Innovating for Inclusion
Advancing U.S. Community Development Financial Institutions (CDFIs) with Data Science
DataKind

With support from
Center for Inclusive Growth
DataKind.org
# Table of Contents

1. Executive Summary
4. Introduction
6. Community Development Financial Institutions and Financial Inclusion
8. Stakeholder Review
10. Key Challenges Faced by CDFIs
15. Where Data Science Fits
24. Final Thoughts
26. References
30. Appendix
Executive Summary

The U.S., despite its status as an economic power, is experiencing entrenched economic inequality, a part of which is due to a lack of equal access and use of financial resources. An estimated 5.9 million households have no formal checking or savings account and are considered unbanked.[1] Another 18.7 million are considered underbanked, or households that have bank accounts but use nonbank services like money orders, rent-to-own services, or payday loans.[2] There are numerous complex reasons why people do not have access to or use formal financial services, ranging from systemic social injustices to being considered too high risk for traditional banks, to a lack of trust in the banking industry. Regardless of the cause, the effects are damaging in the short and long term. Financial exclusion not only reduces the ability of households and small businesses to weather shocks to their income, but it also reduces their capacity to generate wealth for their families and communities and escape poverty.[3] While many interested parties are looking to remedy these gaps, Community Development Financial Institutions (CDFIs) are uniquely situated to address the financial needs of underinvested markets. CDFIs operate as financial institutions with a mission to serve those most overlooked or underserved by traditional institutions. However, despite recent increases in investment and attention, CDFIs are experiencing a challenging tension between a growing need for their services and their capacity to meet that need, reducing their ability to be effective agents of change for financial inclusion.

This landscape reviews the history and role of CDFIs in financial inclusion, key challenges the industry faces, stakeholders that work in this space, and data science opportunities to leverage. This review finds three core issues affecting the ability of CDFIs to grow as businesses and as agents of financial inclusion: (1) their target markets constrain them, (2) their resource-intensive approach and small size compared to traditional institutions reduces the capacity for scaling, and (3) they lack competitive use of technology and data to aid in business operations and performance measurement.

[1] FDIC (2023, July 24)
[2] Ibid.
First, CDFIs, by their definition, target underinvested customers who are perceived by conventional banks as riskier and less profitable for a financial portfolio.[4] To operate in these underserved markets, CDFIs rely on public-private partnerships to support their business operations, which often come with specific investment requirements.[5] These requirements can limit CDFIs’ reach and their ability to grow. Related, the markets in which they operate do not have the volume and profitability of traditional financial institutions, making CDFIs doubly limited in their growth. Indeed, CDFIs only hold 1% of the $18.3 trillion in assets of traditional banks and credit unions.[6]

The second core issue is that CDFIs have a resource-intensive customer service and management approach vital to serving the most under-resourced customers, dubbed relationship lending. Relationship lending requires collecting substantial qualitative data about customers to better understand their financial risk, which is less conducive to digitalization than traditional data, like credit scores and tax returns.[7] This approach gives CDFIs an advantage with their unique customer base, but also limits their ability to increase their financial product service volume, which hampers their profitability, and, consequently, the resources available for investing in their own operations and growth.

Third, and weaving throughout these issues, is the fragmentation of data collection and analysis efforts and a lack of technology use, making both operational efficiency challenging and performance difficult to measure. Because CDFIs operate in such specialized markets, with customized and localized portfolios of products to meet a given community’s needs, the kinds of data they collect are highly variable. There is also a lack of investment in technological integration that could facilitate more consistent and in-depth data collection, as profits are used to support the delivery of services and products in their target markets versus investing in business capacity building.[8] Such a dynamic creates a cycle of needing more robust data systems and analysis, but lacking the resources to build such infrastructures internally, thus hindering CDFIs’ ability to increase their growth and impact.

---

[5] For instance, federally certified CDFIs can receive funding from the CDFI Fund but must follow regulations on eligible markets it can target.
Such challenges create an opportunity for data to enter the arena and help maximize the positive impact of these businesses. A key lesson in this space is that data systems support is needed across all areas of CDFI business operations to better coordinate efficiencies and actionable insights. In addition, this landscape finds a number of ways data science can help CDFIs find markets, reach customers, measure their performance, and leverage the same kinds of tools that private-sector banks and traditional financial institutions have used to serve their customers.

In particular, using new ways to define and measure economic distress and identify areas of need via satellite imagery and Geographic Information Systems (GIS) analysis can open up new areas to target CDFI services. Analyzing different metrics that capture or approximate financial exclusion, including persistent poverty at the census tract versus county level, or creating indices that combine different measures can also expand the market picture and identify gaps in financial access. To reach customers, data science can contribute the tools of customer segmentation, tech-enabled advertising, prediction of the next likely product, and loan approval automation, all of which can increase the scale and competitiveness of CDFIs while also allowing them the nuanced insights to reach customers most in need of financial services. Finally, CDFIs can use technology to automate their product management processes, increasing their capacity to explore different ways to understand their social performance. Building a tech infrastructure that prioritizes better data utilization will help bridge the gap between service demand and CDFIs’ ability to deliver them.
Persistent economic inequality is a prominent issue in the U.S., with upper-income households seeing their wealth grow while middle and lower-income families are seeing declines,[9] in addition to a poverty rate that has remained stubbornly consistent over the last 40 years.[10] These trends are exacerbated when the data is disaggregated, with families of color experiencing more significant economic gaps than white households.[11] This inequality is troubling, as it can reduce economic and social mobility and opportunities, perpetuating poverty and geographic and economic segregation.[12] While driven and produced by a number of complex social, economic, and political factors, including systemic racial and gender injustices,[13] one of the proximate barriers to economic equality is the lack of access to or use of financial resources, such as banking accounts and loans.[14] In 2021, approximately 5.9 million households lacked checking or savings accounts with any formal financial institution.[15] When households do not utilize financial services, they are vulnerable to everyday economic shocks. For instance, without a credit card, paying for an expensive car repair can be impossible, impacting the ability to commute to work, and thereby causing downstream effects. Over the long term, these issues and the lack of financial services can affect a household’s ability to own homes, create generational wealth, or escape poverty.[16] There are many reasons for these disparities – like systemic lack of access or lack of trust in banking institutions[17] – and the movement to address this disparity is more broadly called financial inclusion, defined by the World Bank as when “individuals and businesses have access to useful and affordable financial products and services that meet their needs – transactions, payments, savings, credit, and insurance – delivered in a responsible and sustainable way.”[18]

[10] U.S. Census Bureau (2023, January)
[13] Ibid.
[17] FDIC (2023, July 24)
[18] The World Bank (2022, March 29)
Landscape Purpose: Exploring Community Finance with a Data Science Lens

CDFIs are unique and take many forms, but they are all mission-driven businesses that simultaneously work to meet their bottom line and serve customers typically overlooked or underserved by traditional financial institutions. These underserved customers often live in low-income areas, have low or no credit, or are otherwise deemed too costly or high risk.\[19\] It is worth noting that CDFIs are but one of the many stakeholders targeting financial inclusion. Minority Depository Institutions (MDIs), Microfinance Institutions (MFIs), private philanthropic foundations, and others seek to address the needs of marginalized communities as well. This landscape focuses on CDFIs specifically due to increasing interest in and demand for their services. Signaling the growing demand for CDFIs to address financial service gaps for those that need them most, the U.S. Department of the Treasury announced an $8 billion investment in CDFIs via the Emergency Capital Investment Program in the fall of 2022.\[20\] Despite this, CDFIs experience challenges that hinder both business growth and the achievement of financial inclusion and social performance goals. This landscape provides an overview of the role of CDFIs in financial inclusion, a summary of the data-based hurdles CDFI face, and a review of how data science can address those hurdles to advance the impact of CDFIs.

The Mastercard Center for Inclusive Growth (MCIG) supported this landscape, and as such, it aims to provide useful insights related to the type of work specifically supported by MCIG, as well as broader applications to the CDFI industry.

The history of financial inclusion in the U.S. predates the term and recognizing CDFIs as specific actors within it. The work of finance with a social mission has a long track record. In the 1880s, minority-owned banks emerged to focus on serving low-income areas.[21] Some credit unions in the 1930s and 1940s served African Americans in the rural south, recognizing the legacy of racial injustice that excluded them from traditional banks.[22] In the 1960s and 1970s, Community Development Corporations (CDCs) gained popularity, aiming to revitalize economically distressed or low-income areas through small business support. By the 1980s, non-profit organizations started loan funds to support affordable housing projects.

Formal recognition of CDFIs came about in the 1990s with the renewal of the Community Reinvestment Act (CRA) of 1977 during the Clinton Administration. The CRA mandated banks to serve their entire community without discrimination.[23] During this time, the CDFI Fund was created to serve as the national certifying body and provides funding to support CDFIs.[24] Though not mandatory, some CDFIs pursue certification through the CDFI Fund to qualify for certain funds and technical assistance. Certified CDFIs must follow the rules outlined in the CRA.[25] Today, CDFIs primarily service small businesses, housing projects, or individual consumer goods.[26] They receive support from public-private partnerships, which allows them to operate in lower liquidity communities.[27] While not a total count of CDFIs, in 2021, the CDFI Fund reported 1,271 certified CDFIs, with the plurality being loan funds (45%), followed by credit unions (30%).[28]

CDFIs fall into two main categories: regulated and unregulated. Regardless of the form, community representation is a common feature to remain connected and informed about community needs. Regulated CDFIs encompass community development banks and community development credit unions (CDCUs). Community development banks offer lending and investment services that target redevelopment in economically distressed areas. They are typically for-profit companies regulated by the Federal Deposit Insurance Corporation (FDIC),

[22] Ibid.
[23] Ibid.
[25] See rules for CDFI certification here
[27] Getter (2022)
[28] Ibid.
Federal Reserve, or state banking agencies.[29] In addition, approximately 50% of community development banks are also categorized as MDIs, which are defined as institutions owned or directed primarily by members of certain minority groups.[30] On the other hand, CDCUs focus on ownership, savings accounts, and affordable credit service. They are non-profit businesses owned by their members and regulated by the National Credit Union Administration, an independent or state agency, or both.[31]

Unregulated CDFIs include community development loan funds (CDLFs) and community development venture capital funds (CDVCs). CDLFs focus on providing financing and development loans for businesses, organizations, or individuals in low-income communities. They usually are non-profit organizations.[32] CDVCs provide equity and debt-with-equity services for small and medium businesses. They can be either for-profit or non-profit.[33]

The diversity of business types, customer profiles, and addressed needs makes CDFIs relatively more nimble than other financial institutions. They can provide access to funds for loan repayment to high-risk credit consumers or provide digital banking to unbanked communities. Yet this agility brings unique challenges, perhaps the most significant being the ability to collect and report data in a systematized and standardized way that can be utilized for a variety of analytical purposes. Underlying hardware and software challenges and the relative lack of digital maturity among CDFIs exacerbate this issue.

While certified CDFIs that receive funds from the CDFI Fund have particular rules that govern who they can target and how they must report on their business,[34] these regulations do not apply to CDFIs who do not receive direct funding from the CDFI Fund. Further, different lending sectors, such as business or housing lending, have distinct data needs, leading to a wide variety of data collection, reporting, and analysis across all CDFIs. As a result, the CDFI data landscape is abundant but fragmented, making data science a crucial tool to advance the CDFI industry into a new era.

[29] Opportunity Finance Network (2023)
[30] Mosley (2019); Note: MDIs are not inherently CDFIs and vice versa. MDIs have their own unique history (see review here)
[31] Opportunity Finance Network (2023)
[32] Ibid.
[33] Ibid.
[34] See all rules here. Note: certified CDFIs must complete an Annual Certification and Data Collection Report (ACR) using the CDFI Fund’s Awards Management Information System (AMIS)
Section III - Stakeholder Review

There are many actors involved in the CDFI industry. Understanding these key players and their power dynamics is crucial as they set the stage for operationalizing different solutions. This section provides a concise overview of the key stakeholders, but a more comprehensive table is available in the appendix.

Knowledge Producers

To understand the large trends in the process and impacts of CDFIs, most knowledge producers are positioned within think tanks, like the Urban Institute, or university research centers, like the Center for Impact Finance at the Carsey School of Public Policy of the University of New Hampshire. Government entities such as the CDFI Fund are also a vital source of information on CDFIs, providing regular reports and research on the status and impact indicators of certified CDFIs. The CDFI Fund and other government agencies (e.g., FDIC, the Federal Reserve, etc.) are also the sources of public datasets that can be used for CDFI analysis purposes (e.g., unbanked household surveys, CDFI business surveys, census data, etc.).

Agenda Setters

Since CDFIs as they currently exist became formalized through renewed attention to the Community Reinvestment Act, the CDFI Fund (and the U.S. Congress as its appropriator) is arguably the largest stakeholder in setting the policies and agenda of how CDFIs operate. It holds the keys to certification, which opens the door to much-needed funding and thus also sets the agenda for what kind of data CDFIs collect for their reporting requirements. The FDIC is another important player here as they are the largest insurer of CDFI banks. Beyond government actors, a robust set of private sector organizations also work on lobbying for CDFI support and serve as industry voices via network coalitions (e.g., Opportunity Finance Network) or as community development intermediaries (e.g., Local Initiatives Support Corporation).

Funders

While the CDFI Fund is a primary and prominent funder for CDFIs, these institutions also have the opportunity to seek funding from various other government programs. For instance, the State Small Business Credit Initiative of the U.S. Department of the Treasury, various programs within the U.S. Department of Agriculture, or the
Community Development Block Grant of the U.S. Department of Housing and Urban Development offer potential funding sources.[35]

In addition to government sources, a vast network of private foundations and banks serve as key financial supporters. Bank of America is the largest private investor, with $2 billion in CDFI investments.[36] Private investment is a critical factor for the growth of many CDFIs, as they are often required to apply for private funding to offset the risks of their portfolios to receive public funding.[37] This blend of public and private investment acts as a key catalyst in the growth of CDFIs.

**Data Solutions and Technical Support for CDFIs**

The private sector plays a vital role in the field, particularly in addressing the challenges related to data collection and utilization within CDFIs. Corporations, think tanks, and non-profit organizations fill the gap for CDFIs, offering services like data systems engineering, management, analysis, and impact assessment. For example, some analysis firms help CDFIs understand their performance and impact via visual dashboards, automating the analysis process to make it more accessible for resource-strapped CDFIs.[38] The private sector also provides technical solutions to investors looking to support CDFIs via streamlining impact investment. For example, CNote is a financial investor intermediary, allowing investors to contribute funds to a diverse portfolio of CDFIs.[39] Similarly, Aeris is one of the most well-known impact assessment raters for CDFIs, providing investors with a comprehensive understanding of their impact.[40]

---

[36] Bank of America (n.d.)
[37] Getter (2022)
[38] High Impact Analysis
[39] CNote
[40] Aeris
Section IV - Key Challenges Faced by CDFIs

Over the last 30 years, CDFIs have proven successful in numerous ways, including direct customer benefits, like spurring small business growth and addressing financial gaps within communities.[41] As lenders, CDFIs have exhibited comparable levels of success to mainstream financial institutions, with no more risk of financial failure.[42] However, achieving their long-term financial inclusion and business growth goals presents several challenges for CDFIs, produced by core issues with how CDFIs function and the limitations of their industry: (1) their target markets constrain them, (2) their resource-intensive approach is necessary but limits their growth, and (3) a lack of digitalization has hindered their use of data-enabled operations and insights.[43]

CDFIs operate within inherently challenging markets, requiring reliance on supplemental government and private funding, and resulting in portfolios that are riskier and more costly to maintain. These portfolios also require high-touch, resource-intensive management, further limiting the available resources to expand their businesses. As such, there is a delicate balance between scaling impact and maintaining in-depth customer service relationships, a hallmark of their approach to financial inclusion. These foundational issues mean CDFIs can be uniquely constrained in their growth compared to traditional financial institutions, which puts limits on the kinds of tools, strategies, and capacities they can use. This has created tension; the need for nimble financial institutions like CDFIs to meet the financing needs of the most underserved is growing, but their ability to expand to meet that need is hampered. In an effort to align solutions with business operations to aid in this tension, the challenges in this section are presented along the pathway of providing a financial product: identifying the market, reaching customers, and measuring performance.

Identifying the Market

Since CDFIs are financial institutions dedicated to serving underinvested communities, they often have very specific rules regarding their target audience for those services. In particular, if they are certified, they must adhere to the eligibility rules for Investment Areas or Targeted Populations as defined by the CRA.

[41] Swack, Hangen, & Northrup (2015); McCall & Hoyman (2023)
[42] CDFI Fund (2015, February 15)
[43] Balboni & Travers (2017); Getter (2022), Theodos, González-Hermoso, & Myczkowska (2022); Northrup, Hangen, & Swack (2016)
Those rules stipulate that an Investment Area meet one of the following criteria:[44]
(1) Have 20% or more of the population living in poverty,
(2) Possess a median family income at or below 80% of the metropolitan area or state median family income,
(3) Exhibit an unemployment rate at least 1.5 times higher than the national average, or
(4) Have a net migration loss during the five years preceding the census of at least five percent.

For a Target Population, the rules stipulate that a CDFI can target individuals or an identifiable group of individuals that are either:[45]
(1) Low-income,[46] or
(2) Lack adequate access to finance products or services in the target market.

While these rules ensure certified CDFIs operate in a community with genuine financial need, they can also impose limitations on their market. Consequently, this restricts CDFIs’ ability to meet the financing needs of underinvested communities when the definition of those communities is limited to specific criteria.[47]

Even more, it is evident that there are geographic disparities in where CDFIs operate, even though it is unclear what factors are producing them.[48] In assessing the distribution of CDFIs and their investments, one study found that more populated counties receive more funding per person than less populated ones, even when adjusting for levels of financial need, indicating a rural gap in CDFI services.[49] Further, the amount of funding is not necessarily tied to levels of economic distress. The same study found that counties with the worst poverty received less per-person funding than counties with better poverty rates.[50] These data points illustrate that CDFIs are not equally or equitably distributed, creating opportunities for increasing their markets. Indeed, since CDFIs are targeting higher risk areas, which are both more costly and less profitable to serve, expanding their markets is an important component for future growth and requires defining and identifying those markets in innovative ways.

---

[44] See all Investment Areas rules here
[45] Ibid.
[46] Low income is defined as “(1) For Metropolitan Areas, 80 percent of the area median family income; and (2) For non-Metropolitan Areas, the greater of: (i) 80 percent of the area median family income; or (ii) 80 percent of the statewide non-Metropolitan Area median family income” (source)
[48] Balboni & Travers (2017); Theodos & Hangen (2019)
It is important to note that CDFIs that do not receive funding from the CDFI Fund do not have the same market targeting requirements; they are also not required to report on their operations. As a result, it is difficult to understand the specific market targeting issues they encounter. However, given that the CDFI industry is generally struggling to meet the market need for financial services of underserved customers,[51] it is safe to say that all types of CDFIs could nonetheless benefit from innovative insights on identifying which markets most need their services.

**Reaching Customers**

Once a target market is identified, the next step is to connect a CDFI’s services to customers within that market. This step has posed a significant challenge for CDFIs, partly because of their relatively smaller size and lack of internal capacity. This limits their ability to leverage digital tools for advertising and loan processing. Robust and data-informed marketing apparatuses are often out of reach for CDFIs, leaving them competing with traditional banks, FinTechs, or even predatory lenders. [52] Indeed, online lending has also outpaced CDFI growth, introducing even more competition, particularly with business lending.[53] CDFIs frequently don’t have the staff or technological capacity to provide online or mobile applications for products or to process products quickly (e.g., loan approvals), which means they often miss out on customers that need resources immediately, losing them competitively to online service providers.[54] On the other hand, in rural or low-income communities, households may lack equitable access to reliable internet connections, reducing their ability to participate in digital financial services.[55] This is a gap that CDFIs can fill. The convergence of these challenges – encompassing advertising, industry competition, and digitalization – complicates CDFIs' efforts to secure product adoption, diminishing their business foundation and operational sustainability.

---

[51] Northrup, Hangen, & Swack (2016)
[52] CDFI Friendly America, personal communication (2023, June 1)
[53] Northrup, Hangen, & Swack (2019)
[54] Theodos, González-Hermoso, & Myczkowska (2022); Northrup, Hangen, & Swack (2019)
[55] Friedline, Naraharisetti, & Weaver (2020)
Measuring Performance

After customers connect with the products that a CDFI offers, the next challenge is monitoring and measuring performance. The unique issue for CDFIs is measuring progress toward their social performance indicators.[56] To be sure, there is a wealth of financial output data associated with the business function of CDFIs, such as number of loans closed, returns-on investment, loan performance, etc.[57] However, CDFIs often don’t know who is benefiting from their activities.[58] Although CDFIs can track who signs up and receives their products and services, the picture becomes murkier when trying to understand the long-term impact of their efforts. This challenge is rooted in several core issues along the data continuum for program management and measurement. These issues involve defining impact, collecting data, and analysis, all of which must be addressed to strengthen CDFIs’ understanding of their social performance and expand their impact on financial inclusion.

Defining Impact

First, the diversity of CDFI business operations, lending sectors, customer bases, and services utilized, means there is an equally diverse set of goals and definitions of impact for those businesses.[59] As each CDFI defines its unique goals, it shapes the selection of appropriate measures to capture those impacts. Some CDFIs focus on collecting the minimum data required for reporting to funders,[60] while others seek a broader understanding of their impact. It is also difficult to pinpoint metrics that can capture the longer-term, second-order social impacts (i.e., upward mobility, reduced wealth gaps, improved financial well-being) that CDFIs hope to achieve, limiting their ability to understand and expand financial inclusion in their markets.[61] This creates tension in data needs within the industry. On the one hand, there is a need for standardized, transferable industry-wide impact metrics. On the other hand, specific measures are required to capture the specific needs of each CDFI. The result is contexts that create different data environments, influencing the type of analysis and interpretation that is possible. Navigating these diverse contexts and data environments poses an ongoing challenge for the CDFI industry, but finding a balance between standardized metrics and tailored measures can lead to more comprehensive insights into the industry's collective impact.

[56] McCall & Hoyman (2023)
[57] For example, see standard reports from the CDFI Fund
[58] Chapel (2021)
[60] Mudd (2013); Theodos & Seidman (2016, November 14)
[61] McCall & Hoyman (2023)
Data Collection

The next challenge within this space is efficient data collection, regardless of the articulated goals and metrics for a CDFI.\[62]\) Due to their smaller size and focus, CDFIs have fewer financial and technical resources to leverage for creating robust data collection and analysis frameworks.\[63]\) They heavily rely on public-private partnerships and outside funding (i.e., CDFI Fund, private banks, foundations, etc.) to support their missions, which primarily goes to supporting financial product provision in their target markets.\[64]\) They also utilize relationship-focused lending, a time-consuming process that reduces resources, including staff capacity.\[65]\) This approach also involves collecting qualitative data that provides valuable context for understanding a customer’s financial risk which allows CDFIs to offer them the right product and understand their own risk in providing it.\[66]\) However, this data is less conducive to digitization than more traditional data like credit scores or tax returns, meaning much of the data entry for CDFIs is manual.\[67]\) In addition, collecting longitudinal data is difficult, both in terms of CDFI staffing capacity and the logistical barrier of getting consistent responses from customers. Ultimately, while both supplemental funding and human-centered approaches enable CDFIs to serve overlooked customers, it can hinder their ability to handle higher volume business, quick product approvals, or data systems management, all of which limit their growth and impact potential.\[68]\)

Analysis

The first two issues create a fragmented data analysis landscape for CDFIs. The complexity and diversity in defining metrics of impact and collecting data efficiently and effectively make it challenging to conduct larger-scale impact assessments across the entire industry, as opposed to evaluating the impact of each CDFI’s own operations.\[69]\) The fragmentation of technology tools for business management exacerbates these issues, with specialized tools for client management, loan underwriting, and reporting that often do not integrate with one another,\[70]\) making analysis along the financial product lifecycle cumbersome at best and prohibitive at worst. Addressing these challenges will empower CDFIs to navigate their operations better, improve service, and achieve their long-term goals of financial inclusion and community development.

\[62]\) Mudd (2013)
\[63]\) Neal, Zinn, & Zhu (2023); Theodos, González-Hermoso, & Myczkowska (2022)
\[64]\) Mosley (2019)
\[65]\) Getter (2022)
\[66]\) Ibid.
\[67]\) Ibid.
\[68]\) Theodos, González-Hermoso, & Myczkowska (2022); Mudd (2013)
\[70]\) Theodos, González-Hermoso, & Myczkowska (2022)
Section V - Where Data Science Fits

Data science has the potential to be a powerful catalyst for growth and development in the CDFI industry, as many of its challenges revolve around data. The data needs of CDFIs vary based on the data maturity of the organization, with some requiring assistance with data collection, retention, customer tracking, and business operations. Data professionals can positively impact this space by providing essential data engineering and data management services.

Numerous data science approaches used in the financial sector can also be leveraged specifically for data-mature CDFIs to support their goal of financial inclusion. For instance, predictive modeling using decision tree based machine learning algorithms, such as random forests and gradient boosting,[71] has been used by financial institutions to evaluate creditworthiness and make informed loan approval decisions. By considering factors like credit score, income and employment history, these models predict the likelihood of loan repayment, which benefits lenders, and helps prevent individuals from taking on unaffordable debt, which helps borrowers.[72] These models can also be adjusted to ensure that biases based on race or gender are accounted for and not perpetuated within the loan decision process, promoting fair lending practices.[73]

The rest of this section aligns these kinds of data science opportunities with the challenges summarized in Section IV. It is important to note these suggested strategies may require significant internal capacity and resources, which many CDFIs may lack. Therefore, engaging with technology partners or seeking grants to increase internal capacity[74] can allow CDFIs to embrace data science and utilize it effectively in their operations. Further, the proposed strategies cannot fully address the underlying challenges that the CDFI industry faces due to the inherent constraints of their target markets, but data science can lend efficiencies and insights that can help mitigate them. By engaging with data science, CDFIs can enhance their ability to serve underserved communities, foster financial inclusion, and achieve their social missions more effectively.

[71] See definitions provided in Lok, L. (2022, January 5)
[73] Agarwal, Muckley, & Neelakantan (2023)
[74] E.g., Technical Assistance grants from the CDFI Fund
Identifying the Market

In pursuing underinvested markets, data science can contribute by offering innovative and nuanced approaches to define, measure, and map communities effectively. There are data science approaches that can help assess financial inclusion and pinpoint areas where interventions are needed to ensure widespread access to affordable and appropriate financial services. A few of these include defining and measuring financial needs in different ways, creating customized indices to understand the financial health of a community, and using satellite imagery and GIS mapping to identify gaps in financial services.

Definitions of Target Areas

While the CRA sets specific definitions of eligible communities for certified CDFIs to target,[75] there are alternate ways to capture economic distress or financial needs. One of those is the concept of persistent poverty, defined by the U.S. Department of Agriculture as a population with 20% or more living in poverty over 30 or more years.[76] However, this measure has traditionally been applied at the county level, which is a rough aggregate when considering the diversity of a given community.

To address this limitation, the Opportunity Finance Network (OFN), a national network of CDFIs that offers capital, technical assistance, and networking opportunities to its member organizations, analyzed persistent poverty at the census tract-level. The census tract method identified a much higher number of people living in persistent poverty compared to the county method, indicating gaps in need that could be more precisely targeted.[77] Using these types of alternative analyses and definitions can help CDFIs understand financial need in new ways and expand their markets.

Using Indices to Find Gaps

Related to utilizing different measures of financial need is developing custom indices that incorporate various metrics, such as income levels, access to credit, and educational attainment. These indices are compound measures that aggregate multiple indicators,[78] enabling policymakers to gauge progress toward greater financial inclusion and identify areas requiring policy interventions.

[75] See CRA Rules here
[76] Loethen & Fabiani (2017)
[77] Ibid.
For example, an index might be created to measure access to credit for small businesses or farmers in rural areas, shedding light on gaps in financing options and developing appropriate strategies. In another example, using data from the Federal Reserve’s Survey of Consumer Finances, the Federal Reserve Bank calculates financial inclusion indices for different demographic groups, such as race, ethnicity, and income level. These indices are designed to easily show the level of financial inclusion across groups and allow policymakers to identify specific populations that may be experiencing financial exclusion and tailor their policies accordingly. Creating targeted versions of such indices can provide insights across the CDFI customer journey.

**Satellite Imagery and GIS Mapping**

Satellite images and GIS mapping can identify economic activity, infrastructure, population characteristics, and density in geographic areas. By analyzing these variables spatially, valuable insights can be gained regarding regions where traditional banking services may be inaccessible due to physical barriers or lack of infrastructure, shedding light on the dynamics of economic need.[79] By analyzing satellite images and using GIS to understand spatial patterns, financial institutions and policymakers can better understand where financial services are needed most and design targeted interventions to increase access to finance. For example, the U.S. Department of Agriculture (USDA) uses satellite imagery in its rural development program. In this program, the USDA provides loans, grants, and loan guarantees to support rural development, including housing, infrastructure, and economic development projects. To help identify areas of need and target resources effectively, the USDA uses satellite imagery and machine learning algorithms to analyze population data, economic indicators, and other factors to create a visualization to guide interventions and policy actions.[80]

Hyperlocal mapping and understanding of critical economic indicators like home ownership can also lead to tailored interventions for communities and a more directed policy advocacy. An example is the Foreclosure and Eviction Analysis Tool (FEAT) created by DataKind and New America Future of Land and Housing program. [81] FEAT is an open-source tool that enables local leaders to understand

---

[75] See CRA Rules [here](#)
[76] Loethen & Fabiani (2017)
[77] Ibid.
[79] E.g., see work in Hegerty (2020)
[80] E.g., see examples from the USDA [here](#)
[81] New America (n.d.)
housing loss trends by analyzing housing loss data, identifying patterns, and tracking intervention impact. FEAT helps organizations like CDFIs identify high foreclosure and eviction areas, allowing them to target services and support. By utilizing FEAT, CDFIs can enhance their ability to serve vulnerable populations and promote community revitalization in a more directly targeted way.

**Reaching Customers**

To effectively reach customers in areas of need, CDFIs often begin with data-informed marketing to inform prospective clients about their existence. However, many CDFIs lack the necessary marketing capability. Employing data science tools can help CDFIs better understand their customer base, target products to that base, and embark on advertising to them in more nuanced ways, thus increasing a CDFI’s business and its ability to meet the needs of those it aims to serve.

**Segmentation**

CDFIs can employ various customer segmentation methods based on different criteria. These methods include demographic segmentation (based on factors such as age, gender, income, and education), geographic segmentation (based on location), psychographic segmentation (based on values, attitudes, and beliefs), and behavioral segmentation (based on purchasing habits, usage rate, etc.). For CDFIs, a combination of demographic and geographic segmentation can be particularly useful, as it can help identify distinct communities or populations with shared financial needs and challenges. For example, a CDFI might segment its customer base by age group (e.g., youth, adults, seniors), income level (e.g., low-income, moderate-income, high-income), and geographic location (e.g., urban, rural, suburban). This approach allows the CDFI to tailor its products and services to meet the specific needs of each segment, increasing its effectiveness and impact.

One excellent example of successful customer segmentation is the Latino Economic Development Center (LEDC), a CDFI supporting low-income and minority communities in Washington, D.C., and Prince George’s County, Maryland. Recognizing Spanish-speaking entrepreneurs’ unique challenges, LEDC launched "Nuestro Futuro" (Our Future), a specialized program catering to this specific customer segment. Through tailored financial products, including business loans, credit lines, and financial workshops, LEDC addresses the needs of Spanish-speaking entrepreneurs. The organization strategically collaborates with local
organizations like the Latin American Youth Center and the National Council of La Raza to engage and build trust within this customer segment. As a result, LEDC experienced a substantial increase in lending activity among Spanish-speaking entrepreneurs, with 75% of clients being African American or Latino, and 60% women. LEDC's successful application of the customer segmentation approach has significantly impacted the communities it serves.[82]

**Ad Retargeting**

Ad retargeting is a powerful data-enabled tool used to connect organizations with potential customers and increase engagement. This approach involves displaying targeted advertisements to users who have previously visited the CDFI's website or engaged with their online content. By showing relevant ads to people who have already shown interest in the CDFI's mission, products, or services, retargeting campaigns can effectively re-engage these individuals and motivate them to act.

Retargeting ads can be delivered across a variety of digital channels, including social media platforms, news websites, and streaming services. To maximize effectiveness, it is important to tailor the messaging and creative elements of retargeting ads to the user's previous interactions with the CDFI's brand.

To illustrate how these techniques operate, Starbucks wanted to promote its Mobile Order & Pay feature, which allowed customers to place orders and pay through their smartphones before picking up their drinks in stores. The company wanted to encourage more customers to use the feature, especially during peak hours when lines can be long. Starbucks launched a campaign to promote its Mobile Order & Pay feature. They used email, social media, in-app, and geofencing retargeting to reach customers yet to try the feature. Emails highlighted benefits and offered a free drink or food item. Social media ads showed how easy it was to use, and in-app messages reminded customers of the benefits. Geofencing retargeting sent push notifications to customers near a Starbucks store. The campaign increased Mobile Order & Pay adoption, increasing their business efficiency and customer reach.[83]

**Next Loan or Next Product Prediction**

The idea of next loan/next service prediction revolves around analyzing a customer's financial history, credit behavior, and other relevant factors using

[82] See impact reports here
algorithms to identify patterns and trends that indicate their potential need for future financial products or services. For instance, if a customer has previously taken out loans for education expenses, the algorithm may flag them as likely to require another educational loan in the near future. Similarly, customers who consistently demonstrate responsible credit behavior may be identified as a good candidate for a personal loan or credit card.

By leveraging this technology, CDFIs can take a more proactive approach to serving their customers. Instead of waiting for customers to apply for loans or services, these organizations can anticipate their needs and offer targeted financial solutions at the right time. This not only enhances the overall customer experience and helps to promote greater financial inclusion, particularly among marginalized communities that historically faced barriers to accessing mainstream financial services.

**Loan Approvals and Bias Reduction**

Once a customer base is connected to a CDFI, the next challenge is ensuring that they can effectively utilize the offered product. Some businesses are offering AI-assisted lending services to broaden eligibility requirements. For instance, Upstart is championing an inclusive AI model that aims to rectify the traditionally narrow metrics used to approve loans, looking at other indicators that can increase the likelihood of approval.[84] This approach could be beneficial for CDFIs struggling with quickly providing approvals. However, utilizing technology like this can be a double-edged sword, creating an algorithmic bias that entrenches disparities.[85] For instance, the CDFI Reinvestment Fund created the Market Value Analysis algorithm-based mapping system for urban planning, but it can reinforce redlining disinvestment when attention is not paid to the diverse ways people value their communities.[86]

CDFIs should strive for careful and fair use of algorithmic approaches designed to reduce bias. In an example of bias reduction, the AI company ZEST AI has built a product called ZAML Fair which creates specifically designed models to reduce bias against protected groups. This product uses machine learning techniques to create models that maintain predictive accuracy while simultaneously improving model fairness.[87] Such tools also show promise within financial decision-making, as some actors start to explore ways to improve automated loan processing with

[84] Upstart
[86] Ibid.
[87] Ward-Foxton (2019, April 30)
explicit attention to ethical principles for AI to advance equity within this automation. Any data science solution should actively strive to address these pitfalls. Beyond the approval process, AI could help identify potential funders and write grants. Additionally, for initial customer service processing, it could identify needs and offer customized products.[89]

**Measuring Performance**

*Data Systems Efficiencies*

One of the core challenges to understanding the progress of CDFIs toward reaching their social performance indicators is the initial technological barrier of designing and utilizing efficient data collection and management systems, including hardware and software systems capacity and integration. While this challenge is not a data science analytics issue, quality analysis becomes a challenge without a solid data system foundation. Some private sector organizations see this same gap in capacity for CDFIs and have developed products to help provide more streamlined data systems for them.

Some CDFIs may not even know what their technology and data needs are. One company, ATX Advisory, guides CDFIs through a technology maturity model to identify their needs.[90] High Impact Analysis, LLC creates customizable Power BI dashboards for CDFIs, enabling a better understanding of their data and impact.[91] Fig36, an all-encompassing loan management tool provided by Fig Loans,[92] serves as a one-stop shop for nonprofits seeking to provide funds to low-income areas, allowing them to gather and analyze data about their loans efficiently.[93] LoanWell offers a similar service, providing end-to-end systems that emphasizes automation to help reduce technology burdens for CDFIs.[94] Inclusiv also recently expanded its Financial Inclusion and CDFI Data Analytics Platform, which targets the same goals illustrated throughout these examples – providing streamlined data monitoring, reporting, and analysis.[95]
Ultimately, there are many ways in which stakeholders seeking to support CDFIs can leverage technology to help them find efficiencies for their data management, which will ultimately enable CDFIs to move into a more mature data analysis space. Theodos, González-Hermoso, & Myczkowska (2022) provide a comprehensive summary of entry points for where data management-focused technological solutions can enter into small business CDFIs in particular, but these can be applied to other types of CDFIs as well (Box 1).

**Innovative Social Performance Measurement**

The data ecosystem within the CDFI space is predominantly geared toward collecting and monitoring inputs and outputs, making it challenging to identify and utilize more robust measures of social performance that truly capture progress toward financial inclusion.[96] The main hurdle lies in the tradeoff between capacity and cost-benefit for CDFIs. Conducting a comprehensive social impact assessment can be time-consuming and costly, and funders and investors may not require it.[97] As such, the motivation to find innovative ways to measure impact beyond financial performance can be lacking.[98]

[96] Porteous & Narain (2015); McCall & Hoyman (2023)  
[98] McCall & Hoyman (2023)
That said, there are still avenues for CDFIs to leverage data to gain insights into their impact. A comprehensive recommendation is to increase the adoption of evaluation frameworks, such as utilization-focused evaluation, which facilitates the systemic articulation and collection of measurable impact indicators.[99] Utilization-focused evaluation allows for the design of evaluations to be tailored to specific audiences and does not prescribe rigid methods, making it suitable for the diverse range of CDFIs and their stakeholders.[100] While it may be an uphill battle to get CDFIs to incorporate evaluation approaches into their businesses, it can be a fruitful way to ensure that impact measures are given due attention.

Furthermore, some organizations have piloted different ways to measure social impacts more generally, providing CDFIs with valuable options to assess their impact based on their goals. The Aspen Institute has a tool called MicroTracker, which offers microenterprises standardized metrics to assess their impact.[101] The Global Impact Investing Network’s IRIS+ presents a robust set of social and financial performance metrics, allowing users to create an impact framework and choose metrics that align with their goals.[102] The National Community Investment Fund has utilized this framework to develop its mission-oriented bank set of social performance standards.[103] This approach represents a way to have both standardized and specific metrics to meet industry assessment and CDFI performance needs.

[99] Ibid.
[100] Ibid.
[101] MicroTracker
[102] IRIS+ Metrics
[103] National Community Investment Fund Social Performance Metrics
Section VI - Final Thoughts

CDFIs play a critical role in advancing financial inclusion, catering to the unique needs of their underinvested clients in ways mainstream institutions overlook. CDFIs have a rich history of contributing to the social mission of economic equity, but they are at a tipping point as growing demand and digitization, coupled with strained access to supportive funds and organizational capacity, create unique challenges within the industry.

The current landscape presents challenges that demand careful consideration and innovative solutions. As this landscape has shown, these multifaceted challenges center around technological capacity, data systems, and data analysis, creating an exciting opportunity for data science to enter this space as a partner that CDFIs and investors should welcome. Yet, as the CDFI industry looks to leverage the tools and opportunities available within data science, it is also important to recognize that using data comes with its own risks and potential negative impacts. Exciting tools like machine learning, AI, and algorithmic automation must be tempered with the reality that data are not collected, created, or analyzed in isolation and often carry with them remnants of bias and injustice and risk being perpetuated.

Data is not just numbers and statistics; it represents the stories and experiences of real people. Keeping this in mind, CDFIs must approach data strategies with a strong commitment to equity and fairness. By prioritizing human-centric approaches to data collection and analysis, CDFIs can ensure that their financial services genuinely uplift the lives of those they serve.

Each CDFI should articulate a clear data strategy that aligns with its missions and values, considering the potential impacts of data-driven decisions on the communities they aim to empower (see the Data Ethics Framework of the Federal Data Strategy).[104] With responsible and ethical use of data science tools, CDFIs can strengthen their position as agents of positive change and contribute more effectively to the social mission of economic equity.

[104] See Data Ethics Framework here
Acknowledgements

This report would not have been possible without the efforts of the following:

**Concept and coordination** - John Harnisher, Ph.D. and Mallory Sheff, M.P.H.

**Research, writing, and report design** - Katie Crosley Beem, Ph.D.

**Financial support** - Mastercard Center for Inclusive Growth

For questions about this report, please contact:

DataKind®

Jemilah Magnusson
Communications Director
jemilah@datakind.org
www.datakind.org
References


Northrup, J., Hangen, E., & Swack, M. E. (2016). *CDFIs and Online Business Lending: A Review of Recent Progress, Challenges, and Opportunities*. University of New Hampshire: Center for Impact Finance at the Carsey School of Public Policy. [https://scholars.unh.edu/carsey/287/](https://scholars.unh.edu/carsey/287/)


## Appendix

### CDFI Industry Stakeholder Table*

<table>
<thead>
<tr>
<th>Organization</th>
<th>Primary Function</th>
<th>Relevant Program Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>AERIS</td>
<td>Technical Assistance/Knowledge</td>
<td>Provides CDFI financial and impact performance ratings</td>
</tr>
<tr>
<td>Aspen Institute</td>
<td>Technical Assistance/Knowledge</td>
<td>Created the <a href="#">MicroTracker</a> impact tool</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Part of the CDFI Data Project (from 2004 to 2005)</td>
</tr>
<tr>
<td>Association for Enterprise Opportunity (AEO)</td>
<td>Policy/Advocacy</td>
<td>Network to advance microbusiness and microfinance</td>
</tr>
<tr>
<td>Bank of America</td>
<td>Funder</td>
<td>Partners with CDFIs to fund projects that strengthen family businesses and communities in underserved areas</td>
</tr>
<tr>
<td>Camino Financial</td>
<td>Financial Service Provider</td>
<td>Neo-CDFI, provides small business loans, microloans, startup loans</td>
</tr>
<tr>
<td>CDFI Coalition</td>
<td>Policy/Advocacy</td>
<td>Focuses on lobbying for more funding for the CDFI Fund</td>
</tr>
<tr>
<td>CDFI Connect</td>
<td>Policy/Advocacy</td>
<td>Provides a forum for CDFI networking</td>
</tr>
<tr>
<td>CDFI Fund</td>
<td>Funder and Regulator</td>
<td>Certification of CDFIs and major/primary funder</td>
</tr>
<tr>
<td>Center for Community Investment</td>
<td>Technical Assistance/Knowledge</td>
<td>Provides a framework for communities to strengthen their investment systems and access resources</td>
</tr>
<tr>
<td>CDFI Friendly America</td>
<td>Network/Intermediary</td>
<td>Connects CDFIs to communities, helping bridge the gap between financial resources and communities that need them</td>
</tr>
</tbody>
</table>

*Note: This table is not all inclusive, but intended to be a summary of key players in the CDFI space as a resource
<table>
<thead>
<tr>
<th>Organization</th>
<th>Primary Function</th>
<th>Relevant Program Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cnote</td>
<td>Network/Intermediary</td>
<td>CDFI impact platform for impact investing</td>
</tr>
<tr>
<td>Community Reinvestment Fund</td>
<td>Financial Service Provider</td>
<td>Example of a nonprofit CDFI, supports organizations through the development of products and services that increase the flow of capital to underinvested communities</td>
</tr>
<tr>
<td>Economic Innovation Group</td>
<td>Technical Assistance/Knowledge</td>
<td>Provides insightful economic research, particularly around poverty and economic distress</td>
</tr>
<tr>
<td>Economic Opportunity Coalition</td>
<td>Policy/Advocacy</td>
<td>Coordinates with public, private, and social sector organizations to develop products that solve challenges in getting resources to where they are most needed</td>
</tr>
<tr>
<td>FDIC</td>
<td>Regulator</td>
<td>Insures/regulates a large portion of CDFIs</td>
</tr>
<tr>
<td>FinRegLab</td>
<td>Technical Assistance/Knowledge</td>
<td>Tests new technology and data to inform public policy</td>
</tr>
<tr>
<td>Global Impact Investment Network (GIIN)</td>
<td>Technical Assistance/Knowledge</td>
<td>Created the Investors Reporting and Investment Standard (IRIS) metrics, which are useful for impact investing standardization that CDFIs can benefit from using</td>
</tr>
<tr>
<td>Hope Credit Union Enterprise Corporation</td>
<td>Financial Service Provider and Technical Assistance/Knowledge</td>
<td>A CDFI that also partnered with local Historically Black Colleges and Universities (HBCUs) researching the financial needs of areas where both HBCUs and CDFIs operate, creating the HBCU-CDFI Economic Mobility Guide.</td>
</tr>
<tr>
<td>Organization</td>
<td>Primary Function</td>
<td>Relevant Program Summary</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Inclusiv</td>
<td>Network/Intermediary</td>
<td>CDFI Intermediary, working to understand and measure the impact of CDFIS and advocate for their use</td>
</tr>
<tr>
<td>Local Initiatives Support Corporation (LiSC)</td>
<td>Network/Intermediary and Policy/Advocacy</td>
<td>Community development intermediary, works with CDFIs and other development organizations, acting as the bridge between them and the communities they aim to serve</td>
</tr>
<tr>
<td>Mission Driven Bank Fund</td>
<td>Technical Assistance/Knowledge</td>
<td>Assists mission-driven banks with technical needs and fund management</td>
</tr>
<tr>
<td>More than Fair</td>
<td>Network/Intermediary</td>
<td>Coalition of organizations dedicated to improving access to affordable and inclusive credit, advocating for the use of technology and data science to increase the fairness of lending</td>
</tr>
<tr>
<td>National Community Investment Fund</td>
<td>Funder</td>
<td>Invests in CDFIs</td>
</tr>
<tr>
<td>NeighborWorks America</td>
<td>Technical Assistance/Knowledge</td>
<td>Helps support organizations with missions to provide affordable housing, finance, and healthy communities with technical assistance, including rigorous assessments. Created the Success Measure Program to help with performance measurement, directed towards mission-based organizations, including CDFIs.</td>
</tr>
<tr>
<td>Opportunity Finance Network</td>
<td>Policy/Advocacy, Technical Assistance/Knowledge, and Network/Intermediary</td>
<td>National network of CDFIs, acts as a financial intermediary, engages in lobbying</td>
</tr>
<tr>
<td>Organization</td>
<td>Primary Function</td>
<td>Relevant Program Summary</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>-------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Regional Federal Reserve Banks</td>
<td>Financial Service Provider and Technical Assistance/Knowledge</td>
<td>Federal-level banking institutions that also provide valuable financial research</td>
</tr>
<tr>
<td>Robert F Smith Foundation</td>
<td>Funder</td>
<td>Advocates for the 2% solution - directing businesses to put 2% of their wealth into black communities, including CDFIs as a vehicle for that investment</td>
</tr>
<tr>
<td>The Urban Institute</td>
<td>Technical Assistance/Knowledge</td>
<td>CDFI Research strand within Community and Economic Development Hub</td>
</tr>
<tr>
<td>U.S. Impact Investing Alliance</td>
<td>Policy/Advocacy</td>
<td>Works to raise awareness of impact investing in the United States by advocating for supportive policies, catalyzing investor action, and building the movement of impact investing broadly</td>
</tr>
<tr>
<td>University of New Hampshire Carsey School of Public Policy</td>
<td>Technical Assistance/Knowledge</td>
<td>Has very influential research centers that provide foundational knowledge on CDFIs and their functioning: Center for Impact Finance CDFI Research Consortium</td>
</tr>
<tr>
<td>Uplift America Fund (Opportunity Finance Network)</td>
<td>Funder</td>
<td>Example of a past partnership that leveraged federal resources, bank financing, and private grants to target capital to persistently low-wealth areas</td>
</tr>
</tbody>
</table>
Existing secondary public data sources relevant to CDFIs:

+ Bureau of Labor Statistics Quarterly Census of Employment and Wages
+ U.S. Census Small Area Income and Poverty Estimates
+ Consumer Financial Protection Bureau, Rural, and Underserved Counties
+ HUD Neighborhood Stabilization Program Targeting
+ USDA Food Access Research Atlas
+ U.S. Census Longitudinal Employer-Household Dynamics (LEHD)
+ Climate and Economic Justice Screening Tool (CEJST)
+ CDFI Transactional Level Report (TLR) data
+ National Credit Union Administration Call Report data
+ Current Population Survey (CPS)