

AI Solutions for Global Economic Empowerment



*Identification & Scoping for
US Challenges*

DataKind[®]

Nov 11th 2022

with support from

Google.org

datakind.org

Contents

Executive summary	1
Objectives	1
Key Findings	2
1. Learning Objective #1	2
2. Learning Objective #2	2
3. Learning Objective #3	3
Introduction	4
Sector Selection	4
Desk research	5
Access to capital	6
Access to services	8
DataKind Discovery	9
DataKind Historic Project Review	10
Individual Organization Discovery	11
Group Discovery Event	12
In-Depth Analysis of Economic Empowerment Thematic Areas	13
Access to Capital: Financial Inclusion	15
Access to Services: Safe & Affordable Housing	20
Post-Secondary Education	24
Employment & job retraining	28
Services that support mobility from poverty to stability	32

Executive summary

Economic empowerment is the goal of improving quality of life by increasing access to economic opportunities, while also providing the resources necessary to take advantage of those opportunities. Artificial Intelligence (AI) can provide new opportunities for data driven decision making, and a framework for solutions. This report details DataKind's process for research to understand the challenges and opportunities at the nexus of economic empowerment and data science, honing in on five key thematic areas to more deeply understand and engage in this sector. These thematic areas include: Financial inclusion; Safe and affordable housing; Post-secondary (university) education; Employment & job retraining; and Services that support mobility from poverty to stability.

Objectives

This report documents a landscape analysis - desk research and discovery conversations with over 20 social change organizations - on the broad topic of economic empowerment, and to share systems maps that complement these landscaping conversations. Synthesizing this information, we share key areas of opportunity for data science, and DataKind, to make an impact across the different thematic areas.

Key Findings

Over the course of our 2021-2022 partnership with Google.org, DataKind had a five-point learning agenda. Here, we provide key learnings against three of these points.

1. Learning Objective #1

What are areas where can AI be applied to make a significant impact in the economic empowerment space? Where are there gaps; where is work not happening because of a lack of organization taking it on and/or data that exists?

We operate from the premise that data science, machine learning, and AI can be deployed to effectively do three things: **observe** the current state of the world, **reason** to understand what will happen next, and **act** to impart outcomes and bring us closer to our goals. We saw opportunities in each thematic area for the application of data-driven solutions of varying sophistication. The areas of greatest sophistication arise from the existence of large amounts of homogenous data, which allow for the development and deployment of increasingly complex models.

Typically, these were identified in problems relating to *financial inclusion* and *post-secondary education* attainment. For less data mature thematic areas, we saw (and continue to see) the value of sophisticated methods for data processing - e.g. extracting information using image detection to obtain information from court filings (safe and affordable housing), or using natural language processing to manage unstructured textual data such as from 211 or 311 (services

that support mobility from poverty to stability). We see gaps in being able to apply sophisticated techniques at scale working with jurisdiction-level data across these themes, as each city, county, or state has its own approach to data collection and storage, which limits the generalizability and applicability of these tools. At a thematic level, the theme we found the fewest opportunities for immediate impact of data-driven approaches was in job retraining - while there potentially are methods that would support this problem, we did not identify clear opportunities.

2. Learning Objective #2

Where are there opportunities in the economic empowerment space for AI solutions to scale across multiple organizations based on the defined landscape analysis?

We see a clear opportunity for scale for financial inclusion solutions through both branch networks of single actors (e.g non-profit organizations that work at a national level, but operate with a local presence), and through national networks of credit unions or other Community Development Financial Institutions (CDFIs).

We also see key opportunities for AI to support post-secondary education attainment and clear pathways to scale through national alliances of higher education, and through university networks themselves. Data collected in these two thematic areas is largely standard across actors.

Lastly, we see promise in developing local solutions by leveraging hyper-local data such as the 211 network and data collected by key local stakeholders, from community-based organizations to local policymakers, as that data accurately reflect the current needs and realities of the geographies we aim to impact. In this space, we see the clearest pathway to scale for any solution through existing networks and channels.

3. Learning Objective #3

What is the overall readiness for AI versus data science techniques with the economic empowerment space based on the defined landscape analysis?

In the five thematic areas explored, we identified potential data science, machine learning, or AI solutions in each arena. We determined that three of the five topics -- financial inclusion, post-secondary education, and services that provide mobility from poverty to stability -- were more sophisticated in their perceived solution space as they are already data mature¹ and data

¹ A data mature organization is an organization that has array of data practitioners—data analysts, data engineers, data scientists -- and an invested leadership to work with DataKind in crafting a data science solution; they have the capacity to implement and sustain these solutions, and have a data culture and the tools to invest in strengthening and supporting their data journey. We use the [following checklist](#) with organizations to explore data maturity with a comprehensive lens.

accessible². Within those three thematic areas, we identified some solutions as being technically feasible, but requiring political will to make an impact.

The other two thematic areas were not as data mature and data accessible. For the safe and affordable housing thematic area, we were able to identify data sciencable problems, but access to data - whether public or organization owned- was uncertain. For the employment and retraining thematic area, data largely depended on the organization's ability to acquire and collect data in order to explore executing on a project together.

² Accessible data means organizations have data that can be shared with DataKind in the scope of the proposed project along the pre-determined and agreed upon timeline. This can include data proprietary to an organization, or open data which has been identified, acquired, and downloaded in a ready-to-share manner.

Introduction

Beginning in 2019, DataKind focused on developing approaches to providing data science solutions to increase economic resilience through our deep exploration of economic empowerment. Beginning with landscaping to support economic mobility worldwide and hosting Discovery events alongside the Global Inclusive Growth Summit, we focused on developing key partnerships to extend our reach and impact, and better understand how we could partner with organizations facing the same challenges across this sector to identify leverage points where data science could have an impact at scale.

Continuing our focus, DataKind was the technical partner for the data.org [Inclusive Growth and Recovery challenge](#) in 2020, where we learned from a global mix of inclusive growth projects. These bodies of work informed our point of view as we approached this landscaping report for United States economic empowerment with support from Google.org.

This report summarizes DataKind's Discovery work to better understand the opportunities and challenges at the nexus of data science and economic empowerment supported by Google.org. DataKind undertook discovery conversations with over 20 social change organizations and created systems maps for each of the five identified thematic areas, or sub-domains, that contribute to economic empowerment, reflecting areas of key potential intervention.

This report is structured in two primary sections: the first offers a review of our desk research and our methodology for Discovery conversations; the second does a deep dive into five thematic areas that support economic empowerment.

Sector Selection

DataKind's Discovery process examines multiple levels - organization, issue area, and systems level. To approach impact at scale, the broad system is selected and defined - e.g. "economic empowerment". DataKind then begins Discovery within this specific issue area through a mix of desk research and informational interviews. To determine if a domain is ready for potential data science interventions, we look for performance across a few topics:

1. **Maturity** - e.g. Has the field existed for more than 20 years? Are the goals of the field largely quantitative? Are a healthy set of operations conducted in the field done digitally?
2. **Size** - e.g. Are there major platforms/consortia acting in the space? Is there national or global reach? Has there been substantial funder investment?
3. **Innovation** - e.g. Are there major government limitations on data/its use? Are pain points understood/defined - e.g. through resource allocation or efficiency? Are there a

healthy amount of practitioners/non-researchers in the space and ripeness for disruption?

Desk research

Following initial sector exploration, we kicked off desk research across two themes of economic resilience and empowerment - access to capital, and access to services - with a focus on identifying unifying theories or proposals of what can and should be done to improve economic empowerment.

We looked to understand the common measures and metrics within each sub-sector, with a particular focus on understanding the possibility of identifying underlying factors that drive broad scale impact. We also focused on understanding the available datasets for particular issues, both traditional (e.g. unemployment numbers) and novel (e.g. scanning comments or social media for evidence of social behavior).

While exploring existing public data, we spent considerable time understanding key socio-economic trends. Digging deep into United States Census Bureau Data allowed us to visualize distributions of poverty, employment, and opportunity across the country (Figure 1). These insights encouraged us to look at supporting, through data science, social impact organizations that would target whole regions of the United States, such as the Sun Belt. This geographic area accounts for half of the population growth between 2010 and 2016, and has increasingly diverse cities as compared to the rest of the country (race, ethnicity, age of residents, socioeconomic strata, and local employment opportunities)³. Yet the growing challenge of rising income inequality, combined with rising housing costs and increased segregation, resulted in a 22% jump in low-income populations (as defined by the U.S. Department of Housing & Urban Development) between 2000 and 2016 across the region's metropolitan areas⁴. Poverty appears to be growing at a faster rate across the Sun Belt than in other areas of the country.

³ <https://www.newamerica.org/future-land-housing/reports/displaced-sun-belt/introduction-why-the-sun-belt-and-why-now>

⁴ <https://kinder.rice.edu/research/urban-sun-belt-overview>

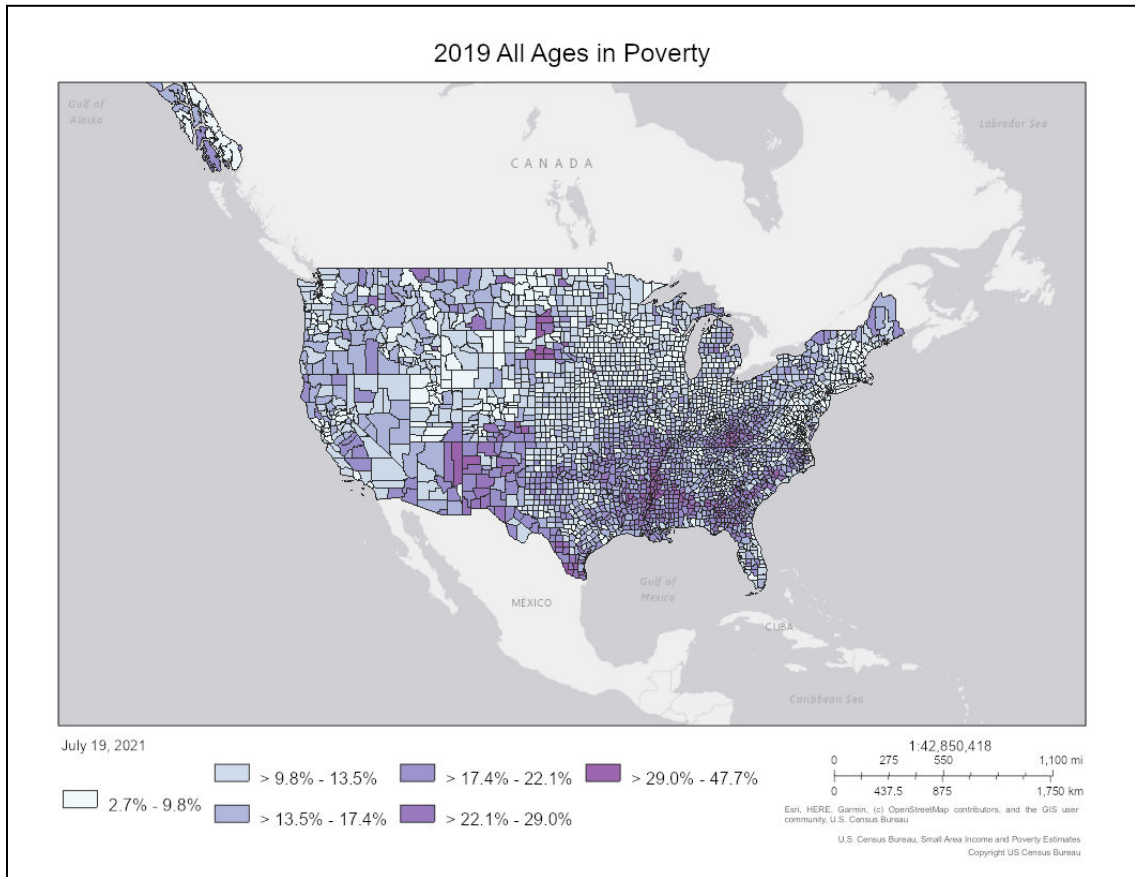


Figure 1. US Census Bureau Data, map representing poverty in the United States

Access to capital

Access to capital is broadly defined as having the ability to obtain and build formal financial assets, such as savings, insurance, and loans, either as an individual or a business. We investigated the problem of financial inclusion - the ability to access capital and participate in the formal financial market - and potential data science approaches as a path to informing opportunities for DataKind to make an impact.

Financial inclusion is a critical step towards reducing poverty and supporting the financial growth and financial health of individuals, families, and small businesses. Having a basic bank account can help raise standards of living: people can save more, borrow money for education or to start a business, buy property, and invest in new opportunities. Yet over 22 million people in the United States do not have a formal bank account, and estimates suggest that unbanked families across the country spend up to 5% of their annual incomes on managing their basic finances⁵. While major financial institutions have been able to invest billions, if not trillions, of dollars^{6 7} into digital transformation to engage in successful marketing, targeting, and other offerings, financial

⁵ <https://www.centerforfinancialinclusion.org/bank-on-battles-financial-exclusion-in-the-united-states>

⁶ <https://www.businesswire.com/news/home/20150629006038/en/Retail-Banks-Spend-16.6-Billion-Digital-Transformation>

⁷ <https://www.bloomberg.com/news/articles/2019-06-20/banks-spend-1-trillion-on-digital-but-few-benefit-study-say>

service providers (FSPs) with the greatest programs for reaching marginalized populations (credit unions and community or regional financial service providers) have not had a similar capability for digital conversion. This results in slow, more expensive, tech-poor solutions meant to facilitate the financial inclusion of millions. And, due to the resource intensity of these solutions, the reality is that the needs of the underbanked are not met.

First, looking at digital finance. Digital finance, due to its wider geographic reach, increased access, and lowered cost of doing business, increases the availability and access of traditional finance, making it an option for the underbanked population. By unblocking the potential for digital transformation and ensuring confidence in the data collected, systems for data-driven decision-making and the creation of predictive models can anticipate both need and opportunity for under-banked populations.

Second, looking at financial inclusion through individual and small business loans.

While loan repayment is a specific problem that has both digital and non-digital repayment solutions, we understand that broadly, increasing adoption of digital financial services reduces the underbanked population. The

underbanked population does not have their needs met or does not have sufficient access to the current financial system, and relies on alternative financial services to meet their financial needs. These services are often associated with high fees, limited consumer protection, and low opportunities for wealth-building; in short, they perpetuate economic inequality. The potential for loan delinquency is a problem that limits financial service providers' willingness to offer unsecured loans to clients who fit all but the narrowest potential profile for business and/or individual loans.

An example of exclusion through business loans occurred during the COVID-19 pandemic. The CARES Act, signed in March 2020, was the first of several relief packages for American workers and small businesses. This national program, in combination with state programs in charge of local distribution of funds, was a lifeline for small businesses across the country as they tried to survive the economic challenges brought on by COVID-19. The funds being distributed by national and local programs are unprecedented in terms of amount, scale, and speed of distribution. Yet across the deep South (Alabama, Arkansas, Mississippi, Louisiana, and Tennessee), the federal Paycheck Protection Program's \$600 billion aid largely bypassed minority-owned businesses and other small businesses despite their critical contributions to the economy (nearly 1 in 4 businesses in these states are minority-owned). It is likely that data science approaches could help shine a light on these inequalities.

For individuals in the United States, one-in-five adults do not have sufficient information for loan officers to make an assessment on them (with one-in-ten being deemed “credit invisible”)^{8 9}. Nearly 11% of these adults, or 2% of the US adult population, can afford to buy a house¹⁰ but do not have sufficient traditional financial history to be considered a safe bet for a home loan. And, according to research from [Upstart](https://www.upstart.com), “just 45% of Americans have access to bank-quality credit, yet 83% of Americans have never actually defaulted on a loan.”¹¹ Banks are cautious to

⁸ <https://www.consumerfinance.gov/about-us/newsroom/cfpb-report-finds-26-million-consumers-are-credit-invisible/>

⁹ <https://www.consumerfinance.gov/about-us/blog/using-alternative-data-evaluate-creditworthiness/>

¹⁰ <https://www.thesimpledollar.com/credit/build-credit/alternative-credit-help-you-get-a-mortgage/> ; <https://singlefamily.fanniemae.com/media/21466/display>

¹¹ <https://www.upstart.com/blog/upstart-ceo-dave-girouard-testifies-in-congress-about-ai-in-credit-underwriting>

increase the number of loans to these individuals. If a loan defaults, the financial service provider no longer has access to the anticipated revenue stream, which puts its own financial solvency at risk. The service provider also loses the ability to intervene on a customer's behalf.

If financial service providers had the ability to accurately predict the likelihood of loan delinquency, they would be more inclined to initially offer loans to qualified borrowers. For loans being serviced, a prediction of potential default would allow lenders to offer early/proactive interventions to their clients to limit past-due and delinquent loans¹². This could increase loan rates of the millions of individuals and businesses currently being rejected for loans and close the fair lending gap.

Access to services

Financial security is jeopardized when shocks to people's lives throw them off-course. While it is important to ensure access to capital, ensuring that people also have access to basic services is similarly key. Not all people have access to dignified work because of systemic challenges working against them. Strengthening the social safety net and ensuring people have access to housing, food, healthcare, transportation, and other basic needs ultimately increases economic empowerment.

The "geography of opportunity" as popularized by Rosenbaum in 1995¹³ has led to an explosion of research and policy recommendations - from improving labor mobility to rethinking transportation infrastructure to building community and improving access to education. It is widely recognized that race, gender, and place intersect with and intensify the challenge of economic inequality. Innovations such as opportunity zones, research into food and healthcare deserts, and other initiatives have shed light on this problem. The impact of geography is clear, and the ongoing transformation of local government data collection to digital, the expansion of local data networks such as 211 and 311, and other digital data opportunities have highlighted that not only are there place-based problems being identified, but there is enough of a data trail and data story such that tools and systems for identification and early intervention can be created.

Using as a specific example the problem of housing insecurity, defined as missing a tenant of adequate housing as outlined by the United Nations, we explore this problem. This situation might include missing tenants such as security of tenure, availability of services, materials, facilities and infrastructure, affordability, habitability, accessibility, location, and cultural adequacy. While a recognized problem, inconsistent and patchwork measures of housing insecurity largely obfuscate the issue at a national level. In the United States, for a majority of states, between 10 and 15% of households are housing insecure, with states with high costs of living such as New York and California having 20%+ of households facing housing insecurity.

¹²<https://www.forbes.com/sites/jeffkaufman/2020/04/16/exclusive-early-data-shows-12-of-online-loans-in-trouble-double-just-weeks-ago/#2df7f38018c1>

¹³ <https://www.tandfonline.com/doi/abs/10.1080/10511482.1995.9521186>

While housing is a clear example, it is simply one example of the place-based issues and opportunities that individuals face by virtue of their geographic location. Our desk research identified multiple opportunities for data-driven solutions to help increase access to housing and other services.

DataKind Discovery

Moving from desk research to project discovery (the identification of potential problem statements where data science and AI can meaningfully contribute to a solution), DataKind reviews our historic projects and partners, leads individual conversations with new organizations, and hosts group Discovery Day events. When DataKind undertakes specific organization-level discovery, it is with the understanding that the solutions we create have the potential to generate positive impact within the sector.

We seek to complete Discovery that reflects the diversity of experiences and problems within that system - or, in other words, we hope to discover and answer “*what if the highest impact DS/ML solutions existed for this problem?*” Our individual project discovery seeks to straddle the tension of impact and usability for an individual organization, with the value and applicability of the solution beyond that specific organization.

DataKind’s principles when conducting Discovery

1. **Learn from experts & listen to all voices:** Engage end-users and end-beneficiaries to understand their pain points, the environment a solution must succeed in, and the constraints on use. This has led us to design solutions that can work with fully offline functionality, or ones with translation built in, or workflow recommendations that remove the need for our tool in the first place.
2. **Focus on leverage points for impact at scale:** DataKind recognizes that data science cannot be a solution for everyone and everything. We try to take the questions that our partners bring up, and identify which of those potential questions has the greatest likelihood of changing and sustainability impacting their system. This leads us to explore where we can be deployed, leveraging data science/ AI/ ML as a tool in the toolbox to effect systems change.
3. **Ethics, privacy, and risk management should be prioritized from the start:** We want to focus not just on what might happen if our project fails, but also what happens if our project succeeds, and the ethical and other considerations of the solutions we build. For example, we remain aware that automation can lead to a loss of jobs, and that a tool that identifies vulnerable communities for positive social outreach could fall into the wrong hands. We build these hard conversations into our process, and build safeguards into our solutions.

DataKind Historic Project Review

Our project discovery learns from DataKind's historic projects, both to inform and strengthen our thematic approach and to recommend potential partners for new projects. DataKind's project history spans years, geographies, and technologies, and we believe that we have significant learnings regarding what makes a project successful. We specifically want to highlight projects that significantly helped inform our thinking as to the data science viability of specific projects and themes within economic empowerment.

Access to Capital

[Microcred](#): Microcred, most recently known as Baobab, is a French digital finance group whose work aims to contribute to the growth of local economies in developing nations by offering simple accessible financial services. Using Microcred's loan application data and internal loan status information, DataKind aimed to predict customer default to better inform decision-making about lending in order to make it more efficient and inclusive. DataKind created predictive models and an analysis pipeline to build default scores for customers at different points in a loan cycle, helping make the loan application process more efficient and effective in determining default risk.

This project highlights the feasibility and impact of developing and applying predictive models in financial inclusion, from identifying loan application and repayment behavior as significant predictors of default, to applying complex data science techniques such as a Monte Carlo simulator to help ascertain expected revenues and expenses associated with customer lending.

[Simpa Networks](#): Simpa Networks is a social enterprise that provides pay-as-you-go solar power to households and small businesses in villages in Uttar Pradesh, India. In collaboration with DataKind, Simpa Networks aimed to use historical customer payment behavior data to predict which new applicants were most likely to be a good fit for their pay-as-you-go solar power systems program by predicting whether a potential customer was likely or unlikely to keep up with payments.

Learnings from this project further reinforce the scalability of leveraging data science, machine learning, and AI to support the financial inclusion of traditionally excluded populations, and through this, their ability to access core services to their wellbeing, such as energy for electricity. Project execution also identified the need for robust customer history to build accurate models, or alternative approaches such as incorporating publicly available data as part of the model development.

Access to Services

[John Jay College](#) (JJC): In collaboration with DataKind, John Jay College of Criminal Justice used historical student data to predict students who were most at risk of taking an extended amount of time to graduate or who would drop out of college. Project outcomes aimed to provide JJC with the data to identify these students earlier in their college career and design proactive rather than reactive interventions to help improve college completion rates. DataKind used statistical analysis and machine learning techniques to analyze historical student data, built a predictive model of the likelihood of student dropout, and identified the characteristics that distinguished students who eventually drop out from those who graduate.

This partnership provides insight into how leveraging data science, AI, and machine learning can sustainability impact post-secondary education outcomes, from a clear understanding of the type of data that is collected across higher education to insight into the key challenges administrators face in trying to support students most at risk. This project also demonstrated the incredible value-add data science could have in a specific area, while remaining cognizant of the importance of human interpretation and contextualization as key to maximizing impact for students.

[New America, Future of Land and Housing \(FLH\)](#)¹⁴: New America Future of Land and Housing works to bridge the gap between policymakers and technologists to ensure and empower individuals' right to property, including safe and affordable housing. In partnership with DataKind, New America FLH created tooling that invites a hyperlocal understanding of housing loss, including eviction and mortgage foreclosure. The creation of complex data science and statistical approaches to identify who was most at risk of housing loss, and the development of census-tract level heat maps to point to where this was happening, enabled partner counties to use these insights to inform relief disbursement under the CARES Act.

Partnership with New America- FLH provided DataKind with numerous learnings, most importantly the possibility of leveraging open and publicly available data to generate insights and learnings at scale. While housing data is disparate and varies between localities, the opportunity of collecting, centralizing, and analyzing uniform open data can have an impact at scale, irrespective of the type or organization or geographic location.

Individual Organization Discovery

DataKind approached new individual organization Discovery through a mix of direct outreach to previous and prospective project partners, open solicitation of potential partners, and targeted introductions through DataKind's network. In Discovery, we are most interested in understanding

¹⁴ Formerly known as Future of Property rights (FPR).

what's most important to organizations to then map out how this impacts the domain at large. We aim to answer questions like, "*What do they want to achieve?*" and "*What's standing in their way?*"

At DataKind, we understand that data can help us dream up ideas, but the questions are most critical to ensuring we form a solution that creates value.

DataKind Partnership Alignment

As [described](#) by DataKind's Center of Excellence, in Discovery, we look for alignment in two main dimensions:

1. **Ability to engage:** DataKind wants to work with partners who are able to work with us on timelines that mutually make sense, and who have the capacity to take on a partnership.
2. **Data maturity:** DataKind wants to work on problems that have the potential to create positive social impact beyond the individual organization. We look to understand mission alignment, data availability, and sector representation.

Group Discovery Event

DataKind regularly hosts group discovery events called [Discovery Days](#). These hands-on work sessions aim to explore what's possible with data science and identify opportunities for high impact projects. The goal of a Discovery Day is to identify pain points and explore high level "data scienceable problems".

To better understand the domain of economic empowerment, DataKind hosted a Discovery Day event to complement our individual organization discovery. Insights from this event allowed us to understand the current challenges faced by organizations to augment their work as identified in the desk research phase. Of the problems heard, organizations were looking to leverage data to more effectively measure their impact and outcome; to identify data-scienceable ways to extend their reach and impact, with an explicit commitment to diversity, equity, and inclusion; and to train colleagues to better understand data and drive data-based decision making.

In-Depth Analysis of Economic Empowerment

Thematic Areas

In the next five sections, we provide a deeper analysis on the five topical areas we identified under access to capital and access to services as our areas of focus for economic empowerment. Specifically, we provide an overview of each thematic area identified that supports and ladders up to economic empowerment in the United States. For each, we provide an overview of the sector challenge, a summary of specific issues identified with potential data science solutions, and a systems map such that we can better identify how influencing one leverage point in that system can positively or negatively impact economic empowerment outcomes for individuals, small businesses, and communities at large.

Systems mapping

Inspired by The Omidyar Group's Systems Practice guide, we created systems maps for the thematic explored under Economic Empowerment. This approach allowed us to present each thematic area as part of complex, dynamic systems, rather than focus myopic causal pathways. Each systems map presents, in turn, the various factors that contribute to individuals, small businesses, and community needs toward economic empowerment, and how each node contributes or hinders this ultimate goal. The review and assessment of the maps provide insight into the leverage points where data science can be used as a tool to catalyze and amplify successful outcomes, while bringing to bear any negative consequences that changes in the systems could effect.

In each thematic area, we see opportunities to leverage data science, machine learning, and AI for positive social impact. We also see places where the technical solutions are feasible, but there is a requirement for political will in order for the solution to have an impact. In other situations, we identify opportunities to develop solutions, but with a need to first develop a strong ethical stance or solutions (e.g. for sharing protected data).

The systems map below provides, at a high level, an overview of the five themes we are exploring and how they are linked across a system that aims to support individual, small business, and community economic empowerment. As we synthesize findings and map out the system, we determine that impacting one or two of the key thematic areas further provides positive repercussions across additional domains. For example, increasing access to safe and affordable housing provides student stability to academically excel, thus encouraging them to graduate with a Bachelor's degree (post-secondary education). This high level map captures how data science and AI could impact and strengthen economic empowerment across the United States.

Access to Capital: Financial Inclusion



The inability to access capital is a barrier to growth and resilience. The failure of traditional banks to adequately and equitably loan money can be attributed to many factors, including their aversion for risk, which is compounded by the difficulty of evaluating risk for small businesses or borrowers in disadvantaged communities. Combined with the high transaction costs of borrowing, this results in an inadequate supply of capital. Asymmetric information, imperfect competition, systemic racial bias, and avoidance of small-sized loans results in many businesses deemed valuable and important by local communities unable to grow, let alone survive, economic crises.

Summary of Issues Identified

Loan repayment prediction

A key issue heard across organizations was regarding financial stewardship - particularly when offering business loans. The potential for loan delinquency is a problem that limits banks' and CDFIs' willingness to offer unsecured loans to clients who fit all but the narrowest profile for business loans. If financial service providers had the ability to accurately predict the likelihood of loan delinquency, they would be more inclined to offer loans to new and qualified, but untapped, borrowers. For loans already distributed, predicting potential default would allow lenders to offer early and proactive interventions to their clients to limit past-due and delinquency. The current data that financial service providers have is limited to what they internally collect.

Small-to-midsize financial institutions often do not have the resources to extract and combine multiple data pipelines, which makes it challenging for them to accurately assess applicant viability or proactively assess if an existing loan will be serviced. Without intervention, they'll

continue to reject these loan applicants because they cannot create more sophisticated repayment models.

We see this as an area where data science is well suited as a tool to support the goal of “providing more loans without increasing delinquency rates.” It is possible that more sophisticated loan repayment models could solve this problem. We believe that combining data sources from multiple financial service providers, identifying alternative financial services behaviors, and creating alternative loan application review systems could potentially reduce loan default. We identify a potential data science solution as creating a model that predicts repayment behavior (across different loan products) for both new and existing clients to support banks’ and CDFIs’ ability to make data-driven decisions about borrower risk, and offer high-impact services to traditionally excluded communities. In a related issue, we also see data science as being able to provide insights into how business loans have been distributed, and identify inequalities in those approaches.

Digital services personalization

A second issue encountered in our Discovery processes was access to digital financial services, and particularly, the ability for financial service personalization. While digital financial services are seen as a key method for bringing active banking to communities that are traditionally excluded from the formal financial system due to distance or presence of a bank branch across varied communities, language barriers or financial literacy for interaction, or ease of transaction (e.g. immediate access to funds.) These barriers can be removed with digital financial tools through which financial institutions can more actively serve underbanked individuals and provide tailored and personalized financial services to support their clients’ economic growth. Access to and use of digital banking was a stark societal divide during COVID-19 when most physical banks were closed. However, with the adoption of “pro-poor” products, services, and policies by many fintech and conventional banking organizations, this digital divide could be closed for those who are underbanked and non-adopters.

We see potential to create, build, and refine predictive models to assess existing clients’ propensity for adopting and using digital financial services to enable banks/ CDFIs to drive more personalized interactions and reach clients otherwise unable to access formal financial institutions. Data science approaches can help financial service providers understand their customers and identify those clients who are currently underbanked, but who might be adopters of digital services. Data science can assist in identifying those underbanked (dormant bank accounts, using alternative financial services), those who live outside of an easy radius of a physical bank or might have a language barrier for in-person banking, and those defined as “financially reachable” but currently marginalized. Once identified, this information could be used to help financial service providers design training and adoption programs and interventions with target populations - perhaps through the use of prioritization or rank ordering. These interventions are often initiated through agent networks, where, armed with information, financial service agents will go into a community to work with individuals to increase adoption of the digital tool. The information provided through a data science analysis is believed to be stronger and more nuanced to aid in this intervention. Once adoption has occurred, the digital data created through the digital financial system can allow for algorithmic approaches to

recommending additional financial services and behavioral nudges. This can include customer segmentation algorithms to best serve client needs once the technology is adopted.

Alternative credit scoring

The third problem we regularly hear in conversation is similar to our first - this problem is focused on providing access to capital for individuals. With modest funds and no credit history, marginalized, unbanked individuals are expensive to serve in the conventional banking channels and are often thought to have low potential to create value. While the traditional and most common way to evaluate the creditworthiness of individuals is their credit history, this is simply not available for the unbanked. They predominantly rely on cash and don't have any traceable financial history; their thin credit files become hindrances to growth in a banking industry that prefers thick ones.

Data science solutions have been proposed and deployed in some respects for this type of problem - Experian Boost is a particular example. Alternative credit scoring and predictive models are a solution that would make credit lines for the unbanked faster, cheaper, and credit worthy. For consumers, credit scoring is the key to home ownership and consumer credit. For businesses, especially small and medium-sized enterprises, credit-scoring increases access to financial resources, reduces costs, and helps manage risk.

However technically feasible, this solution will only work if credit issuers are willing and able to take alternative information into account. A key issue of alternative credit scoring is that these scores have not demonstrated "stickiness" or staying power in the marketplace. While many organizations are conducting thoughtful research in this space, we don't see DataKind's model as well-situated to developing and supporting such a crucial gatekeeping tool.

Systems map visualization

Based on the systems map of financial inclusion presented below, we can identify that access to credit/ loans is a critical step in achieving economic empowerment. Leveraging data science to unblock one of the key pathways to accessing loans -- such as supporting small businesses' ability to repay a loan by predicting default such that financial institutions can provide pre-emptive support and remediate delinquency -- is key to economic empowerment. Accessing loans unlocks key pathways for economic mobility and empowerment by stimulating the local economy (new or growing businesses provide additional employment opportunities), stabilizing one's health and well-being (employment can provide health benefits; savings can offer the chance to seek medical care), and by building overall wealth (accrued savings can be invested into new assets).

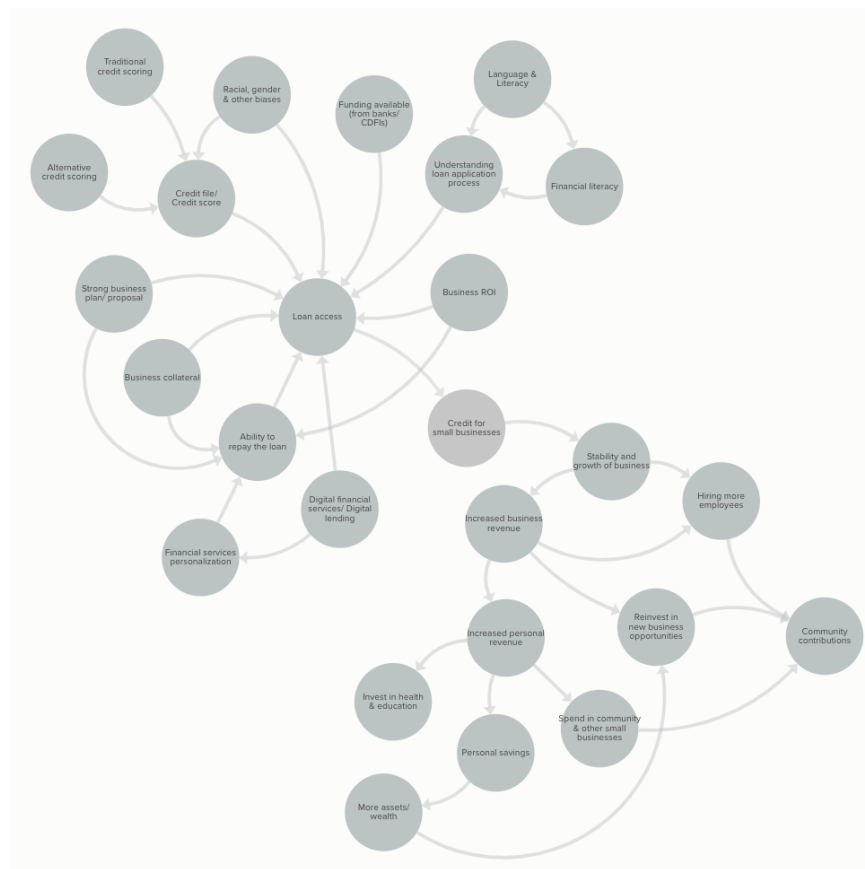


Figure 3: Systems Map of financial inclusion
 [For an interactive map, [click here](#)]

Summary of data science techniques

Fundamentally, financial service providers face a problem of risk management - they want to minimize loan default and expand their financial services footprint, yet don't have robust enough models that would allow for recommendations to make a banking decision. We see potential activities around risk assessment, branch location or agent location optimization, and machine learning for targeting.

DataKind believes there are opportunities to recommend generalized solution frameworks and approaches here.

- **First**, most loan models in use rely on internal data collected by the financial institution; we recommend augmenting with external information regarding geography, demography, and sector performance.
- **Second**, we recommend approaching the problem using machine learning techniques that can handle the asymmetry of historic data (i.e. imbalanced classes) which has limited the success of modeling endeavors in the past. The data science techniques we have identified in the financial inclusion space include the use of predictive analytics to reduce and remediate loan default, and data segmentation to provide tailored financial recommendations to clients based on individual profiles, including assessing current and

predicting future needs.

Readiness for DS/ML/AI assessment

The financial inclusion space is ripe for data science/ AI engagement for multiple reasons.

First, the data collected in this space is similar across financial service providers, irrespective of the type or scale of the organization. The data in this space is client centered, meaning that organizations will collect data on their clients, but also about the client- financial institution interactions. Examples include socio-demographic characteristics of the individual; location and business information (revenue, growth); and loan amount and length, if already received.

Second, the financial inclusion space is a data mature environment in which data already drives much of the decision making processes. However, there is space and opportunity for data science and AI to automate and amplify this work to drive towards increased levels of financial inclusion. Given this, projects supporting the financial inclusion of small businesses across the United States not only have the potential for scale given a similar data set, there is potential to develop tools or methodologies (e.g predicting loan repayment) that can be generalizable for the creation of a global good.

Access to Services: Safe & Affordable Housing



Access to safe and affordable housing is a key determinant of well-being and resilience for individuals and their families. Yet every year, 5 million Americans lose their homes through evictions and mortgage foreclosures, a displacement that is linked to poor outcomes, including adverse health impacts, gaps in educational attainment and chronic homelessness. In addition, access to stable homes and the ability to purchase homes as a wealth-building opportunity is challenging due to equitable access to services for rental and other assistance, as well as racial and other disparities affecting access.

Summary of issues identified

Lack of local information to drive decision making

A key issue we heard across many access-to-services organizations, but most acutely when discussing with housing organizations, is the lack of access to local information to drive decision-making. While a recognized problem, inconsistent and patchwork measures of housing insecurity largely obfuscate the issue at a national level. While data on housing and income is available, they are not used extensively for analytics or scenario-based planning. Locally collected data can provide early indicators of housing insecurity that are likely or known to be predictors of vulnerability or eventual housing loss such that interventions can be deployed. The patchwork of indicators, databases, and intervention across local geographies make it challenging for end users to build the systems and proactively intervene to increase housing stability through direct social intervention (e.g. loans, loan forgiveness, cash transfers, public works, etc.). Decisions about where and how to allocate funding or services that support individuals, families, and communities in need are often made ad-hoc on the basis of

experience, but without comprehensive data. Building on the national success of work DataKind and New America Future of Land and Housing have led, we recommend to first leverage housing loss data at national and sub-national levels to create data heat maps and other visualizations to generate easy to use insights into where housing loss is most acute, and where resources should be focused to support individuals most at risk of losing their homes. With a foundation of data trust, we next recommend developing hyper-local predictive analytics to anticipate future needs.

Generational wealth building opportunities through homeownership

Homeownership is the fastest way to build generational wealth in the United States. For homebuyers, 75% are unaware of the financial and other resources available to them to invest and build their wealth. In this instance, we see a clear linkage between safe and affordable housing projects and financial inclusion projects. Potential approaches to further explore and unblock this challenge include identifying and mitigating biases that removes proxies for racial biases for banks such that they lend to credit worthy individuals who want to buy homes, and supporting CDFIs to expand access to credit in minority and lower-and middle-income communities. Prospective homebuyers, especially Black Americans, who have been most affected by national wealth disparities caused by inequitable housing policies¹⁵, would then be able to equitably access loans. This can be compounded with predictive models that provide tailor-made recommendations for loan size, duration, and repayment options to reduce default and therefore promote homeownership as an opportunity to build wealth and accrue assets.

The exploration of this data scienceable problem falls at the nexus of financial inclusion and safe and affordable housing, and is therefore an example of the interplay of multiple thematic areas in support of economic empowerment. DataKind remains cognizant, however, of the sensitivities around algorithmic assessment of proprietary algorithms developed and owned by entities with whom we are not directly working, and would recommend focusing on the financial inclusion route as a data-science first exploration of this sub-theme.

Systems map visualization

Based on the systems map of access to safe and affordable housing presented below, one's access to homes - whether rented or purchased for wealth-building opportunities- proves to be a critical foundation upon which individuals across the United States can become economically empowered. Leveraging data science to unblock one of the key pathways to accessing and strengthening the stability that safe and affordable housing imparts is key to economic empowerment. Examples include unblocking rental assistance such that governments and community-based and other organizations can provide support to individuals and families with the highest need, or ensuring unbiased opportunities for diverse individuals to access competitive market rate loans to purchase their first homes such that they can build their and their children's wealth. Safe and affordable housing is key to one's economic growth, as it provides individuals with the safety and stability needed to succeed in their academic and career aspirations.

¹⁵ <https://www.brookings.edu/blog/up-front/2020/02/27/examining-the-black-white-wealth-gap/>

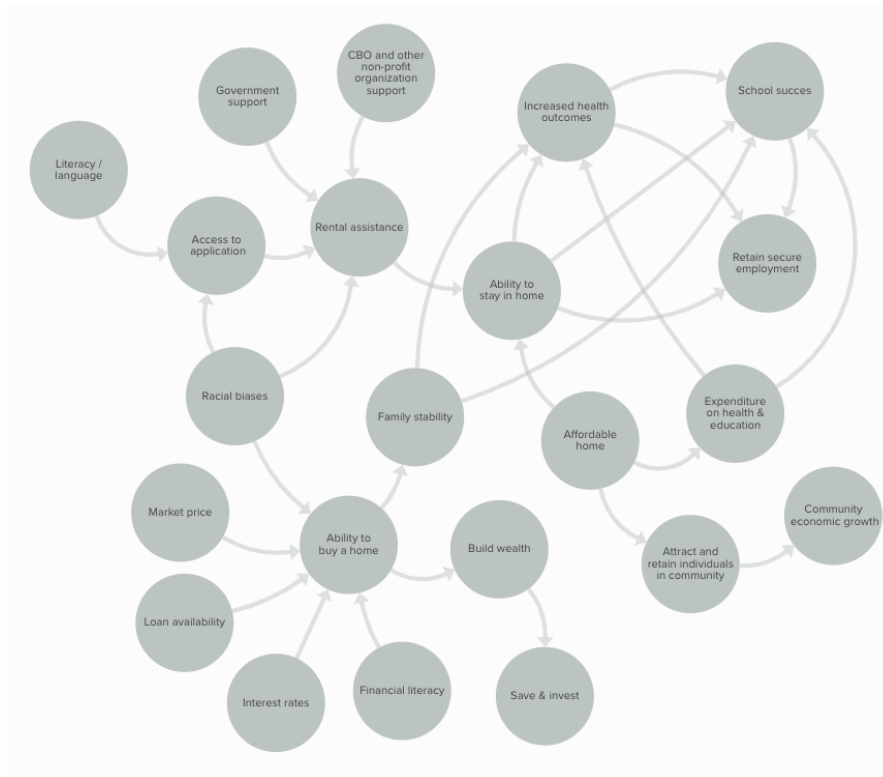


Figure 4: Systems Map of safe and affordable housing
 [For an interactive map, [click here](#)]

Summary of data science techniques identified

With a primary goal of understanding what variables might help predict future housing loss so as to provide recommendations on how to intervene on housing insecurity, we identify several data science techniques that put an emphasis on the intersection of housing vulnerability, race, and equity.

From preliminary studies, we identify principal components analysis as an approach that is useful for explaining housing loss. We also identify geographically weighted regression as a way of exploring and identifying known truths - we know that housing types (e.g. owner vs. renter) are different, and we know that there are spatial dependencies (e.g. certain areas experience higher rates of housing loss). When geography plays a critical role in the predictor-response relationship, this approach will handle spatial effects. We recommend similar regression approaches such as a poisson regression as it is generalizable to more jurisdictions. However, to begin these analytics, first a data pipeline must be established, which is the key barrier to solution-building in this respect. Machine-learning and AI approaches can be helpful in establishing the data pipeline, as there is known valuable information held in image and PDF files (e.g. court records) that could be extracted with image detection methods, and there is valuable court docket information that could be parsed with natural language processing approaches.

Readiness for DS/ML/AI assessment

While there is great potential for high impact projects under the safe and affordable housing theme, ability to execute on projects that leverage data science and AI tools largely depend on the organization's readiness and data maturity, notably access and availability of data within the proposed time frame. Furthermore, DataKind's experience in this space has highlighted that housing data is highly reflective of the disparate housing policy environment across the United States. As a result, data varies between states, and even county by county. Given this, housing projects have the potential to be highly impactful at local or even regional levels, but the opportunity to scale nationally will prove difficult without a national partner or other commonality where an aspect of housing loss data is standardized.

Post-Secondary Education



Graduating from a college education is one of, if not the most effective way for individuals to increase their social mobility and advance economic opportunities. According to data from the American Institutes for Research, only 60% of college students graduate from four-year colleges and universities within six years¹⁶, and only 41% of bachelor's degree earners graduate within four years. Students are generally unable to graduate with a degree due to a variety of factors: students have difficulty learning, or students learn but are unable to remain in school due to lack of motivation, family, or other challenges. If students return to the school to fulfill their credits, the college or university often fails to have the appropriate tools to provide tailored guidance for students to feel re-engaged.

Summary of issues identified

Predicting student drop out

In speaking with numerous post-secondary education organizations, we learned that predicting student drop-out is a key concern. While this builds on our single-institution work, we understood it to be a problem experienced by administrators of programs at both the associates and baccalaureate levels. Administrators do not have the visibility into students who are having the most trouble maintaining satisfactory grades and academic engagement and who eventually drop out. Understanding who is most at risk such that they can preemptively intervene to

¹⁶ Source: <https://www.communitycollegereview.com/blog/the-catch-22-of-community-college-graduation-rates>

support these students with the appropriate services is time consuming and requires sifting through vast quantities of data with no clear understanding of where interventions would have the highest impact.

This nuance extends DataKind's previous understanding of the problem in two key ways.

- **First**, there is an interest in predicting risk of dropout earlier in a student's academic career. Where DataKind's previous mandate was to work to develop risk models for students who were near graduation yet dropped out, we have heard a need to also understand dropout rates earlier in the process, particularly within the first three semesters of an associates program.
- **Second**, we now understand transfer students from associates to baccalaureate programs as a unique student population, and there is a specific need to provide support early in the transfer (but mid-academic career) as a stage of intervention.

However, we believe that the same machine learning principles can be applied to each solution space. We suggest exploring to build a predictive model to identify which students are most at risk of dropping out, including their behaviors and academic characteristics, such that the administration can anticipate and be proactive in providing services to students to ensure they remain in school and graduate with their degree.

Developing academic schedules that best support student success

A second problem we heard from administrators is how to best provide courses that meet their students' needs - effectively, how to provide programming that meets demand. Students in community colleges often face myriad social and family challenges and pressures that impede them from successfully graduating with their associate's degree. These factors such as the need to support family with child care, needing child care themselves, their engagement in full or part time employment, or the provision of other external support takes priority over college classes to graduate with an associate's or bachelor's degree. We see opportunity to draw from demand models in other scheduling industries to build a predictive model leveraging past course data to assist colleges in scheduling classes at times more suited to when students are able to take them. This will encourage and motivate students to complete their courses (and, ultimately, their programs) at a higher rate.

Systems map visualization

Based on the systems map of post-secondary education presented below, the ability for youth to access community or 4-year colleges and graduate with Bachelor's Degrees is critical to their economic growth and well-being. Leveraging data science to unblock one of the key pathways to successfully graduating from university, such as increasing access to financial aid by matching individuals with available scholarships or ensuring university administration has the data-driven information needed to support students with the appropriate services needed to graduate, is key to long-term economic well-being and empowerment. These unique leverage points bolster a student's chance of graduating with a 4 year degree, thus ensuring that they enter the marketplace with a competitive advantage, leading to secure employment, a

sustainable salary, and health and other benefits, all of which lay the foundation for stable economic growth.

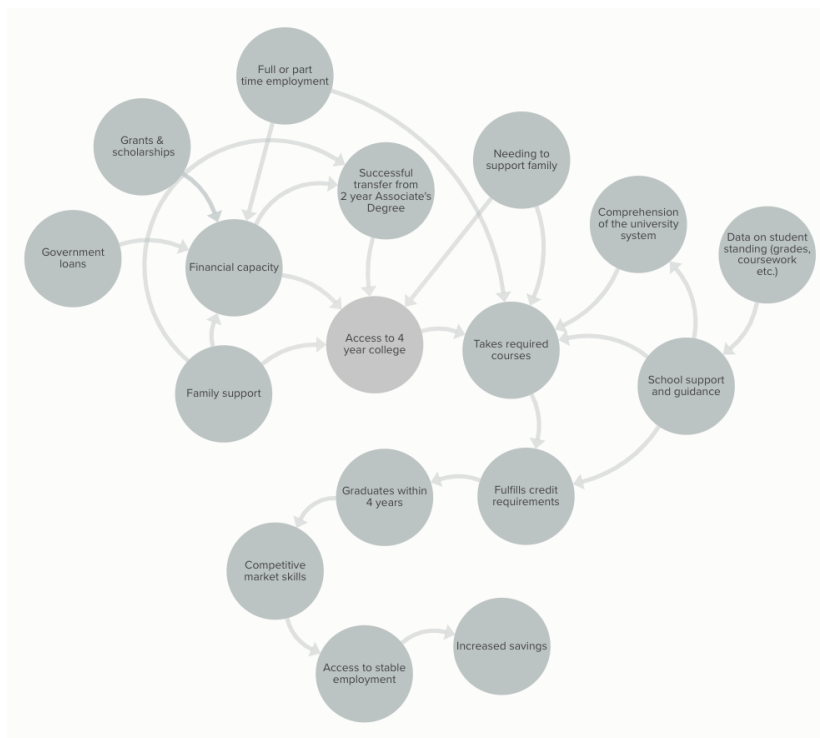


Figure 5: Systems Map of post-secondary education
[For an interactive map, [click here](#)]

Summary of data science techniques identified

We recognize that education delivery occurs in both formal and non-formal settings for assistance in projects around resource optimization, program effectiveness, and program usage/attrition. We see a clear opportunity to deploy predictive models focused both on student performance and university performance. Drawing from other sectors, we see demand modeling (e.g. airline or hotel models) being a useful framework for approaching university staffing to meet student demand and program expansion. In addition to the machine learning modeling approaches, we also recommend modeling approaches adapted from the biological sciences, such as survival models.

Readiness for DS/ML/AI assessment

The post- secondary education space is ripe for data science/AI engagement for multiple reasons. First, the education space is a data mature environment in which data already drives some of the decision making processes, such as using data from student feedback surveys to improve class instruction, and exploring measures of university success by aggregated GPA or graduation data. New datasets of school records on student performance, online mentoring interactions, data collected by NGOs, and open sources of government data offer the promise of boosting impact and scaling the services of social change organizations exponentially. These massive datasets could help schools and organizations react to student needs in real time,

understand dropout and graduation trends, and optimize school operations. However, data alone is not enough. To unlock the power of these new datasets, a combination of skills in computer programming, statistics, machine learning, data visualization, and software design is critical.

Further, the data collected at community colleges or universities is very similar across the board; data including classes or credits taken, attendance, grades and GPA are universal US benchmarks for academic performance and achievement. As such, there is great opportunity for scalability and generalizability of a tool to increase student performance and encourage graduation to support individual economic empowerment.

Employment & job retraining



Traditionally underserved populations have faced a long history of structural barriers in labor markets in pursuit of high-skilled employment. While these populations often live in close proximity to open jobs in high-value sectors (e.g. technology), they often lack the opportunities to build the technical skills needed to be competitive in that market, or know where to find jobs that will be best suited to their skill set. Tools are needed to inform workers of labor market trends, diagnose skill gaps, and help them find jobs or skills that will empower them to plan for future career opportunities. We learned of specific opportunities for policy prototyping¹⁷ that could be data-supported or data-driven, and within those areas, specifically focused on areas where we might directly apply data science.

Summary of issues identified

Job retraining and skills building

Many workers struggle to find education or training programs that could unlock career pathways for them. To successfully pursue new opportunities, workers must understand and articulate the skills they possess while also understanding what new skills are required to be successful and competitive in the current job market to support more advanced careers. Without this knowledge, professionals and workers do not have a clear pathway for career advancement and economic growth. We see the potential for a twofold solution here using data-driven approaches.

¹⁷ Afua Bruce, DataKind's former Chief Program Officer, contributed to this report while at DataKind: https://ash.harvard.edu/files/ash/files/colab-hks_5-6-2020_1_1.pdf

- The first is to aggregate job postings from different job boards so as to make it easy for job seekers to see available opportunities.
- The second is to use the information from those openings to identify the most frequently identified skills and opportunities so that organizations supporting employment and job retraining have an understanding of the marketplace and tailor their job training programs accordingly.

We see a smaller, related issue in providing support using data-driven approaches to application and resume completion. Individuals entering the job market, or retraining and seeking jobs in new markets, are often hindered by a lack of professional collateral. Building smarter data tools for job-seekers and the organizations serving them is an essential component of economic empowerment initiatives.

Systems map visualization

Based on the systems map of employment and retraining presented below, one's access to employment and skills building opportunities to remain competitive in an ever-changing market is a critical step in achieving economic empowerment. Leveraging data science to unblock one of the key pathways to accessing these opportunities-- such as understanding local employer demand and the skills required for those roles, or having the ability to find and access courses to build skills in highest demand to remain competitive-- is key to ensuring individuals have access to stable jobs and a career ladder that will support their economic growth. Data science, AI, and machine learning can work to influence one of the key systems points illustrated below to unlock these career advancement opportunities for long term advancement.

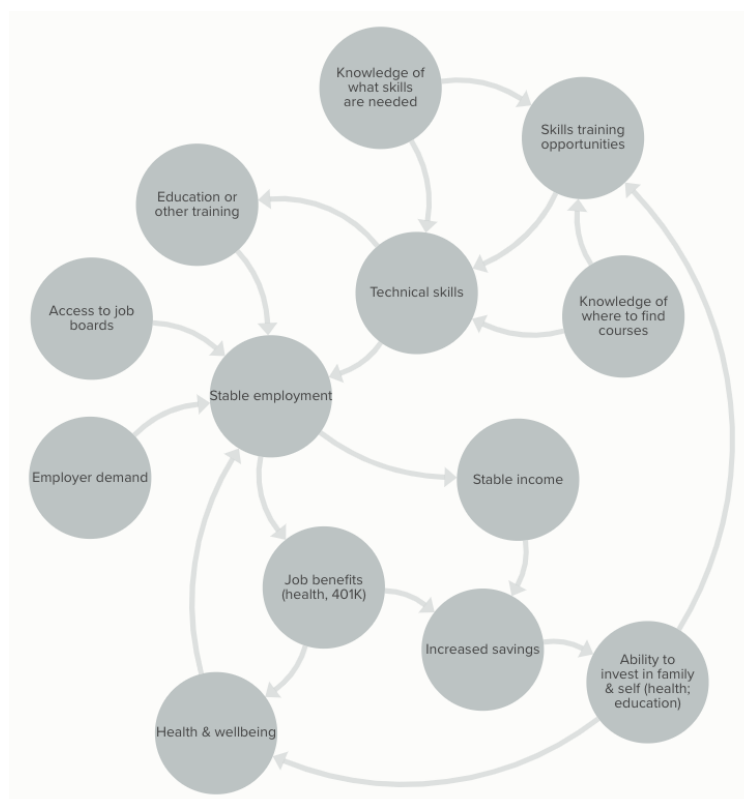


Figure 6: Systems Map of employment & job retraining

[For an interactive map, [click here](#)]

Summary of data science techniques

Throughout this theme, we see approaches ranging from building job aggregators for target markets to completing analytics on text data and doing smart job matching.

While job aggregators are not a new technology - sites such as Indeed.com, LinkedIn, and others have been around since the early 2000s - the software is typically designed with recruiters and job seekers in mind. As a result, software costs for those who want to run a job board create a barrier to entry for nonprofits or other social impact organizations who want to support the specific community they are engaging with. We see potential in deploying data science capabilities toward building a job aggregator utilizing an open-source, open-access code structure with a foundation built in Python. We specifically see the potential of packages [BeautifulSoup and Selenium](#), low- or no-cost API solutions such as [Google Cloud Talent](#) and visualizations using [Plotly DASH](#). This flexible structure would allow for the addition of new job board sites.

We also see the potential to apply natural language processing techniques to conduct a periodic review and synthesis of the common skills and requirements listed on active job listings. We see success in an approach¹⁸ successfully demonstrated on recruitment data in a public forum. We specifically look to apply feature detection for specific duties.

We highlight one example approach to demonstrate the relative sophistication of problems. For example, DataKind successfully deployed word2vec and GloVe approaches for information extraction from short-form textual information in the [colandrapp.com](#) project and the BLOC resume building¹⁹ project, which helps solve for the secondary goal of building tools to help with resume and application completion tooling. The insights provided from this approach would allow partners to facilitate more precise matching between jobseekers and employment opportunities aligned to their skillsets.

Readiness for DS/ML/AI assessment

While there is potential for high impact projects under the employment and job retraining theme, ability to execute on projects that leverage data science and AI largely depends on the organization's readiness and data maturity, notably access and availability of data within the proposed time frame. We also see some limitations in translated sector-specific experience between sectors and matching to specific jobs, which is a limitation of applying data science to common pain points encountered in this thematic area.

It furthermore appears that once knowledge around what skills would be most relevant for professional growth in a competitive environment, some of the problems identified revolve predominantly around, for example, training delivery and the ability to commit time to

¹⁸ <https://www.kdnuggets.com/2017/05/deep-learning-extract-knowledge-job-descriptions.html>

¹⁹ <https://www.datakind.org/blog/applying-ai-to-societal-challenges-in-us-cities-datakind-microsoft-launch-ai-accelerator>

skills-building opportunities. There is an interesting opportunity for developing tools for the marketplace, but that would be dependent on having a clearly-identified partner who would be willing to be an expert advisor and adopter of the tool, though they could be a data immature organization themselves.

Services that support mobility from poverty to stability



Across the United States, communities are challenged by a lack of information about how to connect people with the assistance they need. This has been most apparent during COVID-19, during which calls for help increased by 1000% in some places. Currently, communities, cities, counties, and states do not have a central hub with the ability to gather, analyze, and act on real-time data. The lack of consistent central data capacity impacts civic leaders and residents alike. Leaders need better data about gaps and inequities for regular planning and crisis response, and residents need better access to data to meet their immediate needs. The patchwork of indicators, databases, and intervention across local geographies make it challenging for end users to build the systems and proactively intervene, and state and national-level data masks hyper-local context and variability. As a result, funding decisions, policies, and resource allocation are based on incomplete, outdated information and anecdotal evidence.

Summary of issues identified

Matching resources to need

In order to best support individuals and communities in most need, organizations need to better understand where challenges are most acute, and what resources can provide the most opportunity towards stability and upwards mobility. This proves to be particularly challenging, as each geographical area presents a different profile of need. Creating hyper-localized profiles of need in each community where organizations work is time and effort consuming, most of which

is best used focused on outreach and distribution of critical resources. We recommend developing predictive models so that organizations supporting specific communities can target their services offerings to key areas of need and predict vulnerabilities for early response so that individuals and communities can become stable and resilient.

Systems map visualization

Based on the systems map of services that support mobility from poverty to stability presented below, one's access to critical social safety net services such as health, transportation, and affordable and nutritious food is critical to building economic stability. Leveraging data science to unblock one of the key pathways to accessing these services-- such as matching government or other community-based organization's ability to leverage data-driven models to match resource distribution to areas of highest need, or ensuring applications to these programs are available to individuals with different literacy levels or language skills-- is key to ensuring access to these sometimes life-saving and life-changing services.

There are multiple data-scienceable pathways to unlock this key leverage point, thanks to which individuals can, for example, increase their health and well-being through access to health services or affordable and nutritious foods, or even seek stable employment with child care or readily-available transportations services.

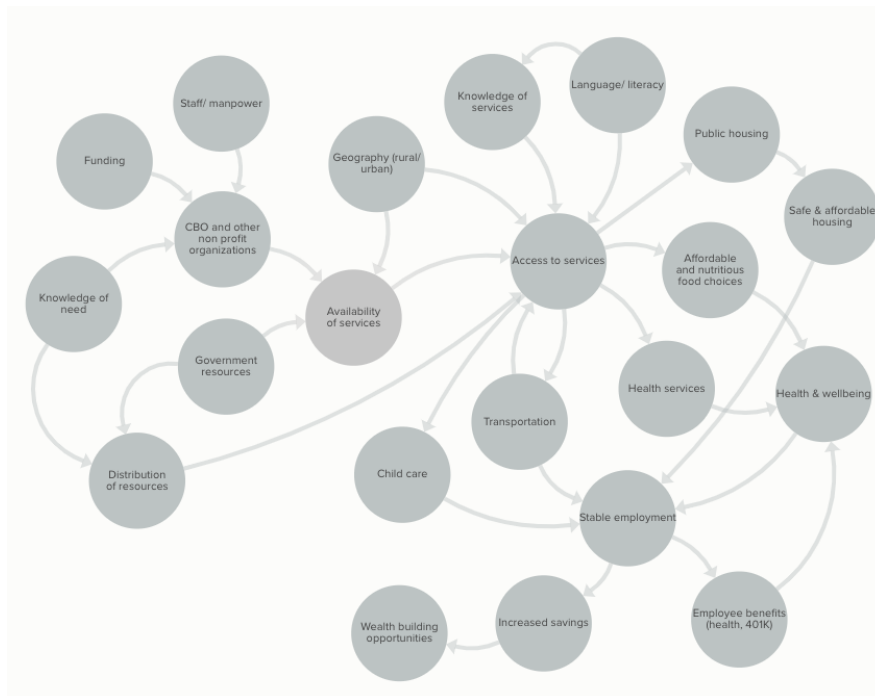


Figure 7: Systems Map of services that support mobility from poverty to stability

[For an interactive map, [click here](#)]

Summary of data science techniques identified

We see an opportunity for a variety of machine learning models to be applied to topics ranging

from social service personalization, efficient triage of social service requests, increasing transportation effectiveness, and helping match-and-track individuals across government assistance forms to minimize the amount of effort required to apply for services. DataKind could support organizations in documenting, designing and prototyping their data ecosystem (entity relationships, data models and definitions, storage, extract-load-transform processes, dashboards, interpretation) to ensure that local data is collected and assessed at a regular interval. Creating machine-level systems to observe data is a key milestone toward using the collected data to reason through potential decisions and act using the acquired information. This data-driven approach will help improve individual care and sharpen partner's ability to drive measurable community outcomes.

Readiness for DS/ML/AI assessment

The ability to leverage data science to engage and impact this field is promising, as there are multiple areas explored that ladder up to national-level impact. Yet the opportunity to scale in this space is largely a function of the type of organization we partner with with national reach, rather than a common data set or set of variables collected across the thematic areas, irrespective of organization size or type.

Working with national-level organizations that have local-level chapters provides us with the ability to leverage data that is owned by these local organizations, or explore publicly available data that will accurately reflect the current needs and realities of the geographies we will aim to impact. As such, in this theme, we see the clearest pathway to scale for any solution to be working through existing networks and channels for large-scale impact.