CAN THE NEW DATA ECONOMY GIVE BACK TO COMMUNITIES?

CHRIS KINGSLEY

Data has earned a bad reputation within the social sector, with the most acute complaints coming from the front-line workers and communities it is meant to benefit. Teachers disparage "flawed, unfair and incomprehensible" new uses of statistics to measure their performance.[1] Social workers are "drowning" in data they are required to collect but lack the training to use.[2] And some communities, citing long histories of having been over-researched but underserved, are organizing to reassert rights over how their data are used and to insist on broader definitions of what *kinds* of data and evidence matter.[3] While it has been promoted as a tool to help organizations continually prove and improve the value of their work, data is more often associated with production management philosophies, narrow registries of evidence-based programs that come with mandatory certification from remote experts, and complex matrices of indicators imposed by different funders with competing theories of change.

It is not as though proponents of data and evidence have been running down a blind alley here. These activities have value, and a focus on measurement and outcomes imposed from the top down is *one* part of the answer to what we might call the Practitioner's Prayer ("God grant me the courage to fix programs and policies that don't work, the resources to expand those that do, and the data to know the difference"). The education and social

sectors direct more than $1 trillion each year, and it is eminently reasonable for taxpayers, practitioners, students, and other participants to know what their contributions are achieving. Moreover, there is a strong moral case for collecting the data necessary to understand and begin to correct the effects of decades of discriminatory practices and policies on the well-being of communities of color.

But there are consequences to teachers, social workers, and nonprofits relating to the use of data almost exclusively as a tool to define, limit, and control their programs and organizations rather than to interrogate, explore, and strengthen their work. A top-down approach to measurement causes expensively developed, quickly abandoned systems to proliferate in the back offices of agencies and nonprofits. It breeds cynicism among frontline staff about new data collection activities that detract from doing their jobs without returning anything of obvious practical value. And, over decades, it has eroded relationships with students, clients, and communities who are too frequently required to sign over access to so much personal data without being invited into conversations about how it's being used. (In the words of Chicago Beyond: "Why am I always being researched?"[4])

These are the complaints of results-oriented people, many of whom would agree with Mark Friedman's rallying cry that "trying hard isn't good enough" but are keenly aware that the data-driven regime that has been built around them is serving somebody else's purpose.[5]

A NEW DATA ECONOMY

The social sector's leaders have a tremendous opportunity to overhaul this broken information economy and, in so doing, put data in the service of innovation, systems reform, and rebuilding cooperation between agencies, nonprofits, and their communities. And the starting point for that transformation is to restructure the market for data and evidence.

Why do so few tools exist for families to manage their own social services profiles or to compare the efficacy of different providers? How can it be that the high-stakes testing systems built to evaluate schools and teachers do not return timely, useful management information to principals and superintendents? Much of the data infrastructure supporting services to children and families was built to the specifications of public and private funders to facilitate payment, auditing, and outcomes reporting. The occasions when these investments in better data also result in local innovation

and improvement are more the result of happy accidents than design; funding incentives are rarely aligned to sustain and scale them.

So, what does the alternative look like?

We can look first to organizations that are serious about designing tools from the perspective of clients and practitioners, incorporating—or at least emulating—the missing "market demand" of communities and nonprofits. For example, when Code for America launched its Integrated Benefits Initiative,[6] staff began as clients would, by applying to different public benefits and documenting the impediments they would have to correct to create more friendly, uniform services for families.

More than capital "R" research and evaluation, public leaders need data partners that can contribute to rapid-cycle analysis and problem solving. This is the kind of service the University of North Carolina's Charlotte Regional Data Trust provided Charlotte-Mecklenburg Public Schools when they identified hundreds of students receiving housing and homelessness services,[7] unknown to the district, and qualified them through the McKinney-Vento Homeless Assistance Act for additional funding available as well as resources like transportation services and expedited enrollment.

The need for this kind of responsive analysis is particularly acute during moments of crisis, as Colorado discovered in the early days of the COVID-19 pandemic when it turned to the state's Evaluation and Action Lab at the University of Denver to quickly connect licensed childcare workers to centers that urgently needed them[8] to serve the children of essential workers. This kind of disaster response illustrates a more broadly generalizable lesson that funders should take to heart: By the time this kind of data infrastructure becomes mission critical, it is too late to build it from scratch. It is *exploratory* analysis within the social sector, not summative reports on the result of a program, that can create the space for new thinking and different responses. A data economy that supports this kind of exploratory analysis is one that can help build, test, and scale innovative solutions. Some examples:

- During the pandemic, Los Angeles County, California, leveraged its long established integrated data infrastructure[9] to support people who were experiencing homelessness and at a greater risk of contracting COVID-19. By linking information from healthcare and homeless management information systems (HMIS) datasets, county researchers working with the University of

Pennsylvania and UCLA were able to assess discrete levels of vulnerability among the aging homeless population and propose housing and service models that matched their level of risk. They also were able to estimate potential cost offsets to Medicaid and the county that would help recapture funds needed to help stabilize people in housing.

- New York City's experience during Hurricane Sandy prepared them to be a reliable, community-engaged partner when NYC was the epicenter of the COVID-19 outbreak in 2020. The city's Center for Innovation through Data Intelligence's (CIDI) used its existing cross-agency workgroup to quickly map vulnerable populations,[10] drawing on integrated client data from NYC Health and Human Services and overlaying information on public housing, retirement communities, and shelter sites. NYC's immediate aid targeted these most at-risk populations and contributed to a more equitable response and recovery to the pandemic.

- Cuyahoga was the first county in the United States[11] to receive social impact financing after building their project on cross-agency data analysis that suggested there could be tremendous benefits—both to families and to the county budget—to providing coordinated housing and social supports to mothers with children in the foster care system, to more rapidly stabilize and reunite them.

These are examples of work that take seriously the needs of agency and nonprofit practitioners and their clients, and that use data to interrogate problems, explore new solutions, and put authority in the hands of decision makers who are closer to the point of service. They are initiatives that use data as a flashlight and not as a hammer. Much more of this is possible.

WHAT IT WILL TAKE

This chapter has been critical of a data economy that revolves around the planning decisions of large government systems and private funders rather than one that reacts to market forces reflecting the needs of nonprofit practitioners and communities. That result was not inevitable, however, and creating a different economy for data and evidence will require infrastructure and new capacity within communities. All of us have a role to play in laying that foundation—funders included, and especially.

Infrastructure

The kinds of data projects that can strengthen the decision making of communities and practitioners share common elements. They are developed by data intermediary organizations that center on practitioner needs and have built trusting relationships with their agency and nonprofit partners, often with formal governance arrangements that include business and legal agreements. When these projects use external technical expertise, that expertise is martialed through organizations like Code for America and the U.S. Digital Service with deep knowledge of the pain points, incentives, and limitations of their public and nonprofit partners. These organizations are vehicles of a more responsive and innovative economy for data products and tools. The networks that connect them are the roads by which new tools, policy analyses, and initiatives propagate.

These intermediary organizations and networks are chronically undersupported parts of the sector. Several of the projects described here were created by social entrepreneurs operating outside the bounds of their professional responsibility, sometimes against the incentives of their funding. To create and sustain this kind of adaptive data capacity requires more than project-oriented grants and capital dollars for modernizing technology. It takes patient support for the crucial "soft" work necessary to understand the priorities of agency and community leaders, negotiate terms of access to their information, and prove that this kind of data infrastructure can solve real problems. Once established, these data intermediaries—whether university-based policy labs, state offices like Kentucky Stats, or local nonprofits like the New Orleans Data Center—tend to persist and expand into new domains where they can rapidly and cost-effectively build projects. It is in the enlightened self-interest of government and philanthropic funders to help develop these practitioner- and community-oriented organizations within the counties and states where we work, and to start *before* there is an urgent need for evidence.

Sharing and Building Power

Fans of data and evidence should recognize that our goals of using data for good depend on earning social license, described in Amy O'Hara's chapter as something that exists when the public trusts that data will be used responsibly and for societal benefit. A caution flag has been flying for several years that parents and communities—particularly communities of

color—are dissatisfied with their place in the development of this data economy. For parents and educators, it was the 2013 public launch of a centralized data sharing platform, inBloom, that catalyzed three years of protest and hundreds of pieces of privacy legislation[12] aimed at curbing the collection and use of information on students. Recent and more pointed arguments from civil rights organizations like Data for Black Lives and the Leadership Conference for Civil and Human Rights have focused on the need to renegotiate limits on the use of data and technology tools,[13] and to foreground issues of race and racism. The data of communities of color is often collected and used, but communities are rarely included in framing the field's research priorities.

Organizations like Actionable Intelligence for Social Policy have responded by collaborating with some of these critics (colloquially, "frenemies" of data) to develop a roadmap for centering race in data use, integration, and governance.[14] And a few local and state data intermediaries are making earnest attempts to give parents, nonprofits, and communities of color real power at those governance tables where decisions are made about what kind of evidence is important to build. But these efforts are nascent and more difficult to manage the further organizations get from neighborhoods and schools. The National Secure Data Service envisioned by the 2018 Evidence Act, for example, has tremendous potential to contribute to racial equity analyses of U.S. programs and policies at all levels, but questions remain whether or not Americans will tolerate the federal government directing such a powerful tool.

The Annie E. Casey Foundation argues that the social sector should lean into this debate[15] about data infrastructure and data innovation, and that, rather than trying to "abolish big data,"[16] funders and civil rights advocates alike should create and enforce standards for the fair and good use of data. The good uses are potentially vast: from AI tools that already triage crisis calls to the Trevor Project[17] and target lead remediation efforts[18] to new talent screening models[19] that promote more diverse technology workforces by rewarding aptitude rather than educational pedigree.

These "fair and good" uses are not inevitable, however. Civil rights critics of these technologies—and of data's use in the social sector—are right to point out that, in some ways, the more *likely* outcome is the opposite, that data science will be deployed in ways that systematically disadvantage poor and minority communities through greater surveillance and actuarial dis-

crimination. This, too, is already happening, as anyone who reads Upturn's weekly newsletter[20] knows.

The difference between these two competing visions parallels the problem this chapter began with, and its fault line is the willingness of our data economy's most powerful actors to cede some control over whose questions take priority and whose decisions new data tools interrogate. By giving nonprofits and affected communities a greater stake in the creation and ownership of this kind of evidence, we enable a much more dynamic and fair market for new ideas and solutions. This is the right moment for the social sector to recommit itself to uses of data that are not only useful to practitioners but also *empowering* to the communities that the Great Society, which inaugurated so much of this kind of policy analysis, was created to help.

NOTES

1. American Federation of Teachers press release, "Federal Suit Settlement: End of Value-Added Measures for Teacher Termination in Houston," October 10, 2017, https://www.aft.org/press-release/federal-suit-settlement-end -value-added-measures-teacher-termination-houston.

2. Putting Data to Work to Improve Child Well-Being: A Post-Convening Report. Casey Family Programs and the National Governors Association for Best Practices: Washington, DC, 2006.

3. See "Why Am I Always Being Researched," at the Chicago Beyond website, https://chicagobeyond.org/researchequity/, as well as the work of groups like Our Data Bodies, https://www.odbproject.org/.

4. Ibid.

5. Mark Friedman's seminal book *Trying Hard Is Not Good Enough: How to Produce Measurable Improvements for Customers and Communities* (Santa Fe, NM: Parse Publishing, 2019) helped popularize results-based accountability, .

6. Code for America, "Reshaping the Safety Net: The Integrated Benefits Initiative State Cohort," Medium, August 28, 2018, https://medium.com/code -for-america/reshaping-the-safety-net-the-integrated-benefits-initiative -state-cohort-32ea621e40d5.

7. See Child and Youth Integrated Homelessness Data Report: Part 2 page at Charlotte Urban Institute website, https://ui.charlotte.edu/story/child-and -youth-integrated-homelessness-data-report-part-2.

8. See "Critical Data Ensures Child Care for Essential Workers," Colorado Evaluation and Action Lab blog, June 24, 2020, https://coloradolab.org/blog -2020-06-24-child-care-for-essential-workers/.

9. See AISP Network: Los Angeles County Enterprise Linkages Project 2.0 page at AISP website, www.aisp.upenn.edu/network-site/la-county-2/.

10. See "Identification and Mapping System for Vulnerable Populations," at NYC Center for Innovation through Data Intelligence website, https://mcusercontent.com/07732221ebbe01a8b8186e8ef/files/b1d68383-43bc-49b8-a7f8-e4b8d5724f08/Vulnerable_Populations_Description_and_Process.pdf.

11. See Integrated Data Is Key to "Pay for Success" page at the Annie E. Casey Foundation website, www.aecf.org/resources/integrated-data-is-key-to-pay-for-success/.

12. "The Legacy of InBloom, Working Paper, Data Society, February 2, 2017, https://datasociety.net/pubs/ecl/InBloom_feb_2017.pdf.

13. See Principles page on Civil Rights Privacy and Technology Table website, www.civilrightstable.org/principles/.

14. "A Toolkit for Centering Racial Equity Throughout Data Integration," Actionable Intelligence for Social Policy, https://aisp.upenn.edu/centering-equity/.

15. The Annie E. Casey Foundation, "Four Principles to Make Advanced Data Analytics Work for Children and Families," The Annie E. Casey Foundation, October 27, 2020, www.aecf.org/resources/four-principles-to-make-advanced-data-analytics-work-for-children-and-famil/.

16. Yeshi, "Abolish Big Data," Medium, July 8, 2019, https://medium.com/@YESHICAN/abolish-big-data-ad0871579a41.

17. Josh Weaver, "The Trevor Project Launches New AI Tool to Support Crisis Counselor Training," The Trevor Project, March 24, 2021, www.thetrevorproject.org/blog/the-trevor-project-launches-new-ai-tool-to-support-crisis-counselor-training/.

18. "Machine Learning Models Identify Kids at Risk of Lead Poisoning," Science News, September 21, 2020, www.sciencedaily.com/releases/2020/09/200916113523.htm.

19. Michele Cunningham, "Time to Use Data for Social Good," Catalyte, www.catalyte.io/data-for-social-good/.

20. See About page on the Upturn website, www.upturn.org/about/.