next decade. Creating a workforce of highly-skilled, diverse, interdisciplinary data practitioners could advance our sector toward more effective and scalable solutions.

But growing this talent pool poses challenges. Obstacles include global scarcity of digital and data skills, wage disparity with the private sector, the variable data maturity of organizations, limited access to data and datasets, low levels of data usability, and lack of tech teams to build and grow the ecosystem.

Add to this a leaky pipeline.

"Further constricting the talent pool is a mismatch in expectation on the part of employees and employers. It means we're losing people in the leaky pipeline," notes Satyam Vyas, Founder and Chief Executive Officer of Arthan, which works to build capacity and strengthen systems for social impact organizations. "Even mature organizations don't have strategies or environments in place to not only hire the right form of talent, but to sustain them with the right set of work."

Despite those obstacles, there are a number of ways in which organizations can access and build data skills, teams, and organizational strategy to not only hire and grow data professionals and strengthen the entire DSI ecosystem. The key? Sustained and intentional investment in the entire pipeline, which has four pathways as defined in the Workforce Wanted report—from new talent to existing talent to transitional talent to leadership.

---

21 2020. "The Future of Jobs Report." World Economic Forum.

"You need to focus investment on all four of these pathways," says Mamadou Biteye, Executive Secretary, The African Capacity Building Foundation, which is working to address the shortage of critical skills, deficits in leadership, and inhibiting mindsets across the African continent. "Yes, you can put your funding toward hiring people away from the private sector and that will get you people ready to act now. But it's not sustainable. We have to build the talent over the medium and long-term or we will continue to face the same gaps we're facing now."

Four Talent Pathways

New talent: Potential DSI talent entering the workforce for the first time with no prior data work experience

Existing social sector talent: Current practitioners in the social sector who could be upskilled in data

Transitional talent: Existing data practitioners outside the social sector who could be upskilled or attracted to work in SIOs

Leadership: SIO leaders and senior executives who could acquire new capabilities and support to make their organizations more data-driven and to attract, retain, and utilize DSI professionals

But social impact organizations continue to wrestle with anemic funding to support the capacity and capabilities of all of their workforce, including data practitioners. "Most funding for nonprofits is focused on programmatic budgets (i.e., goes directly to the end beneficiary), which starves organizations of resources to build themselves, including investing in leadership capacity and data capabilities," says Isha Sharma, Associate Director of India Leaders for Social Sector, who works with organizations to build long-term professional capacity.

Funders, private sector companies, social impact leaders and practitioners must focus on coordinated, multi-year investments across the pipeline, starting with education. Reaching beyond the traditional sources of talent is also necessary— not only if we want to fill the forecasted 3.5 million data jobs available in the next 10 years, but if we want those data practitioners to be diverse across geography, race, ethnicity, gender, and more.

But academic and pipeline programs in data science are scarce in many of the communities that need the most support. In turn, we see the current gulf between talent supply and demand exacerbated by a lack of diversity across gender, racial lines, and geography.

The result? Well-meaning external organizations coming into communities to deliver solutions that have not been informed by data or local culture. Fewer than 40 percent of respondents in our data maturity assessment have dedicated data engineering skills on staff, and over a quarter don't employ or consult anyone with data engineering skills. Only 32 percent of respondents fully agree that their organization has the right skills and capabilities to manage, use, store, and interpret data.

Enter training. Private companies prioritize and resource continual training and professional development. It's one reason the private sector outpaces the public—innovation happens when you invest in knowledge.

"The nonprofit sector is stuck in the 20th century because leaders are unable to use their resources toward talent development. We need constant training and investment to get the best talent to do the best job. And we need funders to understand that," says Deiglmeier.

A good example is Data Science Nigeria, where training was made available to civil service employees. "We launched a training called NoCode.ai to make artificial intelligence more accessible to ordinary professionals without learning to code. We wanted to expose them to AI tools that allow anyone—whether you are a lawyer, an HR executive—to understand their value and application toward social impact," says Adekanmbi.

At data.org, we also understand how critical training is to empowering social impact organizations and local communities to be part of the changemaking process. We're working to train one million purpose-driven data practitioners in the next 10 years. Part of this effort is our Capacity Accelerator Network (CAN), which creates locally-led knowledge hubs that share a common vision with other global hubs and share resources and knowledge, yet set their own priorities, partners, and projects—ensuring community ownership and value creation.

But even if—collectively—we are able to build the army of data practitioners that we need, challenges remain. Will SIOs have the funding to recruit and retain them? Will leaders understand their value and embed them in positions of authority? Which leads to the greatest PEOPLE challenge facing SIOs...

Challenge:
Building the sector requires empowering SIO leaders and funders to understand and truly prioritize data.

In the World Bank's Data for Better Lives study, they assert, "The big push to build an army of data scientists for jobs in the public sector, private sector, and civil society must be complemented with efforts to create enabling institutional and leadership environments that place a high premium on the use of data and evidence."

But according to a 2021 study by Data Orchard[22], only 1 in 20 organizations have leadership that understands how to use data to improve organizational performance.

That's because data is complicated.

It's often difficult for those without formal or informal experience with data to understand the important role it can play beyond measurement, evaluation, and learning (MEL). Coleman shares that in many cases, "there's a fundamental lack of data skills and knowledge in leadership. It's critical that we help leaders in the social good space to understand the role data can play in advancing their work."

Leaders in the early stage of exploring and seeking to understand the opportunities and applications of data need to be supported through training and ultimately—mentorship and guidance. These leaders need help enhancing and shaping new models to support design, experimentation, and advancement of data-driven strategies, initiatives, and talent acquisition. As one expert stated, "The ability to be part of a network and share ideas and experiences is very powerful for nonprofit leaders."

But a review of more than 30 leadership programs in developing contexts revealed that only four could be considered to have a focus on DSI. Yet the overlap between programs focused on social entrepreneurship, technology innovation, and issue-based leadership in the social sector can provide us with some insight. The goals of most leadership programs offer common direction and thus potential for learning. Specifically, the majority of leadership programs we reviewed included a focus on one of three distinct outcomes: i) community building, ii) leadership training, or iii) initiative building. Noticeably missing? A data track.

Add a lack of few visible examples to follow and a lack of diverse, relevant role models, and a limited awareness of specific DSI use cases in relevant sectors, and you get leaders who operate on "gut instinct" instead of data-informed decision making.

---

22 2021. "State of the Sector Data Maturity in The Not-For-Profit Sector," Data Orchard.

"Data can feel for many leaders as an operational or a functional question," says Neera Nundy, Co–founder and Partner of Dasra, a venture philanthropy fund to invest in early stage nonprofit organizations in India. "They're not quite sure how to integrate a data strategy into their work and the whole gamut of complexity and priorities that they are facing."

One place to start is putting data scientists around the decision making table, working alongside subject matter experts within social impact organizations. This helps to build knowledge and capacity at the leadership level and ensures the quality of data and nature of analysis is robust and appropriate to the problem.

But even those that do understand the value of data struggle to justify allocating existing limited resources toward data professionals instead of those on the front lines of their programs. And often leaders have internal resistance from their workforce, who don't fully understand data's value. A study by Qlik[23] shows that 45 percent of employees surveyed heavily rely on "gut feelings" to make decisions.

Another challenge? Funding, of course! An IBM study[24] found that up to 74 percent of nonprofit leaders cited budget constraints, due to competing internal investments and the restrictive nature of grants, as one of their top barriers to advancing analytic capabilities.

"Most funding for nonprofits is focused on programmatic budgets (i.e., goes directly to the end beneficiary), which starves organizations of resources to build themselves, including investing in leadership capacity and data capabilities."

— Isha Sharma, Associate Director India Leaders for Social Sector (ILSS)

---

23 2018. "Lead with Data: How to Drive Data in the Enterprise." Qlik.

24 2017. "Leap Before You Lag: Nonprofits with deeper data capabilities see stronger impact." IBM Institute for Business Value.

Risk:
If we don't involve the communities we seek to empower, we risk doing more harm than good.

"I think the biggest risk that we face as a sector is solutions that we build that can be used for the wrong purposes," says Juan M. LaVista Ferres, VP, Chief Scientist, Microsoft AI for Good. "We definitely want to make sure that we are investing and we are solving the right problems."

Indeed, the 2022 World Economic Forum published A Blueprint for Equity and Inclusion on Artificial Intelligence, which called for the DSI sector to be more inclusive and equitable: "With growing concerns about bias, data privacy, and lacking representation, it is critical to re-evaluate the way in which AI is both designed and deployed to ensure that all affected stakeholders and communities reap the benefits of the technology, rather than any harm."[25]

In other words, it is time for us to put principles of inclusion, diversity, equity, and access (IDEA) into practice across all parts of the ecosystem.

Or, as Carnegie Mellon researchers concluded, "As AI technologies become increasingly powerful and democratized, we note several unfortunate cases where AI is developed for outright malicious use, e.g. DeepNude, or for highly dubious purposes. We believe such cases represent an urgent and complex problem that the AI community, and much more beyond the AI community, need to address together. Reducing such malicious use of AI is equally important as using AI to produce positive outcomes."

Adekanmbi further illustrates the risks associated with top-down deployment of solutions, "In our quest to solve a problem, we create more problems which might lead to exclusion, inequality or fairness, irresponsibility and all that. But I believe that the effort of getting it right may actually make us fast. And that's why I will emphasize the discipline of getting it right, which is, can we create that community, that middle way where everyone that is required comes to the table."

But how do we ensure that data, AI, and other technologies we deploy are not creating new or exacerbating existing problems? Our research found that as cross-sector work gains traction, marginalized communities and stakeholders must be a meaningful part of redesigning data ecosystems—their perspectives are too often overlooked.

Our RECoDE report found there are many points of failure within the data lifecycle, especially when it comes to who is making decisions. More than 400 survey participants were asked about their involvement of people with lived experience—those who carry the burden of inequitable outcomes—in various stages of the data lifecycle. The majority of these survey participants came

---

25 2022. "A Blueprint for Equity and Inclusion in Artificial Intelligence." World Economic Forum.

from either the nonprofit sector (55 percent or government agencies and public institutions (35 percent) with others coming from faith–based organizations, private companies, and beyond.

Involvement of Persons with Lived Experience in the Data Life Cycle

Involvement of Persons with Lived Experience in the Data Life Cycle

STAGE		
Plan	28% | 11% | 24% | 27% | 10%	n = 149
Collect	40% | 15% | 15% | 21% | 9%	n = 139
Access	45% | 20% | 18% | 9% | 8%	n = 143
Analyze	60% | 15% | 9% | 7% | 8%	n = 139
Interpret	40% | 20% | 16% | 15% | 9%	n = 138
Recommend	25% | 14% | 25% | 25% | 11%	n = 139
Report	42% | 27% | 13% | 10% | 7%	n = 135
Disseminate	39% | 22% | 11% | 16% | 11%	n = 134

PERCENT: 0.00 0.25 0.50 0.75 1.00

Involvement
- Not involved
- Informed
- Advises
- Actively engages
- Makes decisions

Overwhelmingly, people with lived experience are not actively engaged or making decisions within multisector data initiatives. Without relevant lived experience, the risk of misinterpretation of data or making false assumptions is high. Data infrastructure and solutions must be local, and must be informed by and supported by the local community in order to be meaningful and sustainable.

Building trust and sharing power with local practitioners and leaders and creating solutions informed by local understanding and context will ensure that the barriers that too often stand in the way of data access are dismantled. Part of the solution lies in greater localization, particularly in low- and middle-income countries where solutions have been historically uninformed by local understanding and context.

RECoDE principles for seeking to increase equity throughout their works

Marginalized communities must be a meaningful part of the initial data ecosystem design.

Systems should be designed by and for the people that they intend to serve.

There must be a balance to protect data security and maintain quality while also giving people control over their own data.

Both perspectives are valuable, and the best way to preserve and advance both is to increase collaboration to develop the skills and expertise to be the best data stewards possible.

Trust is the cornerstone to creating meaningful change. Building trust is time-consuming, but it's the only way to get accurate information about the challenges, risks, and opportunities in the communities we aim to serve.

The process of building trust is just as important as the final data-driven action. According to the Ada Lovelace Institute, data practitioners should prioritize engaging people in the governance of data through participatory data stewardship. Their 2021 report outlines choices about how and when to bring people into the process – from data collection, to linkage, to data analysis. The report says on participatory stewardship, "One of the most frustrating experiences for people is when they are told they will have the power to shape something, but find in fact that consultation is very limited."[26]

---

26 2021. "Participatory Data Stewardship." Ada Lovelace Institute.

"For me, it starts with cultural humility or humility, period," Dr. Wanda Boone, founder and CEO of Together for Resilient Youth. "What is the background of the work you're doing and if you're working with underserved populations, what is your capacity to listen to what the community is saying and respect what the community is saying?"

Equally critical to building trust is empowering people to build data systems with integrity by investing in diverse and interdisciplinary professional development within the tech and data science workforce. In a 2022 Social Science Research Network (SSRN) article,[27] researchers argued that interdisciplinary engagement is essential to the ethical deployment of AI because it enables team members with different expertise to navigate and reconcile their own understanding of the technology and its impact with one another. In other words—diversity and lived experience significantly enhance the creativity, capabilities, trust, and sustainability of data practitioners.

This concept is what drove University of Pretoria Associate Professor Dr. Vukosi Marivate to launch Deep Learning Indaba, an event and organization committed to empowering Africans to be owners and shapers of advancements in technology and artificial intelligence. "AI shouldn't be something that gets dumped on the continent," says Marivate. "But something that we can shape. We need AI on the continent on our own terms.

In 2021, the Urban Institute published a landscape survey on what investments could improve trust in the methods data practitioners use.[28] The report called for funders to "invest in the assessment of the effectiveness of different data-related training (personal privacy, algorithms, governance onboarding, and data use for advocacy, for example) to empower community members to package and scale training for tailored audiences."[29]

The seriousness of the risks mentioned above demand that we strike the balance between risks and opportunities. Repeatedly, the call to include communities in every step of the design process proved paramount. Our interviews with sector leaders substantiated those claims.

27 Mankoo, A. 2022. "Integrating Ethics into Data Science: Insights from a Product Team." SSRN
28 Hendey, L. 2021. "Envisioning a New Future: Building Trust for Data Use." Urban Institute.
29 Hendey, L. 2021. "Envisioning a New Future: Building Trust for Data Use." Urban Institute.

Recommendations:

To improve and enhance how the DSI sector grows and supports talent, leaders, and training, we recommend:

1 Sustained and intentional investment in the entire pipeline, starting with education of new talent through to existing talent to transitional talent to leadership.

2 Investment in constant training to get the best talent to do the best job.

3 Embedding data scientists around decision making tables within social impact organizations to build knowledge, cross–pollination, and capacity at the leadership level.

4 Data infrastructure and solutions informed by and supported by the local community in order to be meaningful and sustainable.

5 Diverse data practitioners with interdisciplinary skills and lived experience to significantly enhance their creativity, capabilities, trust, and sustainability.

data.org resources and guides:

- Building a Data Team **»**
- Capacity Accelerator Network (CAN) **»**
- Using your Data Responsibly and Ethically **»**
- How to Apply an Intersectional and IDEA Lens for Social Impact Organizations **»**

What's Next

What's Next

The field of data for social impact stands at a crossroads. Significant changes are required to address the most important challenges of our time. Now we have a choice. We can continue business as usual, funding and implementing under-resourced data projects that deliver incremental progress. Or, we can think and act boldly.

We have the power to deliver better coordination, build a diverse, inclusive and interdisciplinary workforce, provide intermediary support, and trust sector leaders with funding to build the data-driven social impact sector we all aspire to create. We can change how we work, partner, fund, learn, iterate, and adapt. However, to achieve this vision, people across the sector need to be open, take risks, align incentives, and partner differently.

People working across the globe are the foundation of this report—and their voices and expertise comprise its call for broader, more ambitious action. data.org stands on the shoulders of the many who have come before us, and works closely with aligned leaders and partner organizations driving data for social impact. Together, with our partners worldwide, we urge others to join us. We seek to build the infrastructure, talent, and systems to make the potential of data for social impact a reality.

Recommendations to advance the field

Our reporting revealed concrete and critical next steps: improved coordination, a more diverse talent pool with more interdisciplinary and translational skills, and cooperative and innovative funding practices to support the whole field. Each of these components are necessary to deploy system-level data solutions and, in concert, transform the sector.

1   **Improved data strategies through common governance and tools, data sharing, aligned incentives, and cross-sector coordination.**

Our research found that so long as shared services and shared data collaboratives remain the exception, sector inefficiencies will remain. However, the very prevalence of this issue indicates an equal opportunity to evolve from project-based data projects to systems-level strategies. Improvement lies in better coordination between diverse partners and actors, and long-term, well-resourced multi-sector (public, private, academic, SIO, philanthropy) partnerships based on honest and clear conversations about our diverse incentives. To enable effective coordination top to bottom in the DSI field we need the following:

- Shared language and definitions—basic terms around DSI that the field can unite around to decrease the chance of miscommunication

- A culture of transparency, common governance, and knowledge sharing

- Shared tools and infrastructure, standards, and methods that are free-to-use and open to contribution to increase efficiencies and avoid duplication of effort and fragmentation.

- Stronger intermediary organizations to support the sector or its key functions to broker more sustained and coordinated investment in infrastructure and capacity from funders.

data.org's Commitment

Creation of shared language and definitions

data.org has started this work with the Data for Good Landscape produced by our fellow and co-founder of DataKind, Jake Porway. Next, data.org will be working with partners to define the competency framework for DSI practitioners.

Development of common governance, and a culture of transparency and data sharing

data.org is currently developing an MEL framework for building the talent pipeline for the DSI Field, which will be shared as free to use public good. In parallel, data.org has started experimenting with collaborative grant design, by bringing together a community of DSI field actors to collaboratively create potential concepts to improve critical areas within the field and then convening multiple funders to consider and test these early-stage ideas to identify the most actionable ones.

Deployment of shared tools and infrastructure, standards and methods that are free to use and open to contribution

data.org is working to create open source tools and methodologies across all its funded programs, including the Inclusive Growth and Recovery Challenge, Epiverse, and its Capacity Accelerator Network hubs. data.org is also strongly investing in its own platform to create products and services, such as the Data Maturity Assessment and Resource Library, that can be offered freely to the wider DSI Field. By this investment in platform and open source tools and resources, data.org is testing the model of what it means to be an intermediary player in this space.

2   More diverse and interdisciplinary purpose-driven data practitioners who can drive change locally.

The Workforce Wanted report identified an opportunity for 3.5 million jobs in the data for social impact space in developing countries over the next 10 years. To take advantage of this opportunity in a changing market, we need the following:

- Diverse data practitioners from the communities we seek to support, empowered to use their knowledge and experience to drive change.

- Interdisciplinary data teams embedded within social impact organizations

- Effective communication between disciplines through shared language and standards

- Leaders, organizations, funders, and intermediaries that proactively invest in the time, tools, and efforts required to build a workforce that can learn from the local context and embed subject matter expertise into the data lifecycle.

data.org's commitment

Creation of a pipeline of diverse data practitioners with interdisciplinary training to equip them to face afresh the moral imperative of our time

data.org, together with the Patrick J. McGovern Foundation and Dalberg, published the Workforce Wanted report, quantifying the DSI talent gap and outlining recommendations on how to improve the throughput of talent across four identified pathways. data.org is actively working with partners like U Chicago, GPSDD, and J-PAL South Asia to launch three capacity accelerator network hubs, in US, Africa, and India, respectively, that will create novel applied and interdisciplinary curricula and teach DSI experts both at masters level and via professional development courses. data.org will use its platform to share the curricula and courses and the playbook of how to operate hubs online for free to help amplify the impact. Further hubs and programs are in development, including dedicated courses for SIO leaders, to improve their data skills and awareness. At a recent Bellagio convening on DSI talent, data.org brought together 20 other key actors in this space, from academic institutions to social impact organizations to tech companies to funders, to accelerate action towards our ambitious ten-year program to train one million purpose-driven data practitioners.

3 Stronger funding models with longer time horizons, more flexible funding, and better coordination.

The DSI sector is largely supported by grant-based funding from philanthropic, government, and non-governmental organizations. As our research has shown, many grants are narrow in scope, focused on short term objectives, and constrained by only supporting what can be measured by sometimes outdated frameworks. We must move to a more collaborative funding model, increasing coordination both across verticals within individual funding organizations and across funding partners, more generally.

To achieve these improvements, we must:

- Provide funders with actionable evidence of the need for and advantage of long-term, sustained investment in talent and technology ecosystems.

- Encourage funders to follow their own principles and share data and analysis with each-other to enable a higher degree of systematization in cross-funder decision making.

- Provide clear pathways for funders to invest in longer-term, systems-level interventions, shared infrastructure programs, and intermediary actors that work across sectors and geographies.

data.org's commitment

Creation of stronger funding models with longer time horizons, more flexible funding, and better coordination

data.org has convened a network of funders interested in building the DSI field through collaborative funding processes. These processes start with sharing data and analyses and brokering the creation of joint investments, and build on the success of joint investment in data.org's own key programs: the Inclusive Growth and Recovery Challenge, Epiverse, and CAN, each of which is supported by multiple funders to avoid single points of failure.
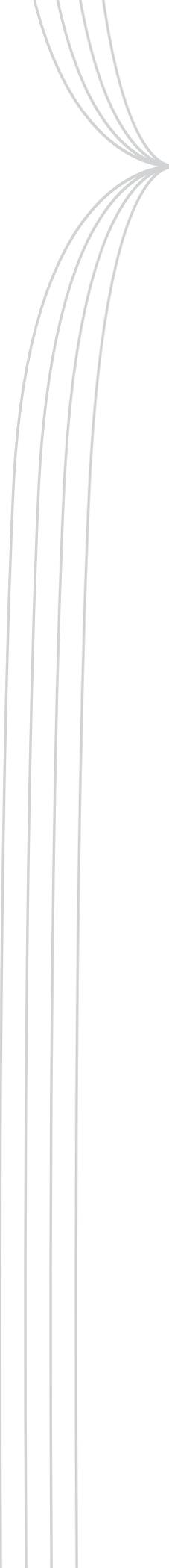
ANNEXES

Definitions

AI localism: the actions taken by local decision-makers to address the use of AI within a city or community

Community based organizations: refers to organizations aimed at making desired improvements to a community's social health, well-being, and overall functioning. Community organization occurs in geographically, psychosocially, culturally, spiritually, and digitally bounded communities.

Data: individual facts, statistics, or items of information

Data ecosystem: A context in which several actors interact with a collection of data infrastructure, analytics, and applications that are evolving in real time.

Data lifecycle: A journey recognizing end-to-end data usage by any organization, including generating, collecting, processing, storing, managing, analyzing, visualizing, and interpreting data.

Data maturity: Data maturity is a way to measure the extent and sophistication of data usage within an organization. It recognizes that organizations have different levels of experience, particularly with defining data purpose, understanding data value, and establishing strategies and management practices for data.

Data professional: A person that contributes to any part of the data lifecycle as a primary or significant part of their professional activity, including generating, collecting, processing, storing, managing, analyzing, visualizing, or interpreting data. For the purposes of this report, we are specifically focusing on data professionals who have and use intermediate and advanced skills as a primary part of their professional activity.

Data science: Data science is an interdisciplinary field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from structured and unstructured data and apply knowledge and actionable insights from data across a broad range of application domains.

Data for social impact: The use of data by social impact organizations to benefit organizations, people, their communities, and the environment more efficiently and effectively.

Data talent for social impact/purpose-driven data professionals: Data professionals who work either as paid employees or volunteers at an organization, institution, or company that is oriented toward social impact.

Intellectual property: includes inventions, art, and other works of "intellect." Intellectual property can be protected under rights including but not limited to patents, copyrights, trademarks, and/or trade secrets.

Measurement, evaluation, and learning (MEL): The purpose of measurement, evaluation, and learning practices is to apply knowledge gained from evidence and analysis to improve development outcomes and ensure accountability for the resources used to achieve them.

Non-governmental organizations: nonprofit entities independent of governmental influence.

Social impact organization (SIO): An organization (for- or nonprofit) working consciously, systematically, and sustainably to address social, environmental, economic, health, or related challenges to drive positive and desired change. Typically, an SIO works to serve marginalized groups. For the purposes of the report, SIO will refer only to organizations that are registered as nonprofits.
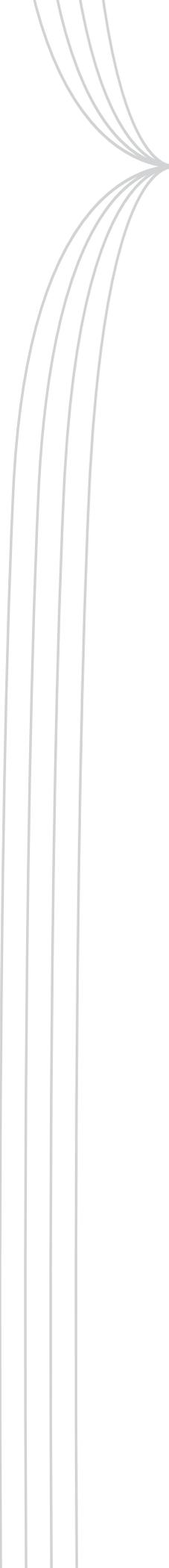
Overview of Methodology

data.org was established in 2020 as a platform to build the field of data for social impact. Since then, this outsized remit has offered the opportunity to engage diverse partners from across continents, learn from social impact organizations across domains, engage a cross-section of philanthropic, academic, private sector, and government partners at the city, state, national, and international level, and build partnerships with bilaterals, international NGOs, and data for social impact innovators.

The Data for Social Impact Sector report is the result of a research effort that was launched in July 2022 to build upon our knowledge and the growing body of evidence in the data field and in the nascent DSI space. More specifically, this study consisted of the following activities over a duration of six months:

- Analysis and review of more than 775 data maturity assessment (DMA) results from SIOs, philanthropies, private sector organizations, and academic institutions. The DMA offers organizations a snapshot view of their data maturity journey today, and relevant tools and resources required to move forward.

- Stakeholder interviews with ecosystem actors. We interviewed over 50 stakeholders across several categories, including training providers, funders, and leaders of social impact organizations, as well as experts in the field of DSI.

- Insights and information from over 1,200 organizations using data for social impact that submitted applications to the 2020 Inclusive Growth and Recovery Challenge, helping us identify both common pitfalls and outstanding exemplars globally.

- A literature review of more than 80 studies, articles, and research reports to understand the existing data on, analysis of, characteristics of, knowledge of, and discourse about the DSI space.
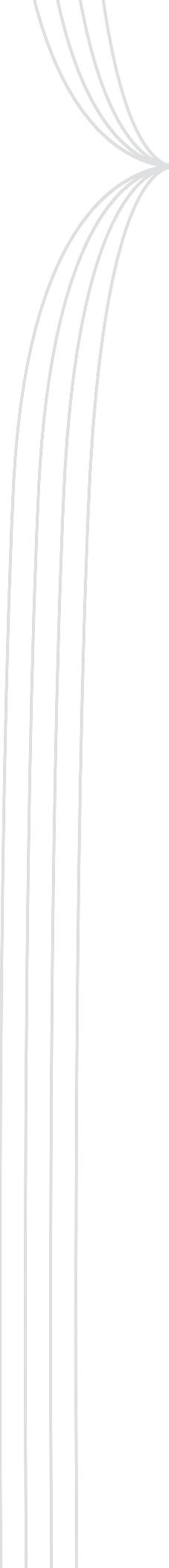
Data Maturity Assessment

We analyzed more than 775 DMA results, which offers SIOs a snapshot view of their data maturity journey today, and relevant tools and resources to move forward.

The following types of organizations completed assessments to identify opportunities for growth within their work:

- SIOs ready to deploy data and technology for significant results and looking for an entry point and neutral resources to help them along their journey

- Philanthropies and private sector organizations seeking strategic ways to invest in social impact through funding and technical assistance

- Academic institutions seeking to identify needs and propose solutions for better use of data and technology to tackle social problems

The data.org DMA provides a framework for assessing organizational data maturity within three categories:

1 **Purpose**:
   What does the organization want to use data to do? Includes the strategy, application, and analysis of data.

2 **Practice:**
   How does the organization plan to use data to achieve its mission? Includes the quality, security, ethics, and infrastructure in place to work with data.

3 **People:**
   Who works with data and makes data–driven decisions? Includes the leadership, talent, and culture of an organization.
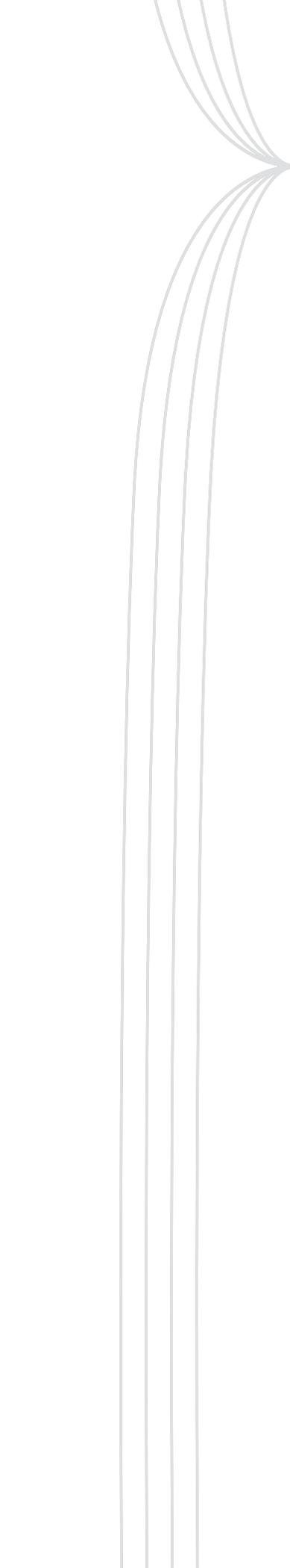
Data Maturity Assessment: Key Findings

Organizations do have access to data:

- 86 percent of respondents report that they collect data from clients or end users

- 57 percent of respondents report taking advantage from external data sources

Organizations need help deriving value from the data they have access to:

- 79 percent of respondents feel they have the technology or tools to collect data while only 65 percent feel they had tools to conduct analysis

- More than 52 percent of respondents report that their organizations only sometimes, never, or rarely use the data they have to better understand their programs

- More than 50 percent of respondents report that they struggle to use data and evidence to design programs and services using data they have collected

Organizations need to invest in data security and responsible data practices:

- Only 30 percent of respondents would describe their data security policies and practices as robust

- Only 51 percent of respondents report that they employ at least one person responsible for data security

- More than 58 percent of respondents report that they never, rarely or only sometimes conduct benefits–risks–harms assessments for the data they collect

Organizations need to continue to invest in data talent:

- Only 52 percent of respondents report that they have a trusted data expert on senior leadership

- More than 70 percent of respondents report that they do not or are unsure that there is at least one data expert on their Board of Directors

- 53 percent of respondents report that the leadership is aware of the value of data, but does not always use it to make decisions

- 90 percent of organizations report that their organization is fully or somewhat committed to investing in data tools, training and staff

List of Interviews

We conducted over 46 stakeholder interviews across several categories, including training providers, funders, and leaders of social impact organizations, as well as experts in the field of DSI or in adjacent fields, focusing on the following three lines of questions:

1   The Field Today
What are the key trends in the data for the social impact field today?

2   Exploring Tension
What are the natural tensions that will play out in the ecosystem that will shape how we as a sector will use data to solve social impact problems?

3   Looking Forward
What would you recommend to organizations seeking to improve their use of data? How do you think investors need to use their power and influence to shape this growing field?

Bayo (Olubayo)	Adekanmbi	Founder	Data Science Nigeria
Davis	Adieno	Director of Programs	Global Partnership for Sustainable Development Data (GPSDD)
Nathaly	Alarcón	Data Science Engineer Team Lead and Women in Data Science Ambassador at La Paz	Mojix, Women in Data Science (WiDS)
Carol	Andrade Bussacos	Co-founder & Senior Advisor	Social Good Brasil
Abhijit	Balakrishnan	ICT Business Analyst, Enterprise Architecture	SOS Children's Villages International
Mamadou	Biteye	Executive Secretary	The African Capacity Building Foundation
Afua	Bruce	Founder & Principal	ANB Advisory Group LLC
Stuart	Coleman	Director	Open Data Institute (ODI)

Shanna	Crumley	Director, Impact Data Science	Mastercard Center for Inclusive Growth
Zulma M.	Cucunubá	Assistant Professor of Epidemiology	Pontificia Universidad Javeriana (Javeriana)
Rodrigo	de Reyes Lanfranco	Country Managing Director	Fundación Capital
Kriss	Deiglmeier	Chief Global Impact Officer	Splunk
Vilas	Dhar	President	McGovern
Arturo	Franco	Senior Vice President	Mastercard Center for Inclusive Growth
Haishan	Fu	Director, Development Data Group, Development Economics	World Bank
Rayid	Ghani	Professor	Carnegie Mellon
Nelson	González	Head, Global Impact Computing	Amazon Web Services (AWS)
Catalina	González–Uribe	Director of Internationalization at the Vice Presidency of Research and Creation and Associate Professor at the School of Medicine	Universidad de los Andes (Uniandes)
Claudia	Juech	Data Philanthropy Advisor	Independent
Zia	Khan	Senior Vice President, Innovation	The Rockefeller Foundation
Deepali	Khanna	Vice President	The Rockefeller Foundation
Tariq	Khokhar	Head of Data for Science and Health	Wellcome
George	Kibala Bauer	Director, Digital Utilities	GSMA

Aparna	Krishnan	Project Director	The Abdul Latif Jameel Poverty Action Lab (J-PAL), South Asia
Juan M.	Lavista Ferres	VP, Chief Scientist	Microsoft AI for Good
Donald	Lobo	Executive Director	Chintu Gudiya Foundation and Advisory Board for IDinsight
Temina	Madon	Member	The Agency Fund
Bilal	Mateen	Senior Manager, Digital Technology	Wellcome
Juan	Mateos-Garcia	Director	NESTA
Claire	Melamed	CEO	Global Partnership for Sustainable Development Data
Jackie	Mwaniki	Energy Sector Lead	Fraym
Neera	Nundy	Co-founder and Partner	Dasra
Angela	Oduor Lungati	Executive Director	Ushahidi
Lance	Pierce	CEO	NetHope
Ben	Pierson	Deputy Director Enterprise Data	Bill & Melinda Gates Foundation
Frederic	Pivetta	Co-Founder and CEO	Dalberg Data Insights
Chaitali	Sinha	Senior Program Specialist	International Development Research Centre (IDRC)
Jenny	Tran	Head of Data	Paul Ramsay Foundation
Mahadia	Tunga	Co Founder and Director Data Science, Research and Capacity Development	Tanzania Data Lab (dLab)

David	Uminsky	Executive Director, Data Science Institute; Senior Research Associate, Department of Computer Science	University of Chicago
Stefaan	Verhulst	Co-Founder	GovLab
Satyam	Vyas	Founder and Chief Executive Officer	Arthan
Chris	Worman	Co-Founder, Chief Partnership and Strategy Officer (Connect Humanity), Founder and Board President (TechSoup)	Connect Humanity, TechSoup Romania
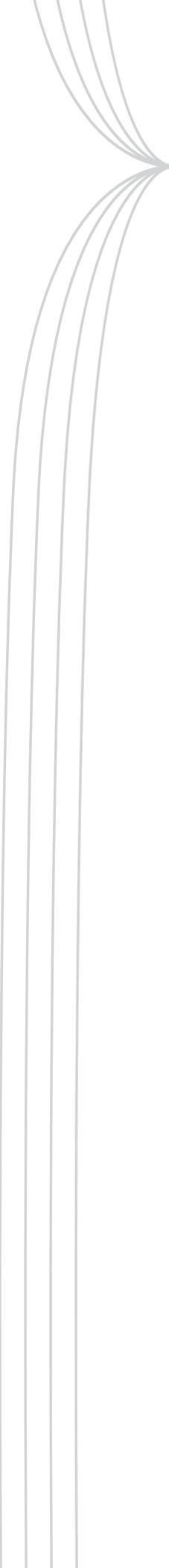
Literature Review

Aaronson, S, A. (2022). "A Future Built on Data: Data Strategies, Competitive Advantage and Trust." CIGI. https://www.cigionline.org/publications/a-future-built-on-data-data-strategies-competitive-advantage-and-trust/

Aaronson, S, A. (2022b). "A New Approach to Digital Public Goods Is Gaining Steam." Barron's. https://www.barrons.com/articles/digital-public-goods-global-trust-51651775649?tesla=y

Abebe, R., Aruleba, K., Birhane, A., Kinsley, S., Obaido, G., Remy, S, L., Sadagopan, S. (2021). "Narratives and Counternarratives on Data Sharing in Africa: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency." ACM. https://doi.org/10.1145/3442188.3445897

Ada Lovelace Institute. (2021). "Participatory data stewardship." https://www.adalovelaceinstitute.org/report/participatory-data-stewardship/

Ada Lovelace Institute, AI Now Institute and Open Government Partnership. (2021). "Algorithmic Accountability for the Public Sector." https://www.opengovpartnership.org/documents/ algorithmic-accountability-public-sector/

Aiken, E., Bellue, S., Karlan, D., Udry, C., & Blumenstock, J. (2022). "Machine Learning and Phone Data Can Improve Targeting of Humanitarian Aid." Nature. https://www.nature.com/articles/s41586-022-04484-9

Arsenyan, J. & Roloff, J. (2022). "The People Versus The Algorithm: Stakeholders and AI Accountability." Academy of Management. https://doi.org/10.5465/AMBPP.2022.14042abstract

Basker, S. & Gosling, M. (2021). "State of the Sector Data Maturity: In the Not-for-Profit Sector 2020." https://static1.squarespace.com/static/5d514d1775e9c90001345670/t/610d3ccc04c8d36395021cfb/1628257509568/Data+Orchard+-+State+of+the+Sector+-+Data+Maturity+in+the+Not-For-Profit+Sector+2020.pdf

Brockman, B., Hersh, S., Hoyer Gosselink, B., Maaganza, F., & Berman, M. (2021). "Investing in AI for Good." Stanford Social Innovation Review. https://ssir.org/articles/entry/investing_in_ai_for_good

Carlson, T. & Deiglmeier, K. (2022). "Here's how we can bridge the 'data divide' for a more equitable future." World Economic Forum. https://www.weforum.org/agenda/2022/01/bridging-the-data-divide-for-a-more-equitable-future/

Conner Snibe, A. (2006). "Drowning In Data." Stanford Social Innovation Review. https://ssir.org/articles/entry/drowning_in_data

Daniel, M, L. (2022). "How Does the Public Sector Identify Problems It Tries to Solve with

AI?." Tech Policy Press. https://techpolicy.press/how-does-the-public-sector-identify-problems-it-tries-to-solve-with-ai/

Data.org. (2021). "Advancing the Field of Data Science for Inclusive Growth: Learnings from the data.org Inclusive Growth & Recovery Challenge." https://data.org/reports/learning-insights-challenge-report/

Data.org. (2021b). "Move Slow and Fix Things: Re-Imagining the Data for Good Sector." https://data.org/news/re-imagining-the-data-for-good-sector/

Data.org & The Patrick J. McGovern Foundation. (2022). "Workforce Wanted: Data Talent for Social Impact." https://data.org/reports/workforce-wanted/

Davenport, T, H. & Patil, D, J. (2022). "Is Data Scientist Still the Sexiest Job of the 21st Century?." Harvard Business Review. https://hbr.org/2022/07/is-data-scientist-still-the-sexiest-job-of-the-21st-century

Davies, T. (2021). "Fostering Open Ecosystems Around Data: The role of data standards, infrastructure and institutions." https://www.timdavies.org.uk/2021/12/03/fostering-open-ecosystems-around-data-the-role-of-data-standards-infrastructure-and-institutions/

Davies, T., Fumega, S., & Gray, J. (2022) "The Global Data Barometer Report." Global Data Barometer. https://globaldatabarometer.org/wp-content/uploads/2022/05/GDB-Report-English.pdf

Deiglmeier, K. (2022). "We can't create shared value without data. Here's why." World Economic Forum. https://www.weforum.org/agenda/2022/05/we-cant-create-shared-value-without-data-heres-why/

Digital Rights Watch. (2020). "State of Digital Rights Report." https://www.dropbox.com/s/vz6w57pzyt4uuaa/State%20of%20Digital%20Rights%20Report%202020.pdf?dl=0

Digital Rights Watch. (2021). " Techno solutionism—very few things actually need to be an app." https://digitalrightswatch.org.au/2021/03/25/technosolutionism/

Dimegani, C. (2022). "AI in Government: Examples, Challenges & Best Practices." AI Multiple. https://research.aimultiple.com/ai-government/

Engin, Z. (2022). "Towards Human-Centric Algorithmic Governance." The Medium. https://medium.com/data-policy/towards-human-centric-algorithmic-governance-44989c34e27f

Global Partnership for Sustainable Development Data. (n.d). "Reimagining Data and Power: A Roadmap for Putting Values at the Heart of Data." https://www.data4sdgs.org/reimagining-data-and-power-roadmap-putting-values-heart-data

Floridi, L., Cowls, J., King, T, C., & Taddeo, M. (2020). "How to Design AI for Social Good:

Seven Essential Factors." Springer. https://link.springer.com/article/10.1007/s11948-020-00213-5

Google. (2019). "Accelerating social good with artificial intelligence: Insights from the Google AI Impact Challenge." Data Orchard. https://services.google.com/fh/files/misc/accelerating_social_good_with_artificial_intelligence_google_ai_impact_challenge.pdf

Harkins, L. & Proctor, B. (2022). "Data Maturity Assessment Impact Report 2022." Data Orchard. https://static1.squarespace.com/static/5d514d1775e9c90001345670/t/62bafb2ac5656a5e41b7da27/1656421170355/Data-Orchard-DMA-Impact-Report-2022.pdf

Hart, N., Karkera, A., & Logan, V. (2022). "Data Literacy for the Public Sector: Lessons from Early Pioneers in the US." Data Foundation. https://static1.squarespace.com/static/56534df0e4b0c2babdb6644d/t/6238f0b087fdb4397d0e7f61/1647898802196/2022-DataLiteracy-PublicSector-final.pdf

Hendey, L. & Pettit, K, L. (2021). "Envisioning a New Future: Building Trust for Data Use." Urban Institute. https://www.urban.org/research/publication/envisioning-new-future-building-trust-data-use

Hewitt, P. (2021). "Understanding the World Through Data." Data.org. https://data.org/news/understanding-the-world-through-data/

Hewitt, P. (2022). "Pathways to Impact: Angela Oduor Lungati." Data.org. https://data.org/news/pathways-to-impact-angela-oduor-lungati/

Hewitt, P. (2022b). "Pathways to Impact: Danielle Getsinger." Data.org. https://data.org/news/pathways-to-impact-meet-danielle-getsinger/

Hewitt, P. (2022c). "Pathways to Impact: Evan Tachovsky." Data.org. https://data.org/news/pathways-to-impact-evan-tachovsky/

Hewitt, P. (2022d). "Pathways to Impact: George Kibala Bauer." Data.org. https://data.org/news/pathways-to-impact-meet-george-kibala-bauer/

Hewitt, P. (2022e). "Pathways to Impact: Ivana Feldfeber." Data.org. https://data.org/news/pathways-to-impact-ivana-feldfeber/

Hewitt, P. (2022f). "Pathways to Impact: Miguel Luengo-Oroz." Data.org. https://data.org/news/pathways-to-impact-meet-miguel-luengo-oroz/

Hewitt, P. (2022g). "Pathways to Impact: Nate Matias." Data.org. https://data.org/news/pathways-to-impact-nate-matias/

Hewitt, P. (2022h). "Pathways to Impact: Neera Nundy." Data.org. https://data.org/news/pathways-to-impact-meet-neera-nundy/

Hewitt, P. (2022i). "Pathways to Impact: Roberta Evangelista." Data.org. https://data.org/news/pathways-to-impact-meet-roberta-evangelista/

Holshof Schmidt, R. (2022). "Managing Nonprofit Technology Change: 2022 Report." NTEN. https://word.nten.org/nten-login/?redirect_to=https%3A%2F%2Fword.nten.org%2Fwp-content%2Fuploads%2F2022%2F03%2FManaging-Nonprofit-Tech-Change-Report_March-2022.pdf

Hoyer Gosselink, B. & Bromberg, C. (2019). "2,602 Uses Of AI For Social Good, And What We Learned From Them." Google. https://blog.google/outreach-initiatives/google-org/2602-uses-ai-social-good-and-what-we-learned-them/

Independent Sector. (2021). "Health of the U.S. Nonprofit Sector." https://independentsector.org/wp-content/uploads/2022/07/sector-health-report-2021-101421-1.pdf

Independent Sector. (2022). "Health of the U.S. Nonprofit Sector." https://independentsector.org/wp-content/uploads/2022/10/Quarterly-Health-Report-Sep-2022_v4-1.pdf

Johnson, K. (2021). "Buidling AI for the Global South." VentureBeat. https://venturebeat.com/ai/building-ai-for-the-global-south/

Kalkar, U., Marcucci, S., Mansuri, S., & Verhulst, S. (2022). "What does AI Localism look like in action? A new series examining use cases on how cities govern AI." The GovLab. https://blog.thegovlab.org/post/ailocalism-blog-0

Koponen, J, M. (2021). "Forget User Experience. AI Must Focus On 'Citizen Experience.'" VentureBeat. https://venturebeat.com/ai/forget-user-experience-ai-must-focus-on-citizen-experience/

Krasodomski-Jones, A., Smith, J., Judson, E., Baines, V., Achieng', G., & Ansari, M. "The Open Road: How to Build a Sustainable Open Infrastructure System." Omidyar Network. https://omidyar.com/wp-content/uploads/2022/06/The-Open-Road-1_report.pdf

Kshirsagar, M., Robinson, C., Yang, S., Gholami, S., Klyuzhin, I., Mukherjee, S., Nasir, M., Ortiz, A., Oviedo, F., Tanner, D., Trivedi, A., Xu, Y., Zhong, M., Dilkina, B., Dodhia, R., & Ferres, J. M. L. (2021). "Becoming good at AI for good." https://arxiv.org/abs/2104.11757

Madianou, M. (2021). "Nonhuman humanitarianism: when 'AI for good' can be harmful." Taylor & Francis Online. https://doi.org/10.1080/1369118X.2021.1909100

Margonelli, L. (2022). "On the Power of Networks." Issues In Science and Technology. https://issues.org/power-of-networks-editors-journal-margonelli/

Microsoft. (2022). "Open Data for Social Impact Framework." https://news.microsoft.com/open-data-social-impact-framework/

Mikhailov, D. (2022). "Key Trends in AI and the Case for a Social Impact Focus." Data.org. https://data.org/news/key-trends-ai/

Mulgan, G. (2021). "The Social Economy and the Fourth Industrial Revolution." Stanford Social Innovation Review. https://ssir.org/articles/entry/the_social_economy_and_the_fourth_industrial_revolution

National Commission To Transform Public Health Data Systems. (2021). "Charting a Course for an Equity-Centered Data System." Robert Wood Johnson Foundation. https://www.rwjf.org/en/library/research/2021/10/charting-a-course-for-an-equity-centered-data-system.html

National Council of Nonprofits. (2019). "Nonprofit Impact Matters: How America's Charitable Nonprofits Strengthen Communities and Improve Lives." https://www.nonprofitimpactmatters.org/site/assets/files/1/nonprofit-impact-matters-sept-2019-1.pdf

NESTA. (2022). "Designing the Collective Intelligence Commons." https://www.nesta.org.uk/report/designing-the-collective-intelligence-commons/

Novitske, L. (2018). "The AI Invasion is coming to Africa (and its a good thing)." Stanford Social Innovation Review. https://ssir.org/articles/entry/the_ai_invasion_is_coming_to_africa_and_its_a_good_thing

NTEN. (2020). "Equity Guide for Nonprofit Technology." https://word.nten.org/wp-content/uploads/2021/07/NTEN-Equity-Guide-for-Nonprofit-Technology_September_2020_v2.pdf

ODI. (2022). "Experimentalism and the Fourth Industrial Revolution." https://theodi.org/wp-content/uploads/2022/11/ODI_Experimentalism-and-the-Fourth-Industrial-Revolution_Final-Report.pdf

OECD., OPSI., & CAF. (2022). "The Strategic and Responsible Use of Artificial Intelligence in the Public Sector of Latin America and the Caribbean." https://oecd-opsi.org/publications/ai-lac/?utm_

Oxford Insights, The International Development Research Centre. (2019). "Government Artificial Intelligence Readiness Index 2019." https://www.oxfordinsights.com/ai-readiness2019

Pascu, C. & Burgelman, J, C. (2022). "Open Data: The building block of 21st century (open) science." Cambridge University Press. https://www.cambridge.org/core/journals/data-and-policy/article/open-data-the-building-block-of-21st-century-open-science/E7D3B6EE5F05606D92A69FA55F76E001

Porway, J. (2021). "Charting the 'Data for Good' Landscape." Data.org. https://data.org/news/charting-the-data-for-good-landscape/

Porway, J. (2022a). "A Taxonomy for AI/ Data for Good." Data.org. https://data.org/news/a-taxonomy-for-ai-data-for-good/

Porway, J. (2022b). "Funding Data and AI that Serve the Social Sector." Stanford Social Innovation Review. https://ssir.org/articles/entry/funding_data_and_ai_that_serve_the_social_sector

Ramanathan, N., Fruchterman, J., Fowler, A., & Carotti-Sha, G. (2022). "Fighting Data Colonialism Through Community Ownership, Consent, and Privacy." Stanford Social Innovation Review. https://ssir.org/articles/entry/decolonize_data

Rosenblatt, G. & Gupta, A. (2018). "Artificial Intelligence as a Force for Good." Stanford Social Innovation Review. https://ssir.org/articles/entry/artificial_intelligence_as_a_force_for_good

Salamon, L, M. & Sokolowski, S, W. (2004). "Global Civil Society: Dimensions of the Nonprofit Sector." Johns Hopkins Comparative Nonprofit Sector Project. https://ccss.jhu.edu/wp-content/uploads/downloads/2011/08/Global_Civil_Soiciety_2_TOC.pdf

Salesforce. (2021). "2022 Nonprofit Trends Report." Salesforce.org. https://www.salesforce.org/resources/report/nonprofit-trends-report/

Savage, N. (2020). "The Race to the Top Among the World's Leaders in Artificial Intelligence." Nature. https://www.nature.com/articles/d41586-020-03409-8

Saxena, S. & Verhulst, S, G. "10 learnings from considering AI Ethics through global perspectives." Medium. https://medium.com/data-stewards-network/10-learnings-from-considering-ai-ethics-through-global-perspectives-3b6f38d88f0f

Schnurer, E, B. (2022). "Democracy Disrupted: Governance in an Increasingly Virtual and Massively Distributed World." The Hedgehog Review. https://hedgehogreview.com/issues/the-use-and-abuse-of-history/articles/democracy-disrupted

Scutari, M. (2022). "Funders Set Out to Advance Public Interest Technology. What's the State of the Field?." Inside Philanthropy. https://www.insidephilanthropy.com/home/2022/7/21/funders-set-out-to-advance-public-interest-technology-whats-the-state-of-the-field?utm_content=bufferca225&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer

Selbst, A, D., Boyd, D., Friedler, S, A., Venkatasubramanian, S., & Vertesi, J. (2019). "Fairness and Abstraction in Sociotechnical Systems." Association for Computing Machinery. https://doi.org/10.1145/3287560.3287598

Shi, Z, R., Wang, C., Fang, F. (2020). "Artificial Intelligence for Social Good: A Survey." Carnegie Mellon University. https://arxiv.org/pdf/2001.01818.pdf

SSIR Editors. (2021). "Recap of 2021 Data on Purpose Conference." Stanford Social Innovation Review. https://ssir.org/articles/entry/2021_data_on_purpose_reading_list_and_recap

TechSoup Global Network. (2021). "Data Handling and Digital Readiness in Civil Society: Global Study 2020." https://page.techsoup.org/hubfs/Downloads/data-handling-survey-2021.pdf?hsCtaTracking=765fc28f-bb2f-42d3-9912-725122e398fb%7Caf37af40-d4fe-47b0-94d3-312f6291f519

University of Oxford. (2022). "NGOs must rapidly evolve to stay relevant, say NGO leaders – Oxford survey." https://www.ox.ac.uk/news/2022-07-14-ngos-must-rapidly-evolve-stay-relevant-say-ngo-leaders-oxford-survey

Vilsack, R. & Laprad, J. (2022). "Data For An Inclusive Economic Recovery." National Skills Coalition. https://nationalskillscoalition.org/wp-content/uploads/2022/05/FINAL-for-Web-Data-for-an-Inclusive-Economic-Recovery.pdf

Whittaker, M. (2021). "The Steep Cost of Capture." Association for Computing Machinery. https://interactions.acm.org/archive/view/november-december-2021/the-steep-cost-of-capture

World Bank Group. (2021). "Data For Better Lives: World Development Report 2021." The World Bank. https://www.worldbank.org/en/publication/wdr2021

World Economic Forum. (2022). "Global Gender Gap Report 2022: Insight Report." https://www3.weforum.org/docs/WEF_GGGR_2022.pdf

World Economic Forum. (2022b). "A Blueprint for Equity and Inclusion in Artificial Intelligence." https://www3.weforum.org/docs/WEF_A_Blueprint_for_Equity_and_Inclusion_in_Artificial_Intelligence_2022.pdf

Worman, C. (2022). "Funding to bridge the digital divide: U.S. philanthropic giving to digital equity causes." Connect Humanity. https://connecthumanity.fund/wp-content/uploads/2022/09/ConnectHumanity_PhilanthropyFundingReport_Final.pdf

Zielinskie, G. & Gottshalk, L. "Rising Equitable Community Data Ecosystems (RECoDE)." Data.org. https://data.org/wp-content/uploads/2022/02/ReCode-Report.pdf

Zoli, A. (2018). "After Big Data: The Coming Age of "Big Indicators." Stanford Social Innovation Review. https://ssir.org/articles/entry/after_big_data_the_coming_age_of_big_indicators