

Workforce Wanted

DATA TALENT FOR SOCIAL IMPACT

Executive Summary

data.org



Patrick J McGovern
FOUNDATION

Dalberg

FOREWORD



Data is everywhere; but how social impact organizations tackling the most complex problems of our day effectively unlock its potential remains a stubborn challenge worthy of our collective action.

Undoubtedly, data is an essential tool to better solve important global problems like climate change, healthcare disparities, food insecurity, humanitarian emergency response, lack of financial inclusion, and systemic discrimination in all its forms. To unleash the power of data across the social impact sector, we must advance how we identify, cultivate, support, and retain diverse data professionals.

Easier said than done. There is a persistent workforce shortage, with equally persistent barriers that exclude women and people of color from the field. Data infrastructure within social impact organizations is often underdeveloped due to lack of meaningful investment from philanthropy and organizations themselves. At the same time, competition is fierce for the skills needed to better use data, as sectors of every kind are competing to attract, upskill, and retain data talent.

Determination to overcome these challenges, coupled with the exciting opportunity to transform how the social impact sector uses data, is why data.org and the Patrick J. McGovern Foundation partnered on a landscape analysis. We ventured into this work to better understand the current data talent training landscape within the social impact sector, uncover opportunities for collective action, and develop a set of recommendations for global partners across academia, philanthropy, government, social impact, and the tech for good sector to build a diverse and inclusive data talent ecosystem. We have a unique chance to build pipelines to opportunity and impact that are filled with talented individuals representative of the societies and communities that have previously been excluded.

Spoiler alert: to be effective, the solutions and recommendations uncovered require all of us to act in concert.

Additional learning is needed, capital is required, and an authentic and unwavering commitment to building a diverse and inclusive data talent pool is essential. We remain energized by the work ahead. There are innovative and replicable ways that organizations that tackle entrenched problems are growing data talent. New talent is emerging, and they are eager to apply their skills and expertise to problem statements that matter. Talent within social impact is looking for pathways to their data skills. Finally, and perhaps most critically, leaders are beginning to understand the potential they can unlock if their organizations use data better to achieve impact.

That is where you come in. The purpose of this report is to shed light on the scale of the opportunity in front of us all, better understand its dimensions, and catalyze a wide array of people and institutions to work together and be deliberate about how we build a diverse data talent pool. Together, we can create an ecosystem that connects the right skills to the right problems and supports people across their careers to develop and strengthen the diverse skills necessary to effectively establish the field of data for social impact.

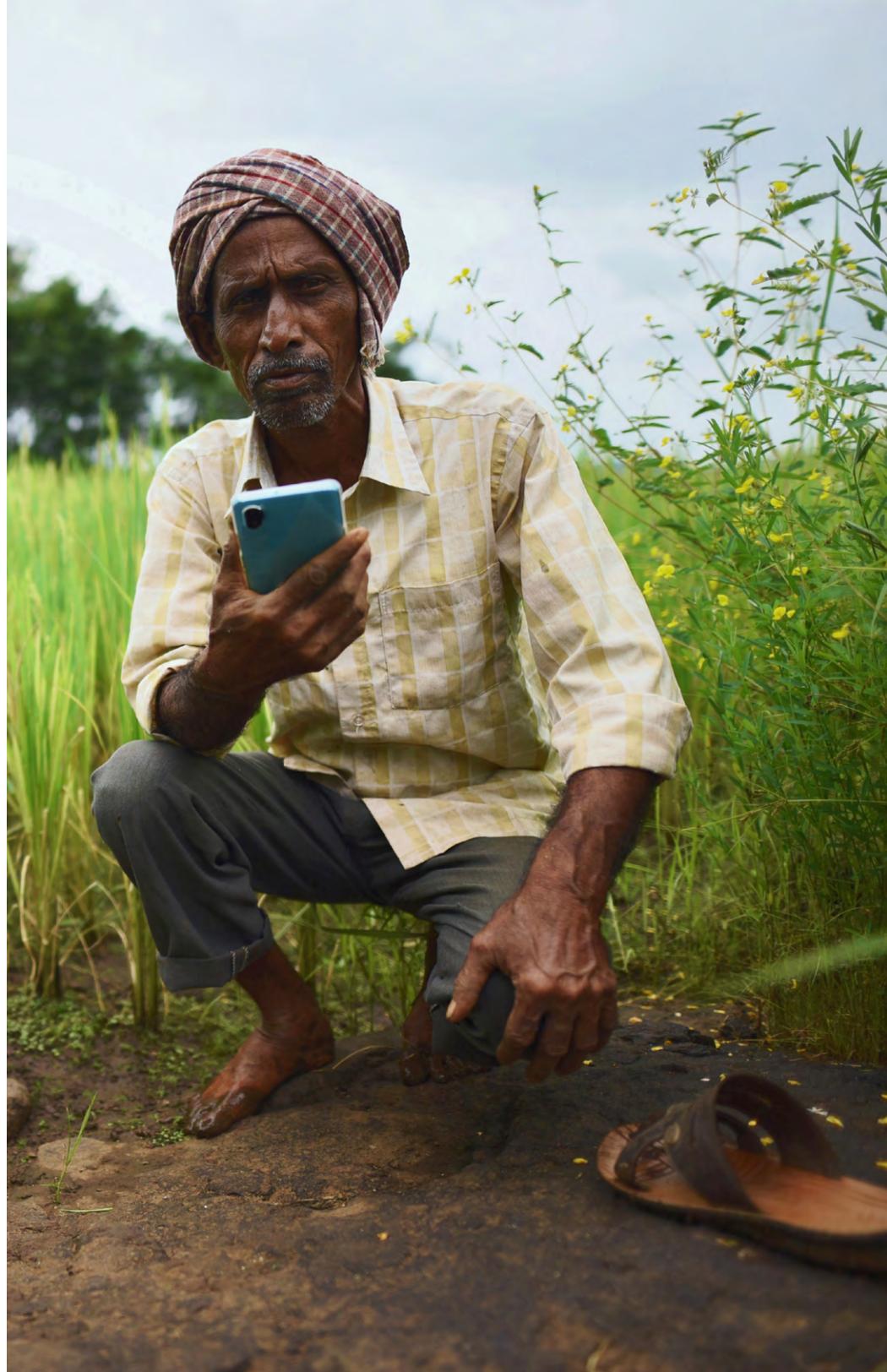
Claudia Juech
Vice President of Data and Society
Patrick J. McGovern Foundation

Ginger Zielinskie
Chief Strategy Officer
data.org

June 2022

ABBREVIATIONS

AI	artificial intelligence
BSc	Bachelor of Science
CEO	chief executive officer
CIO	chief information officer
CTO	chief technology officer
DSI	data for social impact
DQM	data quality management
GDP	gross domestic product
HR	human resources
IDEA	inclusion, diversity, equity, and accessibility
MEL	measurement, evaluation, and learning
MOOC	massive open online course
MSc	Master of Science
SIO	social impact organization
STEM	science, technology, engineering, and mathematics
WEF	World Economic Forum



EXECUTIVE SUMMARY





This is a complex and pivotal moment for harnessing the global impact of data. On the one hand, we see tremendous innovation across public, private, and social sectors, and broad recognition of the critical role data is playing in our daily lives. On the other hand, the industry struggles to understand, unlock, and keep pace with demand—all while navigating serious questions about who is considered when gathering, shaping, and using data, and how we ensure that the data field is enriching human lives and livelihoods, not diminishing them. Lack of resources, technical capacity, and data governance hamper the production of useful data for public policy.¹

Amid this complexity, data.org and the Patrick J. McGovern Foundation (PJMF) believe that there is significant untapped potential to shape the field of data more broadly to tackle the world's most pressing issues in developing contexts.

Quality talent is central to this goal. data.org and PJMF recognize that building the next generation of diverse data talent for social impact is essential for us to effectively address social challenges. This first-of-its-kind report focuses on understanding what is needed to cultivate a diverse, equitable, and inclusive talent base of data professionals that are driving social impact in developing contexts around the world.

¹ <https://www.worldbank.org/en/publication/wdr2021>

This study aims to do three things:

1. Bring visibility to an emerging pool of talent: data professionals focused on social impact in developing contexts;
2. Explore the potential to accelerate this labor market segment, particularly when it comes to inclusion, diversity, equity, and accessibility (IDEA); and
3. Offer recommendations for like-minded efforts to dramatically grow and expand access to purpose-driven data professionals around the world.

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

- What do we know about the current talent landscape and the challenges it faces—particularly in developing contexts, and with an emphasis on inclusion, diversity, equity and accessibility?
- What are the immediate-to-big-picture opportunities for this space?
- What are the pathways for attracting, building, and supporting data professionals focused on social impact? What systemic challenges do the data field, and data for social impact (DSI) ecosystems face? What support, enhancement, or new solutions could accelerate innovation, growth, and scale?
- How can we align interests and develop recommendations for a broad range of committed actors who will foster the growth and expansion of inclusive, diverse, equitable, and accessible data talent labor markets around the world?

KEY FINDINGS OF THIS REPORT

Opportunity

We believe there is an opportunity to shape and develop 3.5 million data professionals focused on social impact in developing countries over the next 10 years. This opportunity is based on multiple factors and includes variation based on different scenarios.²

- The current landscape for purpose-driven data professionals is nascent in terms of its overall size and organization, but shows momentum and growth, fueled by a number of intrinsic and extrinsic factors—including increasing access to the internet and data itself; acceleration of digital transformation efforts around the world; recognition of the value of inclusion, diversity, equity, and access; acceleration of digitization forced by COVID-19; elevation of and investment in advancing social issues like the SDGs, climate, public health, etc.; exceedingly high demand for data professionals globally; and many others.
- Alignment of efforts, partnership, and resilient digital infrastructure is required. There are several complementary efforts that offer the potential to galvanize shared goals when it comes to unlocking DSI talent.
 - » The momentum and investment in supporting digital transformation strategies around the world—particularly in low- and middle-income countries (LMICs)—offers opportunity to further advance structural investments in shaping labor markets for data-driven skills and professionals.
 - » The advancement of research institutions and academia in understanding the opportunities and realities of IDEA in the data ecosystem can help push the data field, including DSI, toward a systematic emphasis on IDEA.
 - » Digital transformation efforts focus on holistic, meaningful connectivity solutions (including cost of data, devices, and enabling infrastructure such as identity, payments, and asset registries), which are critical to addressing the digital divide and unlocking meaningful access to opportunities.

² In sizing the data talent opportunity, we are considering potential, i.e. the possibility of new jobs to be added to the current trajectory, rather than a current employment gap. In order to reach this potential—whereby the percentage of the workforce in the social sector in developing contexts that occupies data roles equals that of developed contexts — it will be necessary also to stimulate the demand for DSI talent. We recognize that these analyses have several limitations, e.g. attempting to size the opportunity across “developing contexts” and for the “entire social sector” within a nascent/emerging field. Most developing countries’ labor market data are limited and incomplete, and data are not aggregated across countries. Moreover, there is no common taxonomy of data professionals across countries. However, we believe that this analysis is critical to perform with the evidence available to us.

The opportunities highlighted above can be realized through the high-level prioritization of data, including long-term financing, investments in human capital, and laws conducive to the safe production, exchange, and use of data. Some investments in better data have paid for themselves.

Capacity

There are a number of ways in which organizations can access and build data skills, teams, and organizational strategy.

- As a framework for considering the different approaches to building and growing skills and organizational strength, we identified four pathways:
 - » *New talent.* Expanding exposure of learners through development of DSI use cases; integration of hands-on, practical learning; incorporation of applied learning into curriculum; and stronger alignment of training models with the needs and demands of the social impact sector.
 - » *Existing talent.* Models for upskilling and reskilling—such as in-house, outsourcing, and sponsorship

models—that recognize the value of existing talent committed to social impact and SIOs.

- » *Transitional talent.* Greater exposure and access to opportunities that allow for more agile flow of talent across sectors; examples include hands-on fellowships, short courses, volunteer opportunities, and rotational leadership programs.
- » *Leadership.* Enhancing and shaping new models to support design, experimentation, and advancement of data-driven strategies, initiatives, and talent acquisition; investment in allies, such as boards and funders, to advance understanding of data-driven solutions.

The DSI field is competing within the broader data skills ecosystem, meaning that DSI professionals are often disincentivized to choose the social impact pathway, particularly when considering public-private wage gaps and limited visibility into career growth opportunities.

A review of nearly 200 data talent initiatives, a literature review of approximately 90 articles and reports, and expert interviews with more than 30 leaders in the field suggest that training and talent initiatives struggle with several systemic issues, including low levels of organizational awareness of how data can be valuable and a need for increased, sustained financing to drive shared growth, as well as:

- *Limited capacity of traditional institutions.* Traditional education programs—specifically, university science, technology, engineering, and mathematics (STEM) programs—are insufficient in terms of both the number of institutions and volume of qualified data professionals; they also lack social impact orientation.
- *Proliferation of non-traditional training models, including massive open online courses (MOOCs) and other online training platforms, lack evidence of efficacy.* Program outcomes are often disconnected from longer-term results such as job placements, and many programs demonstrate a bias towards technical training rather than integrated translational skills and work readiness.
- *Mid- and senior-level talent as both a gap and driver of growth.* Intermediate and advanced skills are underserved relative to the need for growing the talent base; training or bringing in mid- and senior-level data talent can have a multiplier effect based on leadership’s ability to shape ecosystems.
- *Ecosystem constraints.* The ability of training programs to adjust their business models is hindered by a lack of accurate market demand data for skills. While there is a growing appreciation of the value of data for social impact across public, private, and social sectors, the understanding of demand and sourcing of talent remains limited.
- *Need for leadership programs that focus on supporting leaders’ evolution rather than one-off interventions.* Leadership programs and fellowships often focus on individuals rather than holistic interventions that affect the broader data ecosystem; lessons of leadership programs are not embedded into the existing activities of the professional environment through workplace experiments and nudges.



Diversity

- While the values of inclusion, diversity, equity, and accessibility (IDEA) are generally acknowledged as important, these values have the potential to be more deliberately and comprehensively embedded in the data field today, in order to ensure responsible ethical advancement and impact.
- There may be more in common than not when it comes to DSI in “developing contexts.” The use of the phrase developing contexts, rather than developing economies, recognizes that the characteristics of developed and developing economies may not be highly variable when it comes to DSI. While economic structure may be quite different at an aggregated national level, the high levels of variation led us to push for a more granular perspective on where and how DSI talent could be unlocked. This terminology also recognizes that many contexts share common characteristics and constraints (e.g., advancing diversity, competition with private sector for skills, etc.), and that countries often have a wide range of social, economic, technological, and structural realities. As a result, while “developing context” is difficult to cleanly delineate, it signals that our work attempted to focus not on geographic or economic delineations, but rather on countries, regions, sub-regions, or other “contexts” where the infrastructure, social, or economic context is still emerging.

- Based on the findings of the research, a few recommendations emerged for actors interested in advancing the data for social impact space:
 1. *Experiment early and evaluate often.* The nature of a nascent field requires actions that will crowd in others, draw attention to what works, quickly demonstrate limitations, and facilitate frequent pivots.
 2. *Prioritize inclusion, diversity, equity, and accessibility* when considering access to education and training, links between training and placement, and absorptive capacity of maturing data ecosystems (organizations and beyond).
 3. *Recognize the interdisciplinary nature of data for social impact,* where the depth of technological understanding and expertise is matched with the discipline and understanding of social sciences.
 4. *Move from individuals to ecosystems.* Recognize the role an individual leader plays within an organization, an industry or sector, or a broader ecosystem, and align efforts and investments accordingly. Recognize the individual incentive systems already in play and the potential tensions that may exist when seeking to build new data-driven strategies or decision-making processes.
 5. *Invest in applied learning and stronger links to professional placement and advancement,* shifting from a focus on “the number of people trained” as a critical result to “the number of people playing an active role addressing social issues and working within organizations.” Consider aligning funding with intended outcomes. Financing mechanisms that link training to sustainable employment could nudge the sector in an impact-focused direction.
 6. *Coordinate complementary efforts.* Looking across various efforts linked to data, digital transformation, and the advancement of data-driven strategies for nonprofit or social-impact-oriented organizations we see significant opportunities for greater coordination to advance DSI as a field.
 7. *Continuously invest in more and better visibility through data* to illustrate how the ecosystem is functioning—in order to build on what works, better understand gaps, and track the many factors that influence outcomes.

- Because this is a new sector, we *recognize the limitations of this research and the broader ecosystem constraints*. As we developed this report, we were conscious of several challenges in mapping, sizing, and analyzing this emergent field. These limitations include:

- » Availability of granular datasets that disentangle *information* linked to the positioning and movement of talent, particularly in developing contexts, across the social sector and with regard to advancing inclusion, diversity, equity, and access.

- » *Difficulty defining “social impact.”* We know it is not simply a legal construct, such as a nonprofit organization, and that social impact manifests across all sectors, including government and the private sector. However, as a starting point, where possible, we used the nonprofit sector as a proxy for a “social impact” focus.

- » *Lack of a common language* to identify necessary data skills and roles across the public, private, and social sectors.

- » *Lack of visibility* into current and future demand, particularly latent demand.

Recognizing these constraints, we have established a set of principles to guide the analysis:

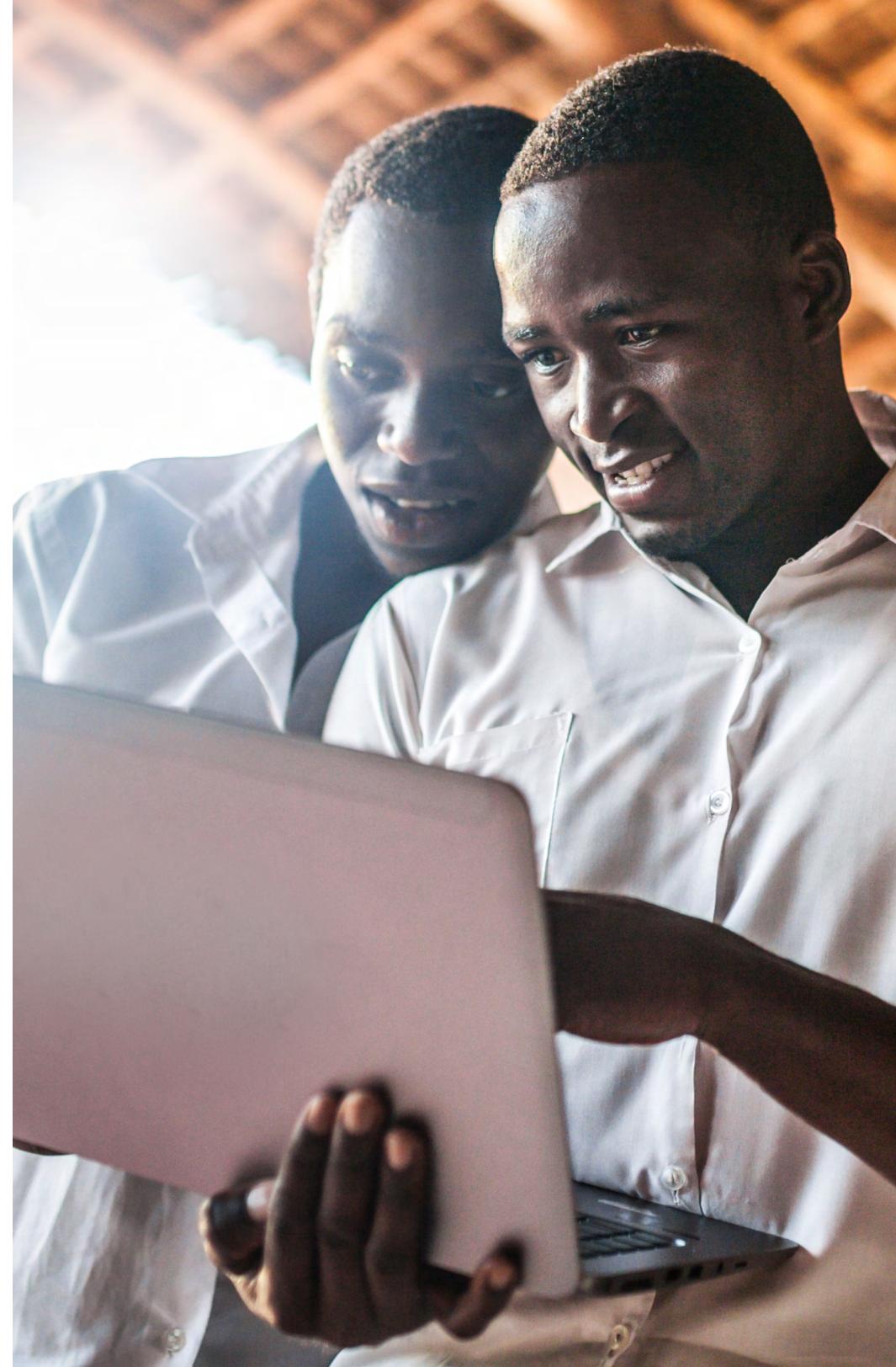


FIG 1 » GUIDING PRINCIPLES FOR ANALYZING THE TALENT LANDSCAPE

- » Reflecting on the important interplay between supply (number of DSI professionals) and demand (e.g., Social Impact Organizations' (SIOs') absorption capacity) in conjunction with limited data availability. Therefore, we use data maturity stages as a proxy for understanding demand
- » Taking a long-term view on building the DSI field, and using a portfolio perspective that combines scaling proven approaches, with active experimentation in order to iteratively learn and advance a strategy that works
- » Recognizing an overarching commitment and opportunity to build and scale a diverse community of DSI professionals by acknowledging deep systemic challenges that influence existing labor markets, as well as targeting opportunities to design new solutions that drive greater and more equitable access for all pathways
- » Anchoring on SIO leadership and senior level talent challenges and opportunities, in order to unlock the size of opportunity and build the field
- » Identifying and learning from outlier organizations (bright spots), that have managed to overcome key challenges in this space, and from analogous sectors that have faced similar barriers
- » Considering the direct and indirect results of investments, particularly as it allows for a multiplier effect on results and/or deeper, more systemic shifts

Because our intention is to grow and shape an emerging field, we believe that, in defining the talent base for DSI, we are offering a starting point from which others will enhance, build, and deepen the research. As one interviewee shared with us,

“We will never know as little about DSI as we do today; that’s the place to build from.” We look forward to the improvements and advancements that other stakeholders will develop in coming years.