



ACCION
INSIGHTS

Unlocking the Promise of (Big) Data to Promote Financial Inclusion

Accion Global
Advisory Solutions

Kathleen Yaworsky
Dwijo Goswami
Prateek Shrivastava

MARCH 2017

Contents

- 1.0 Executive Summary
 - 1.1 What's the Big Deal With Data?
 - 1.2 How Will This Guide Help Me?
 - 1.3 What's Stopping Financial Service Providers From Acting?
 - 1.4 What Are the Opportunities?
 - 1.5 How Do I Get Started?
- 2.0 The Evolving Role of Data in Financial Inclusion
 - 2.1 Where Are You on the Data Continuum Today?
 - 2.2 The Global Reach of Big Data for Financial Inclusion: Who's Doing What, Right Now?
 - 2.3 How Can Data Advance Financial Inclusion?
- 3.0 A Framework for Deriving Value From Data
 - 3.1 Step 1: Understand Your Data Goals
 - 3.2 Step 2: Establish a Prioritized Road Map of Objectives
 - 3.3 Step 3: Map Business Processes and Build a Data Inventory
 - 3.4 Step 4: Identify Gaps or Redundancies in Your Current Data Collection
 - 3.5 Step 5: Understand Your IT Systems Architecture
 - 3.6 Step 6: Revise Your Processes and Culture
 - 3.6.1 Start With Clear Communication
 - 3.6.2 Align Incentives and Build Agility
 - 3.7 Step 7: Don't Go Big Before You Are Ready
 - 3.7.1 Focus on the Fundamentals
 - 3.7.2 Hire the Right Talent
 - 3.8 Step 8: Plan a Data Governance Strategy and Stewardship
 - 3.9 Step 9: Roll Out Across the Organization
- 4.0 Conclusion
- 5.1 Annex 1
- 5.2 Annex 2

1.0

Executive Summary

1. WHAT'S THE BIG DEAL WITH DATA?

While access to financial services has steadily increased, three billion people are underserved by the world's financial sector, with limited or no access to quality savings, credit, insurance, and payments. Previous industry reports¹ have highlighted big data's incredible potential to create a financially inclusive world. The sheer volume of data produced today is staggering: every half hour, the internet – combined with internet-connected devices such as phones, computers, and home appliances that use electronic sensors to record and transmit usage information – will generate digital data equal to all of the written works in human history.² Facebook users alone share 30 billion pieces of content each month.³ This has key implications for the financial inclusion movement. A recent Omidyar report⁴ suggests that, in the world's six biggest emerging economies alone — China, Brazil, India, Mexico, Indonesia, and Turkey — big data has the potential to help between 325 million and 580 million people gain access to formal credit for the first time. **But it isn't just big data that can help extend high-quality financial services to those who are failed by the world's financial sector: data of all shapes and sizes can help financial service providers improve operations and accelerate financial inclusion.**

2. HOW WILL THIS GUIDE HELP ME?

Believing in data's potential to provide both social and economic benefits, Accion set out to conduct research with the goal to develop a practitioner's guide, building on the existing literature, to provide financial service providers (FSPs) with actionable insights and first steps on how to better leverage data for financial inclusion. Introducing data-driven innovations is a multi-dimensional challenge for any FSP, requiring the talents and coordination of teams across the organization. This guide sets out to help FSPs begin to address these complex challenges by providing a case for change and outlining the initial strategic and tactical steps an FSP can take to first assess and then enhance its data capability. This guide breaks down the multiple components required for an FSP to strengthen its data capabilities, while referencing the growing number of specific use cases for data-driven innovations in fostering a financially inclusive world.

3. WHAT'S STOPPING FINANCIAL SERVICE PROVIDERS FROM ACTING?

An FSP's data is only useful insofar as its information can be mined and analyzed to extract meaningful insights. Traditional FSPs, such as microfinance institutions (MFIs) focused on financial inclusion, want to understand how to use data more effectively, but may not have the internal capacity to do so or might face resource constraints. Financial technology (Fintech) startups, likewise, often do not have the bandwidth or financial resources to pursue research about the broader impact of their innovative business models. This gap reveals the need to provide more 'proof points' to advance the industry's understanding and underscore the social and economic benefits of using data to promote inclusive finance.

Introducing data-driven innovations is a multidimensional challenge for any FSP, requiring the talents and coordination of teams across the organization.

4. WHAT ARE THE OPPORTUNITIES?

Among the early adopters of data-driven innovation for financial inclusion are firms that specialize in credit analytics using new digital or alternative data sources, such as mobile call data records, utility payments, social media activity, and others to assess risk and extend credit to “thin-file” or “credit-invisible”⁷ customers. Yet the potential benefits of leveraging data for financial inclusion extend far beyond alternative credit underwriting. Indeed, the smart analysis and application of data can improve every aspect of a business’s operations. McKinsey⁸ suggests several ways that businesses can harness the value of (big) data, including: creating transparency (e.g., a comprehensive view along the supply chain or operating process); enabling experimentation to discover needs, expose variability, and improve performance (such as through A/B testing); segmenting populations to customize actions (such as by creating more targeted product design or marketing); replacing or supporting human decision-making with automated algorithms; and innovating new business models, products, and services. The permutations of potential products, services, and use cases leveraging each of these sources of value are limited only by creativity.

5. HOW DO I GET STARTED?

Accion conducted a wide range of interviews with industry experts and practitioners and found that the size of datasets may not be a first-order concern for many FSPs working with the financially underserved. Maria Fernandez Vidal, Financial Sector Specialist with CGAP’s Digital Finance Frontiers Initiative, explained that FSPs working toward financial inclusion are often operating with hundreds or thousands of data points per user, not hundreds of thousands or millions (as might be the case with larger financial institutions).⁹ Accordingly, FSPs focused on financial inclusion face an immediate challenge in learning how to leverage “small data” before moving on to “big data”.

Most organizations’ readiness to leverage data falls along a continuum. Accion assesses an organization’s capacity using a “Data Continuum Diagnostic Framework”, summarised in Figure 1. The first step in building capacity is to assess where your organization currently stands. Basic analytics can yield productive insights. Benefits can be realized at any level of maturity and these increase as the FSP moves along the continuum. McKinsey has identified multiple levels of maturity in terms of capacity to leverage data. The most basic level is digitizing and structuring the data in preparation for further analysis. The next level of sophistication requires making the data available so that multiple datasets can be integrated to create more meaningful business insight. The highest level of maturity involves the application of advanced analytics, such as customized and automated algorithms and real-time data analysis. Together with the Data Continuum Diagnostic Framework, this assessment informs Accion’s recommendations and helps FSPs evolve along the spectrum.

1. https://www.omidyar.com/sites/default/files/file_archive/insights/Big Data, Small Credit Report 2015/BDSC_Digital Final_RV.pdf;
<http://www.cgap.org/sites/default/files/Focus-Note-The-Potential-of-Digital-Data-Jan-2015.pdf>
 2. <https://www.cgap.org/publications/potential-digital-data>
 3. <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>
 4. https://www.omidyar.com/sites/default/files/file_archive/insights/Big Data, Small Credit Report 2015/BDSC_Digital Final_RV.pdf
 5. <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>
 6. <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>
 7. “Thin-file” refers to customers with limited credit history; “credit invisible” or “no-file” refers to customers lacking any formal credit history.
 8. <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>
 9. Interview with Maria Fernandez Vidal, August 10, 2016.

This guide breaks down the multiple components required for an FSP to strengthen its data capabilities, while referencing the growing number of specific use cases for data-driven innovation in fostering a financially inclusive world.

FIGURE 1. DATA CONTINUUM DIAGNOSTIC FRAMEWORK

	PEOPLE	SYSTEMS	STRATEGY
Early	Lack of systematic data analytics function Ad hoc ability to answer specified data queries	Most or all data is collected and stored manually Some processes in place for digitizing data Digitized data is used for making some business decisions	Building initial capacity for using data Awareness of importance, but unclear which direction to take
Growing	Full-time resources dedicated to data analysis Staff capable of posing and answering data queries for business decisions	A post-facto management information system (MIS) produces periodic, static reports A real-time MIS and business intelligence reporting systems enable some automated business decisions	Strengthening existing business decision processes Building data capability to take business decisions that cannot be taken today
Mature	Business team actively collaborates with data analysts/teams Cross-vertical experts can comfortably work within both business and analytics teams	Nearly all control points in the business workflow are digitized and generating data Internal data systems are integrated with stable external data feeds Nearly all business decisions are automated	Systematically adapting processes to facilitate data-driven decision-making Building a comparative advantage to gain information unavailable to industry competitors

The Evolving Role of Data in Financial Inclusion

BIG DATA, MACHINE LEARNING ALGORITHMS, ARTIFICIAL INTELLIGENCE

These terms dominate today's fintech lexicon. In the 1980s, the financial sector was responsible for pioneering the development of advanced analytical and business intelligence tools for commercial purposes.¹⁰

As business intelligence tools grew in sophistication and capability, organizations beyond the financial sector began to look for ways to better aggregate and store all their operational data on one platform. Walmart was one of the first large retail corporations to build a data warehouse (a platform that aggregates all available data generated by an organization under one roof for analysis). It started as an accident: as the story goes, one of the retailer's computer operators, tired of retrieving archived historical sales data, secretly "borrowed" excess storage space on a company server, where he downloaded and stored the data from the most-requested archives. When Walmart managers found out, they quickly realized the value of timely and widespread access to data and soon transactions from thousands of retail locations were available for analysis in the data warehouse within minutes, enabling near real-time reaction to sales and supply data. Walmart eventually became the first commercial data warehouse to collect a terabyte of information in 1992.¹¹

As the cost of data storage decreases and the processing power to analyze the data and transform it into meaningful information increases,¹² data itself has become an increasingly strategic asset, across all industry sectors, both in terms of ownership or access and the ability to leverage its insights for business decisions.

Coming full circle back to the financial sector, a 2016 report from the Institute of International Finance, *The Business of Financial Inclusion: Insights from Banks in Emerging Markets*, outlines how the Commercial Bank of Africa (CBA) had access to millions of data points from mobile banking phone use to build and refine the M-Shwari credit scoring algorithm.¹³ For smaller FSPs focused on financial inclusion that do not command a larger commercial bank's resources, that scale can appear overwhelming. In the Center for Financial Services Innovation's 2016 Banana Skins report, leading industry experts and practitioners cited the development and implementation of a strategy in the face of rapid technological change as one of the biggest challenges facing the financial inclusion industry.¹⁴

Figure 2 decodes the various characteristics of data that are often conflated when people use the overarching term "Big Data."

10. Trippi, Robert R., and Efraim Turban. *Neural networks in finance and investing: Using artificial intelligence to improve real world performance*. McGraw-Hill, Inc., 1992.11. <https://www.cgap.org/publications/potential-digital-data>
11. <https://www.healthcatalyst.com/wal-mart-birth-of-data-warehouse/>
12. Often referred to as Moore's Law, after a 1965 observation made by Gordon Moore, in which he extrapolated that computing would dramatically increase in power, and decrease in relative cost, at an exponential pace. <http://www.intel.com/content/www/us/en/silicon-innovations/moores-law-technology.html>
13. http://www.centerforfinancialinclusion.org/storage/documents/IIF_CFI_Report_FINAL.pdf
14. https://static1.squarespace.com/static/54d620fce4b049bf4cd5be9b/t/57989070d2b8575b1b001466/1469616256985/Banana+Skins_07-16_v8.pdf

FIGURE 2. DEFINING DATA

Data	Facts and statistics collected for reference or analysis.
Analog Data	Data that is represented in a physical way or stored in physical media (e.g., the surface grooves on a vinyl record, magnetic tape on a VCR cassette, etc.); may also be known as organic data or real-world data. ¹⁵ This term as such is not used in financial inclusion, but could refer to the information a loan officer writes down on a paper loan application form.
Digital Data	In the context of financial inclusion, digital data is often used to refer simply to data stored in an electronic format, or data that has been "digitized", i.e., previously stored in a physical format but now available electronically. More broadly, digital data seeks to capture elements of the physical world and simulate them for technological use, for example by storing complex audio, video, or text information in a series of binary characters, traditionally ones and zeros, or "on" and "off" values. ¹⁶
Traditional or Alternative Data	In the context of financial inclusion, alternative data refers to data acquired from what banks consider non-traditional sources (such as utility bill payments, social data, or call data records) that can then be used by a financial institution to assess a potential borrower. Traditional data sources contain financial information available in a credit report or a customer's banking records. ¹⁷ Since its inception, microfinance has relied on what could be considered "alternative" data – in the sense that microfinance customers often do not have a credit report or bank account – and the archetypical microfinance credit methodology requires a loan officer to reconstruct a customer's cash flow statements and gauge willingness and ability to pay based on an in-person assessment of a customer's home and business environment and reputation in the community. However, microfinance institutions do not typically consider this type of information (one that they have been collecting for decades) as alternative data. Yet it can be a rich source of analysis, especially once digitized. In fact, this paper recommends that financial institutions look first at this internal data already at their disposal, before seeking alternative data from external sources.
Big Data	Data of a very large size, typically to the extent that its manipulation and management present significant logistical challenges. ¹⁸ As McKinsey pointed out in a 2011 report, the term big data "is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data." ¹⁹ Few definitions establish a minimum threshold. Gartner's information technology (IT) dictionary definition of big data emphasizes complexity over size, with what it refers to as the three V's: high-volume, high-velocity, and/or high-variety information assets that demand cost-effective, innovative forms of information processing, which enable enhanced insight, decision-making, and process automation. The term big data is often used interchangeably with alternative data in the context of fintech and financial inclusion.

15. <https://www.techopedia.com/definition/24871/analog-data>
16. <https://www.techopedia.com/definition/24872/digital-data>
17. <http://lending-times.com/2016/06/20/over-12000-variables-of-alternative-data-for-36-billion-in-alternative-loans/>
18. <https://whatsthebigdata.com/2013/06/15/big-data-arrives-at-the-oxford-english-dictionary/>
19. <http://www.gartner.com/it-glossary/big-data/>

Organizations need to be fully aware of their own potential to generate meaningful data before looking elsewhere. As Paulo Marques of Capgemini put it, “Data is a journey where you’re heading towards the unknown. So, you have to start with the known.”

WHERE ARE YOU ON THE DATA CONTINUUM TODAY?

Any organization can begin its data journey by asking two simple questions: What is the process for making critical business decisions today? and, How would this process change if more data were available, or if the data currently available were better organized? Understanding these gaps is the first step towards developing a data strategy. The high-level framework in Figure 1 can serve as a self-diagnostic tool for an organization to assess where it stands and to chart a course for the future.

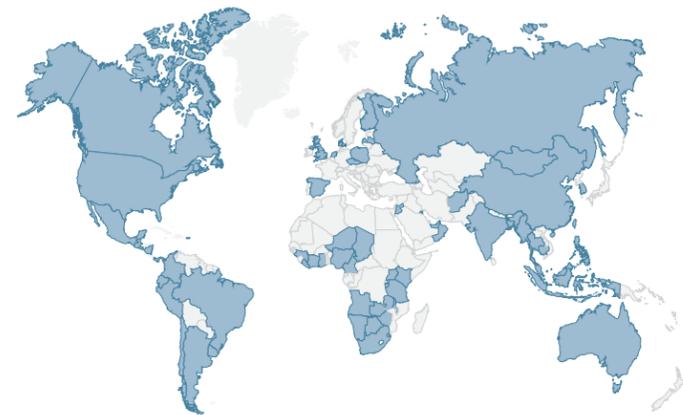
It is essential to perform a clear-eyed assessment of existing data capabilities and resources. Many FSPs focused on financial inclusion in emerging markets do not have the necessary skills and budget to support strategic projects using data. Ironically, it is due to these very constraints that these organizations can most benefit from powerful data analytics. Counterbalancing the common lack of human or financial resources, many FSPs underestimate the value of the data they already have. A recurrent theme in conversations with experts was the need for organizations to be fully aware of their own potential to generate meaningful data before looking elsewhere. As Paulo Marques of Capgemini put it, “Data is a journey where you’re heading towards the unknown. So, you have to start with the known.”²⁰

20. Interview with Paulo Marques, August 17, 2016.

21. <http://www.cgap.org/sites/default/files/Focus-Note-The-Potential-of-Digital-Data-Jan-2015.pdf>

22. As there is no standard definition for big data applied across these organizations, we have listed organizations that self-declared their use of big data, or are explicitly using alternative data sources to expand access to financial services. Peer-to-peer (P2P) lending platforms are not included in these examples.

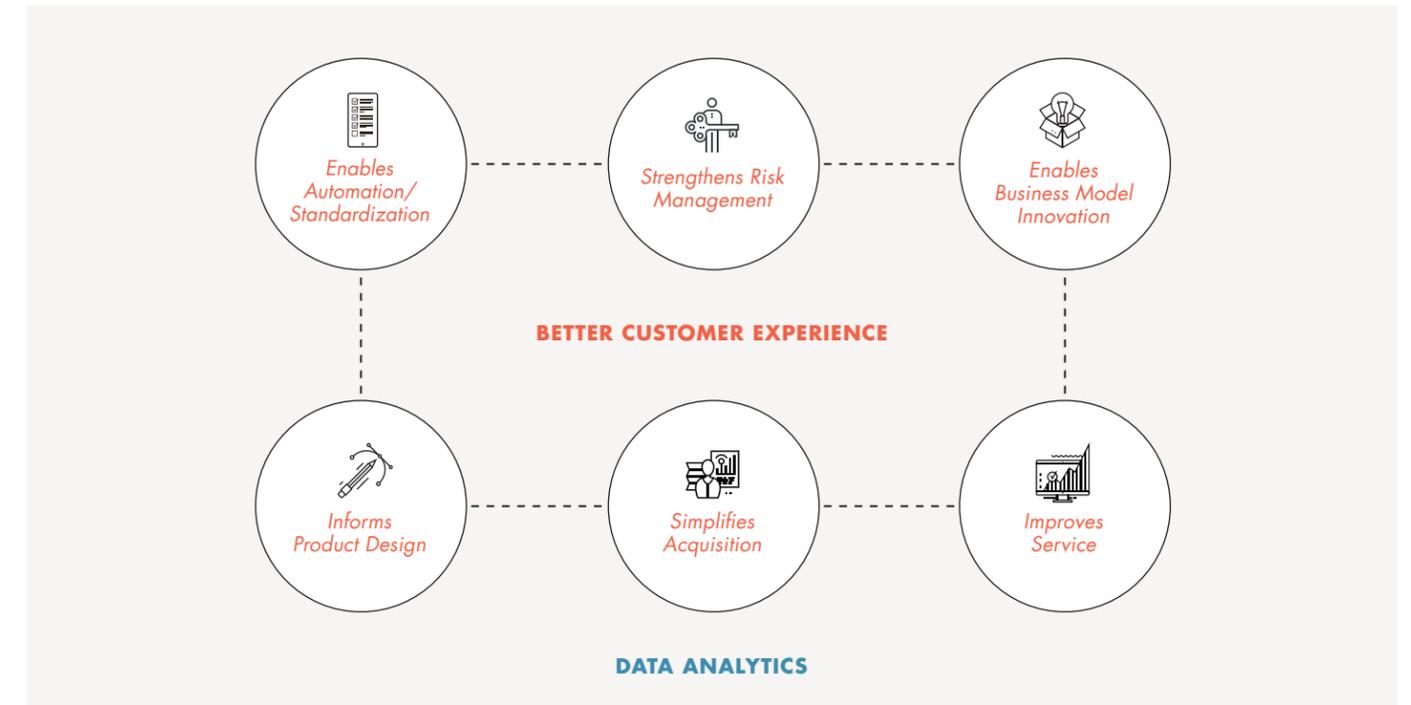
FIGURE 3. MARKETS WHERE FSPS ARE LEVERAGING ADVANCED ANALYTICS FOR FINANCIAL INCLUSION²²



THE GLOBAL REACH OF BIG DATA FOR FINANCIAL INCLUSION: WHO’S DOING WHAT, RIGHT NOW?

Figure 3 represents the geographic footprint of FSPs reportedly leveraging big data in their core business model. An annex to this report lists these organizations, and offers a brief description of how each organization is using data. This list builds on previous work by CGAP, and is not meant to be an exhaustive review of the global industry.²¹ It is, however, meant to illustrate how the use of digital data and advanced analytics for financial inclusion has rapidly expanded in both mature and emerging markets and capture some of the diversity in business models and use cases of big data for financial inclusion.

FIGURE 4. DATA-DRIVEN INNOVATION IN FINANCIAL SERVICE PROVISION



HOW CAN DATA ADVANCE FINANCIAL INCLUSION?

The rapid expansion of mobile services and access to the internet, globally, has opened new frontiers for understanding customer behavior. Simultaneously, while the amount of data available has increased and its sources have proliferated, the analytical tools for making sense of this data have also become more sophisticated thanks to, among other things, advances in machine learning and artificial intelligence. FSPs are consequently trying to strategically position themselves to take advantage of the new digital footprints generated by their customers to target their customer base more effectively than their competitors.^{23 24} All else being equal, the organization with differential access to data is more likely to succeed.^{25 26} In line with our experience, CGAP²⁷ notes that, once digitized, an organization’s data can benefit FSPs through applications in three broad areas: finding new customers, deepening customer relationships, and

managing risks. These forces have led to a proliferation of data service providers and partnerships to advance financial inclusion, many of which are still early-stage innovations that have yet to evolve into mature financial products.

Figure 4 illustrates a few examples of business processes that an FSP can improve by applying more automated and data-driven decision-making. Data analytics is a foundation upon which to build stronger risk management, enable automation or standardization of processes and innovation of business models, simplify acquisition, inform product design, and improve service, all of which ultimately lead to a better customer experience. Each of these pieces is interrelated and mutually-reinforcing, addressing various inefficiencies or pain points in traditional business processes and customer interaction.

23. <https://www.microfinancegateway.org/library/mobile-phone-data-key-promoting-financial-inclusion>

24. <http://mastercardcenter.org/insights/increasing-financial-inclusion-mobile-phones/>

25. <https://hbr.org/2012/10/big-data-the-management-revolution/ar>

26. <https://www.pwc.es/es/publicaciones/tecnologia/assets/Seizing-The-Information-Advantage.pdf>

27. <http://www.cgap.org/sites/default/files/Focus-Note-The-Potential-of-Digital-Data-Jan-2015.pdf>

BOX 1. EXAMPLES OF SUCCESSFUL DATA-DRIVEN INNOVATION IN FINANCIAL INCLUSION

Many of the examples cited here could be easily classified under multiple headings, which underscores the interrelated and mutually-reinforcing benefits of data-driven innovation outlined in Figure 4.

CUSTOMER ACQUISITION

MiMoni, an online lender in Mexico that provides instant credit decisions, worked with Kenshoo, a social media analytics company, to identify its most responsive audience and optimize its social media strategy around that segment's activity. Leveraging Kenshoo's big data analytics for MiMoni's Facebook advertising campaigns led to a 60 percent decrease in its customer acquisition costs.³⁰

An identity verification service provided by **Lenddo** leverages Facebook and Android social media data to verify customer identity within 5 seconds.³¹

DemystData provides customer verification services to financial institutions by piecing together and analyzing scattered data points on identity, employment, income, etc., reducing the need for manual checks. DemystData reports that 15 percent of customers who would have been rejected due to incomplete information can actually meet requirements for loan approval, enabling its U.S.-based microlender clients to reach new customers.³²

MARKETING

In a partnership between Airtel and HFC bank in Ghana, **Cignifi** analyzed detailed call records of both partners to understand the call patterns of low- and high-net worth individuals. Cignifi found that SMS use could be a potential proxy for discretionary income, while simultaneously serving as an effective marketing channel to reach higher value customers.⁴¹

28. <http://blogs.accion.org/fin-tech/new-report-msmes-small-business-credit-gap/>

29. Interview with Jonathan Whittle, August 9, 2016.

30. <http://kenshoo.com/mimoni-case-study/>

31. <https://www.lenddo.com/pdfs/Lenddo-Verification-Factsheet-2015.pdf>

32. <http://www.cgap.org/blog/big-data-financial-inclusion-boring-better>

AUTOMATING PROCESSES

Digitizing workflows can improve efficiency and provide organizations with richer and more structured internal datasets for further analysis. In a case study previously documented by Accion, **Ujjivan**, an Indian microfinance institution, worked with **Artoo**, a digital CRM service provider, to digitize its loan application workflow. The project increased Ujjivan's loan officer productivity by 134 percent.³³ At the same time, Ujjivan used its new digital platform to perform real-time analysis of its loan cycles.

New digitally native entrants automate the application process from inception through online lending platforms, such as **Konfio** in Mexico or **Lulalend** in South Africa.³⁴

PRODUCT DEVELOPMENT

The emergence of new digital channels for customer engagement and financial service provision also enables a secondary service layer of business-to-business (B2B) innovation. In Kenya, for example, **Kopo Kopo** built a new application for a previously untapped customer base on top of the existing M-PESA infrastructure. M-PESA began as a P2P remittance system; as its popularity increased, customers wanted to use it to make payments at merchants as well, but merchants faced limitations using their personal accounts for business. As a result, M-PESA was seeing low adoption rates in the merchant business (less than 0.01 percent of the total market).⁴² Kopo Kopo developed a business management platform for merchants that allowed them to automate transaction entries by triggering payments from M-PESA. Within three years of launch, Kopo Kopo had built a product that was generating transactions worth US\$3 million a month.⁴¹

CUSTOMER SEGMENTATION

Data-driven innovation allows organizations to learn more about their customers and offer more direct interaction, delivering relevant content and personalized products. **Juntos Finanzas** provides a real-time messaging service that analyzes customer responses to gain insights on and influence customer behavior. It aims to build customer trust by structuring guided conversations via text. The Juntos team spends the first few months of every client engagement rapidly testing and iterating its messaging to ensure the platform adapts to the local context. Pilot projects have led to an increase in customer savings balances, a reduction in transaction costs at branches, and better customer segmentation. According to the founder, Ben Knelman, one of the greatest value additions of Juntos has been the insights it gains into customer needs and behavior that were previously difficult for banks to gauge through traditional customer interactions.³⁹

CUSTOMER ENGAGEMENT

RevolutionCredit uses behavioral analytics to complement underwriting processes. It helps banks identify customers that will perform better than their credit score would indicate by offering customers a series of educational videos, quizzes, and games designed to increase financial capability, and then tracking how customers engage with the content.⁴⁰

33. https://www.accion.org/sites/default/files/ujjivan_dfa_study.pdf

34. <http://blogs.accion.org/fin-tech/new-report-msmes-small-business-credit-gap/>

35. Accion internal analysis

36. Interview with Aneesh Varma, June 22, 2016.

37. <http://aire.io/partners/>

38. <https://www.ft.com/content/6f5453d6-1b69-11e6-8fa5-44094f6d9c46#axzz4ADUi6riM>

39. <http://www.cgap.org/sites/default/files/Working-Paper-Juntos-Finanzas-A-Case-Study-Oct-2015.pdf>

40. Interview with Ben Knelman, August 3, 2016.

41. <http://www.savings-banks.com/SiteCollectionDocuments/Cignifi.pdf>

42. <http://www.gsma.com/mobilefordevelopment/programme/m4d-impact/kopo-kopo>

CREDIT UNDERWRITING

In Accion's experience, leveraging client data points for credit scoring has helped increase productivity in several institutions globally. One MFI operating in the Caribbean responded to low productivity by building a streamlined credit scoring mechanism based on client socio-demographic data and other characteristics. Introducing a faster, data-driven process for credit evaluation helped this MFI double its loan officer productivity over one year, using a simple Excel-based system.³⁵

Aire integrates a 3-minute virtual interview into lenders' online loan application forms that assesses the applicant's character and ability to pay, leveraging insights from research in behavioral economics. FSPs use the scores from these virtual interviews to extend credit to individuals with little or no formal credit history. On average, Aire has increased loan disbursement by 14 percent.^{36 37}

InVenture asks customers to download its mobile application, which scans the customer's phone and uses that data to provide a credit score within 20 seconds.³⁸

BUSINESS MODELS

NeoGrowth, an SME lender in India, is leveraging business transaction history to appraise small enterprises that do not traditionally have access to formal credit. They have systematically built partnerships with major banks and developed proprietary software that integrates with the point-of-sale machines of such enterprises to build a unique transaction database that gives them insight into small enterprises, which have typically been disregarded by traditional lenders.⁴²

Alibaba's Ant Financial represents one example of how e-commerce players and others leverage transaction data to offer tailored financial services for the small businesses on their platforms.

Currently, the most common market applications of data for financial inclusion are business models to provide insights on customers with little or no formal credit history. In addition to leveraging alternative data sources to develop new ways of scoring more customers, innovations in underwriting and digital lending platforms add value by enabling automated decision-making, which can increase efficiency and reduce acquisition costs.

But the potential benefits of leveraging data for financial inclusion extend to every aspect of business operations. On the preceding page, Box 1 provides examples of data-driven innovation in several of these areas. Jonathan Whittle of Quona Capital commented: "FSPs need to pay attention to innovations by any and all organizations serving their target customer base, even if they operate outside the financial sector." Alibaba's Ant Financial represents one example of how e-commerce players and others leverage transaction data to offer tailored financial services for the small businesses on their platforms. The business process improvements enabled by advanced digital data analytics can create links across traditionally distinct sectors, in addition to internal links across business verticals. At the same time, this means competition can come from all sides.

3.0

A Framework for Deriving Value from Data

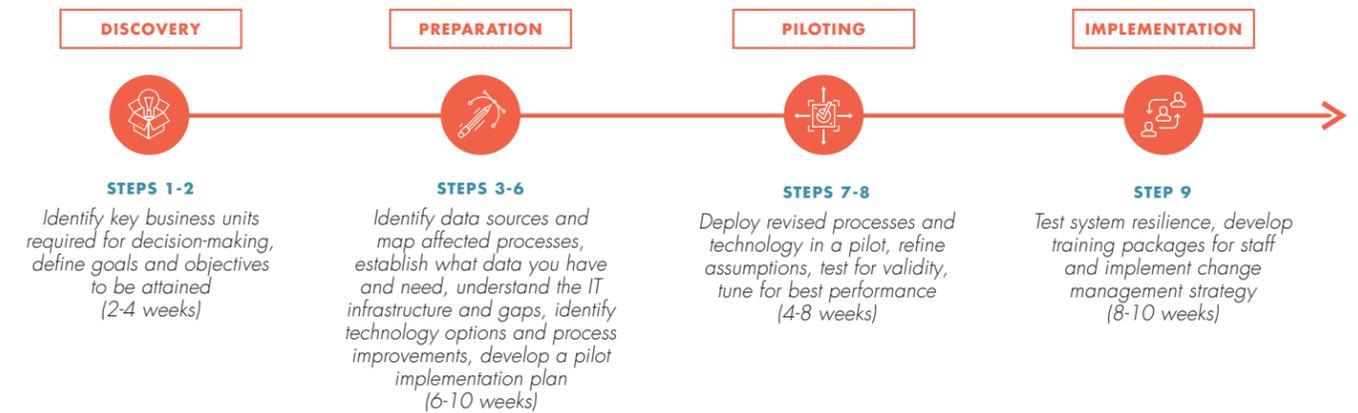
OPPORTUNITIES TO IMPROVE BUSINESS MODELS

Common to these examples is the organizations' ability to derive value from data, regardless of its source, scale, or complexity. As mentioned, however, opportunities to improve business models do not arise from data analytics alone. Building the operational capacity to respond to data is equally, if not more, important for success.

A data strategy and implementation plan will vary depending on several factors particular to any given organization, such as its goals, existing technology infrastructure, the level of control it can exert on processes and systems, and its appetite for investment. Like many product planning and design processes, developing and implementing a data strategy demands appropriate time and consideration in four phases: discovery, preparation, piloting, and implementation. Figure 5 outlines these

phases and the following sections describe the steps required for each in more detail. The goal is to develop a data strategy and implementation plan. Of course, not all FSPs will be ready or willing to do this in-house, but an FSP should address at least some of these questions internally before passing them on to a third party. A sense of organizational ownership and responsibility is a prerequisite for success.

FIGURE 5. FRAMEWORK FOR DERIVING VALUE FROM DATA



Alex Holderness, Senior Vice President and Head of Digital at Accion in the U.S., finds it helpful to frame process-mapping conversations with staff as if you were the database. This facilitates taking a systems perspective: ask, for example, what is that information used for and how does it flow through the system?



STEP 1: UNDERSTAND YOUR DATA GOALS

Data's best, most precise applications are for predictable business outcomes and, as such, enterprises should use data to evaluate core business goals. When implementing a data project, an organization should start by asking "Why are we doing this? What business goals can be better evaluated or advanced through data usage and analysis?" That goal could be, for example, increasing productivity or profitability. It should reflect some concrete business aim. Research from the Harvard Business School shows that businesses that use data to make decisions are more productive and profitable.⁴³

Given the strategic nature of data and its long-term impact on the business, it is imperative to consider cross-cutting organizational goals (e.g., increased customer retention through improved product design, marketing, delivery channels, customer service, etc.), in addition to goals specific to individual business verticals. Regardless of the scope of the project, the exercise will require coordination and collaboration among all stakeholders of the business. A good practice is to establish a working group made up of representatives of each business function or department (including upper management and others involved in setting the strategic direction of the organization) to see how their data management issues are interrelated. Jonathan Whittle of Quona Capital observed that data drives efficiencies across an organization, breaking down communication barriers and departmental silos, and enables a comprehensive big-picture view across business verticals.⁴⁴

43 <https://hbr.org/2012/10/big-data-the-management-revolution>

44 Interview with Jonathan Whittle, August 9, 2016.



STEP 2: ESTABLISH A PRIORITIZED ROAD MAP OF OBJECTIVES

Using the organization's goals as a starting point, next develop a series of prioritized objectives or initiatives that can be implemented within a period of 9-12 months. These should be prioritized based on their relative strategic importance and ease of implementation (given your organization's resources, capacity, and current location on the data continuum depicted in Figure 1) to create a project road map agreed on by all stakeholders. It is important to map out all objectives to clearly identify any dependencies and common data sets.

In the short term, an organization's data strategy and capabilities should address immediate needs to inform business decisions that require urgent analytical support. In the long term, organizations should build a strategy to acquire a unique information advantage.



STEP 3: MAP BUSINESS PROCESSES AND BUILD A DATA INVENTORY

Identify what existing data you have to help meet the prioritized objectives and document where these data are generated and stored to create a "data inventory." During a regular business operation, such as a loan application, staff involved at each stage of the process often collect and use data in disparate or disconnected systems. These flows and intersections need to be understood and reconciled against the written policies and procedures to discover how people actually work and make decisions. This is not an audit to uncover noncompliance, but rather a process-mapping exercise to identify pain points and areas of improvement. What are the data points needed to complete the operation, how does that information enter the system, where is it stored, and how is the decision ultimately made?

Mapping the entire business workflow can help answer these questions. Alex Holderness, Senior Vice President and Head of Digital at Accion in the U.S., finds it helpful to frame process-mapping conversations with staff as if you were the database. This facilitates taking a systems perspective: ask, for example, what is that information used for and how does it flow through the system? How does that information inform any decision? If we automate it, how do we get the system to learn what's important?⁴⁵ Every human assumption needs to be picked apart to get at the logic underneath if the process is to be successfully analyzed and automated.

45. Interview with Alex Holderness, August 29, 2016.

46. Ibid.

47. Interview with Paul Makin, August 11, 2016.

48. Interview with Johann Bezoudienhoudt, August 9, 2016.

Paul Makin, co-founder of M-PESA, believes organizations should first allow their business team to play around with available data on tools such as Excel, simply "looking for interesting patterns." The questions they begin to ask from the data should be used to inform the goals of any data project.

Part of the mapping process should involve asking the key decision-makers at each step to describe what they do on a daily basis (this can be accomplished through a few hours of interviews and job shadowing). Analysis should look for hedging language and 'soft' verbs that imply human value judgments, such as 'might,' 'can,' 'depending on the following,' etc., to identify the decision-making rule versus how exceptions are made. Good questions to ask at this stage include: "Can I rely on all these fields being completed all the time? If, for example, this field is left blank, does the process stall? What is required for the process to move forward?"⁴⁶

How to get started? Paul Makin, co-founder of M-PESA, believes organizations should first allow their business team to play around with available data on tools such as Excel, simply "looking for interesting patterns." The questions they begin to ask from the data should be used to inform the goals of any data analytics project. This workflow enables business operations teams to gain a better understanding of the data available on their systems and ensures business insights guide data analytics, rather than the other way around.⁴⁷ Johann Bezoudienhoudt of Bankable Frontier Associates emphasized this point by stating there is a large difference in the value one can extract from data if it is looked at by someone with deep operational insights, as opposed to individuals who are only technically competent in statistical analysis. He asserted that individuals with a deep understanding of business operations are critical to a successful analytics function.⁴⁸

After process-mapping, you should have a clear picture of your “data inventory,” which can be used to identify redundancies or gaps in your current data collection. This is the foundation from which to build your data strategy.

As you begin, assess how comfortable your business operations team is with handling and interpreting its departmental data. The data continuum of Figure 1 categorizes organizations with individuals who can comfortably operate within both the business and data verticals as having ‘mature’ data capabilities. Selorm Adadevoh, Chief Executive Officer of Digicel Haiti, built this capacity in his team by sharing his decision-making framework.⁴⁹ This helped the team understand what kind of information he needed before making a business decision. Over time, the team incorporated this framework into their thinking, allowing it to seamlessly support the CEO’s decision-making.

In summary, look at the key processes required to arrive at the decision to support the goals and objectives. Map the data points, the location of each of those data points, and permissions required to access those data points (by role) in the organization for each step of the process. This data-process-mapping will result in a clear picture of your “data inventory,” which can be used to identify redundancies or gaps in your current data collection. This is the foundation from which to build your data strategy.

49. Interview with Selorm Adadevoh, August 12, 2016.

STEP 4: IDENTIFY GAPS OR REDUNDANCIES IN YOUR CURRENT DATA COLLECTION

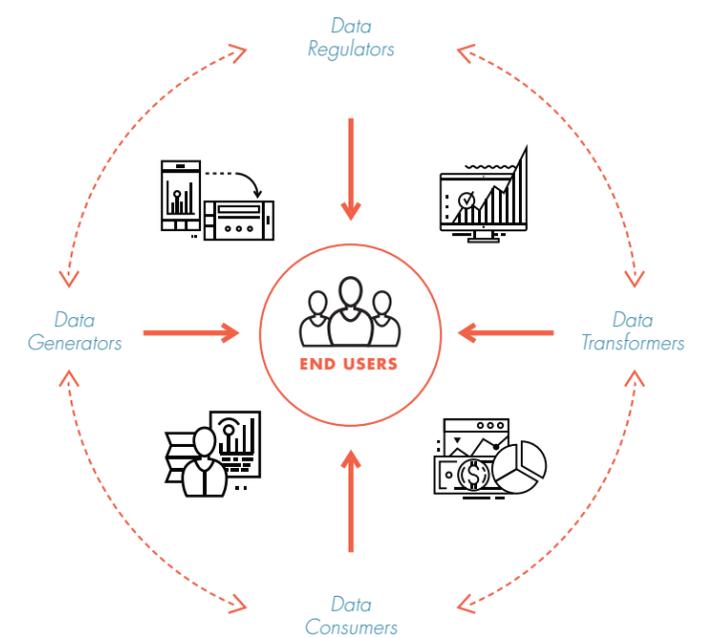
Using the process map and data inventory as working documents, have the established task force 1) identify areas for improvement in the processes, and 2) create a wish-list of data points that, if available, would make decision-making and responsiveness to the customer more efficient. Include in this exercise any current third-party data sources, such as credit bureaus, tax return information, etc., and if applicable, consider additional sources of “alternative” data that you may want to integrate, whether now or in the future. During this workshop, you may want to ask, “What information would you like to have, if cost were not an issue?” In addition to building capability to better leverage internal data, FSPs should consider whether to integrate supplementary data sources, including social media or government datasets, bill payment history, or call data records. Consider the potential sources of that data and the feasibility of adding them to your data inventory, in terms of both cost of acquisition and ease of integration.

Figure 6 depicts the data ecosystem, illustrating how the actors involved in the life cycle and value chain of any given data point – from its creation to its analysis and use for other business purposes – can be understood as overlapping entities. For example, an FSP “generates” data on the financial performance of its customers that it can “transform” by applying advanced analytics to create business insights and “consume” by applying those insights for credit underwriting, customer engagement, or any of the applications discussed above.

50. Interview with Miguel Herrera, August 4, 2016.

Miguel Herrera of Quona Capital believes that a major strength of successful fintech operations is their ability to discover the minimum number of variables required for a business decision, and to “run with them”.

FIGURE 6. THE DATA ECOSYSTEM



The value of the data that is currently being generated or could potentially be captured can be ranked according to its relative importance in the decision-making process to help prioritize within the workflow. Prioritizing the most critical data points on current business insights can help focus energies in the most impactful areas. Miguel Herrera of Quona Capital believes that a major strength of successful fintech operations is their ability to discover the minimum number of variables required for a business decision, and to “run with them”. It is critical to question and re-question assumptions of why a given data point is necessary in order to distill this minimum viable data set.

In any technology project, success often hinges on the thoroughness of the requirements definition process...if the goals and requirements are unclear, the project is likely to fall short of expectations, or at the very least, encounter serious setbacks and delays.

STEP 5: UNDERSTAND YOUR IT SYSTEMS ARCHITECTURE

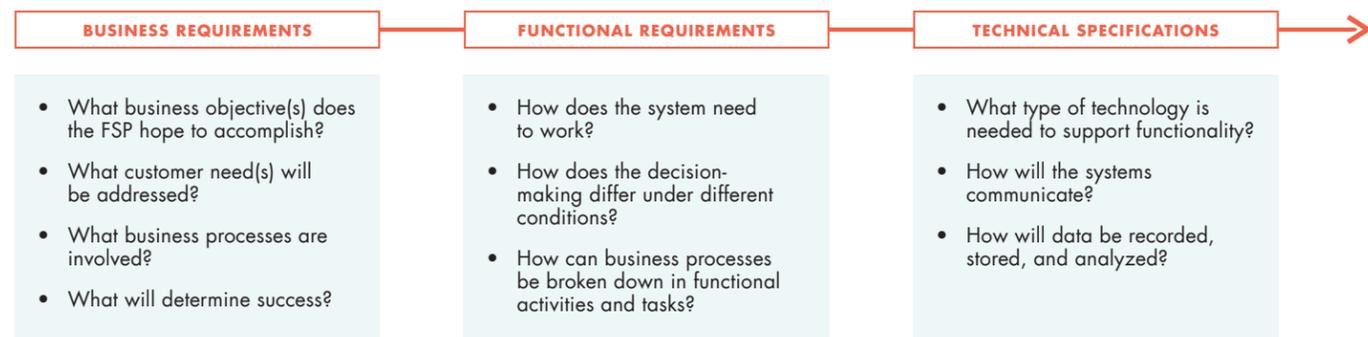
One of the key components of a data implementation program is to understand the technology infrastructure of the organization. What systems are in place? How are they operated? Is there resilience built into the systems and infrastructure (in terms of system sizing and disaster recovery, for example)? What is stored on individual computers and what is available in a shared environment? Based on this understanding, outline the technology architecture required to support the data implementation and plan to fill any gaps in the existing architecture by identifying the appropriate vendors and providers. Knowledge of your IT systems architecture also informs the technical specifications of vendor Requests for Proposals (RFPs), as described below.

Before signing a contract with a data service provider, organizations must take the time to think through what they want to achieve from such an engagement and what

it will take to get there. A technology vendor will likely do what is asked of them; if the goals and requirements are unclear, the project is likely to fall short of expectations, or at the very least, encounter serious setbacks and delays.

Figure 7 outlines some of the key questions to ask when translating business goals into functional and technical requirements to ensure that the RFP ultimately developed for vendor selection matches the institution’s needs and expectations. Any known constraints, assumptions, and dependencies must be clearly documented while describing each business requirement, which is why it is important for business units, rather than IT staff, to determine those business requirements as they better understand the business needs. The functional requirements should describe as much as possible the solution’s expected behavior and responses under different business conditions. The technical specifications then outline the technology needed to support the business and functional requirements.

FIGURE 7. BUSINESS REQUIREMENTS DOCUMENTATION



51. Bill Gates famously once said: “The first rule of any technology used in a business is that automation applied to an efficient operation will magnify the efficiency. The second is that automation applied to an inefficient operation will magnify the inefficiency.” <https://www.capgemini.com/blog/bpo-thought-process/2015/01/tempted-to-rewrite-bill-gates-rules-on-automation>
 52. Interview with Ganesh Rengaswamy, August 9, 2016.

Put simply, manual systems are more challenging to monitor and adjust rapidly in response to analytical insights and make it difficult to generate those insights in the first place. Digital data is easier to aggregate and analyze, which speeds response time.

It is important to note that the process is an iterative one. That is, there might be cases in which some envisioned functional requirement might not be possible or supported with the existing technology, or vice versa. For example, if it is difficult to ensure there are reliable or real-time data fields (such as when accessing public data sets, like credit bureau reports), then systems will need to be designed to manage the expected degree of uncertainty within the operational model. If access to credit bureau data is unreliable, then the loan processing workflow should be designed to collect all other available information independently, rather than halting activity until the applicant’s credit score can be confirmed. Developing the business case and requirements becomes much easier when organizations can articulate the expected output. In any technology project, success often hinges on the thoroughness of the requirements definition process.

Critical to the technology architecture is the security of the data, especially in a distributed or cloud-based environment. Ensure controlled access is in place based on agreed-upon roles and responsibilities. While building the security foundation of the systems architecture, you will need to ensure compliance with your country’s confidentiality laws. This is also important from a process point of view and it ensures that the staff is trained and aware of compliance requirements that may have implications for how they handle data.

BOX 2. DIGITIZE YOUR DATA, STRATEGICALLY AND SYSTEMATICALLY

The process-mapping exercise (along with the discovery and preparatory activities in Steps 1 – 5) forms the foundation for planning digitization efforts. Before beginning to digitize data and automate processes, it is important to ask if the data point adds value; otherwise you risk compounding inefficiency.⁵¹ In a loan application form, for example, data fields that are frequently left blank without affecting the ultimate business decision should be removed from the existing process to streamline the ultimate outcome, rather than automate the collection of irrelevant data.

Complete digitization of data and automation of processes is a long-term, later-stage goal for more mature data organizations (see Figure 1). Determine if your organization is collecting and storing any operational data manually. Does your organization have a long-term plan for digitizing the data and automating the processes? Put simply, manual systems are more challenging to monitor and adjust rapidly in response to analytical insights and make it difficult to generate those insights in the first place. Digital data is easier to aggregate and analyze, which speeds response time.

This does not mean that manual processes cannot inform business decisions. NeoGrowth, for example, intentionally began many of its processes manually to develop a deeper understanding of its clients when it was still fine-tuning operations, selecting the most relevant data gathered from these manual processes to inform its digital transformation.⁵²

The first step of digitizing data can be as simple as hiring a data entry operator to record data that was previously languishing in paper files and, after process-mapping, identifying data that is useful for decision-making. Organizations with digital strategies often automate processes in a phased manner, taking a long-term view of how ad hoc processes can integrate with existing data platforms.

Collecting data with no long-term plan for analysis can be a wasted opportunity.

Data can be structured, semi-structured or completely unstructured. A structured dataset (e.g., an Excel spreadsheet with consistently-labeled rows and columns, or other relational databases) requires minimal cleaning before it can be used for analysis. Semi-structured datasets need to be organized in a meaningful manner so that the data points can be related to an organization's existing datasets; this usually involves identifying and categorizing patterns. Unstructured datasets (e.g., text such as emails, Word documents, social media posts, or audio or visual files) also require identification and matching. However, the line is not always clear; a social media post has a date and time stamp, and may have tags or other structured elements. Unstructured data demand an application program interface (API) to pull the data into an FSP's internal systems. As a result, every time an external API is updated, FSPs may need to rework their data extraction process.⁵³

Manually-collected data should be stored in as structured a manner as possible so that it can eventually be migrated as legacy data onto a real-time system. For example, India's Aye Finance recently digitized its loan management

systems. According to Nitin Bhardwaj, who is responsible for IT application implementation at Aye, migrating the legacy data onto the core banking solution platform was easier because the data had been maintained in a particular structure.⁵⁴ If semi- or un-structured data proves too resource-intensive to analyze efficiently, FSPs could consider deploying middleware services that specialize in drawing unstructured data from various sources and delivering them in usable formats (middleware is the software that connects software components or enterprise applications). Questions of data architecture aside, it is up to each FSP to determine the appropriate business rules to follow, regardless of the types of data incorporated into its underwriting and credit scoring process.

Collecting data with no long-term plan for analysis can be a wasted opportunity. In one case that Abhishek Aggrawal, Chief Regional Officer, Accion, observed, one institution held financial literacy classes followed by examinations for loan applicants. There was no intention to use the exam scores for assessing creditworthiness but when this data was subsequently digitized, it was discovered that exam scores were highly predictive of loan default rates.

53. <https://developers.facebook.com/docs/apps/changelog>

54. Interview with Nitin Bhardwaj, August 16, 2016.



STEP 6: REVISE YOUR PROCESSES AND CULTURE

When asked for his main message to organizations looking to build data systems, Paulo Marques, a senior solutions architect at Capgemini told us: "Be prepared for change. Be prepared to encounter something every few months that will challenge your assumptions."⁵⁵ Organizations should carefully manage expectations and be willing to embrace uncertainty to successfully innovate, whether that means adopting new ways of looking at old data, or incorporating new data and processes into old systems.

Revising processes requires strong change management to shift organizational culture and established behaviors in the desired direction, which is often the most underestimated element of successful implementation of new technology or systems. Organizations must be willing to own the entire process and be receptive to change, whether for building in-house data capability or outsourcing. Ameer Parbhoo, a Senior Investment Officer at Accion, and Mike Ogbalu, Divisional CEO at Interswitch, both cited FSPs expecting fintech partners to drive the change as a reason that new initiatives failed.⁵⁶ Use the task force set up earlier in the implementation process to help develop a change management strategy and shepherd it through the organization.

55. Interview with Paulo Marques. 18th August 2016.

56. Interview with Ameer Parbhoo (16th August 2016) and with Mike Ogbalu (17th August 2016).

When asked for his main message to organizations looking to build data systems, Paulo Marques, a senior solutions architect at Capgemini told us: "Be prepared for change. Be prepared to encounter something every few months that will challenge your assumptions."

START WITH CLEAR COMMUNICATION

Building capacity to respond to data often requires changing organizational processes. Establishing a new status quo for conducting business requires deliberate engagement with all employees within the organization, thinking through the implications for existing roles and incentive structures. As an organization commits financial resources to a new approach, it must guide its employees throughout the process. Employees must have a sense of why the changes being introduced are important for the future of the organization and how they affect their day-to-day work. Salient issues include resistance to new ways of doing business, including changed credit methodologies, product management, or other activities. Wherever possible, employee feedback should serve as an input in the change management process. Clear communication and management of expectations build a sense of genuine ownership that helps gain employees' support.

For example, Accion worked with an MFI in Africa that had engaged with a fintech company to implement an automated credit pre-approval score, only to face considerable resistance from loan officers. They felt the pre-approval score had made their inputs to the application process irrelevant. To the contrary, the system had been designed to free their time to focus on more complicated cases where their expertise could be used more productively. The MFI's loan officers receive

The isolated exercise of analyzing data, without changing behaviors or the decision-making processes based on the results, has limited business use; however, it is hard for staff to justify experimentation when a current process is generating profit unless management explicitly encourages and rewards that experimentation, nurturing an environment where data-driven decision-making is consciously and actively promoted.

incentives for each additional loan processed; since the pre-approval score flagged clear approvals and rejections, it allowed them to focus on cases on the margins, potentially increasing their portfolio and their compensation. However, these changes were not adequately communicated to the staff initially, resulting in an overall lack of enthusiasm and a disappointing uptake for the new system in the field. This example illustrates the importance of developing an employee training and communications plan, one that goes beyond a functional description of how any new system will work to explain the value proposition to the individual employees, their customers, and the organization as a whole.

ALIGN INCENTIVES AND BUILD AGILITY

Clear communication is counterproductive if the value proposition is not aligned with staff incentives. According to Nektarios Liolios, co-founder and CEO of Startup Bootcamp, traditional FSPs rarely provide middle management with any incentives to innovate. To determine if you fall into that camp, consider: what are organizational attitudes toward experimentation and failure? To what extent is the learning process emphasized in key performance indicators compared to the final output? The isolated exercise of analyzing data, without changing behaviors or the decision-making processes based on the results, has limited business use; however,

it is hard for staff to justify experimentation when a current process is generating profit unless management explicitly encourages and rewards that experimentation, nurturing an environment where data-driven decision-making is consciously and actively promoted. Many organizations struggle with the challenge of building a new, data-driven culture that permeates up and down the ranks, but for those that succeed, senior management is key to setting and defining an innovative work culture.⁵⁸

Identifying a key metric for success can free your organization to think creatively about problem-solving. NeoGrowth has implemented a wide range of exciting data innovations to support its growth strategy. Raju Shetty, NeoGrowth's Chief Technology Officer, said much of this innovation was possible due to the organization's laser focus on a single problem; how to reduce customer acquisition costs.⁵⁹ Attention to this challenge enabled NeoGrowth's team to experiment with several parallel interventions that leveraged as many data sources as possible to reduce costs and address the pain points associated with customer acquisition.

High barriers within the decision-making structure make it difficult for organizations to respond to data insights, even with small, incremental experiments. Barriers to knowledge-sharing are lower in young companies, as early years of growth provide more opportunities to collaborate

Many organizations struggle with the challenge of building a new, data-driven culture that permeates up and down the ranks, but for those that succeed, senior management is key to setting and defining an innovative work culture.

across smaller teams. But that ease of collaborating early in a company's growth can lead to underestimating the importance of structuring such efforts to ensure that the spirit of collaboration continues as the organization matures. During an engagement with a leading MFI in Latin America, Accion found that loan officers were using the WhatsApp messaging application informally as a peer support group to troubleshoot and answer questions. Identifying those systems or ad hoc workarounds staff at all levels use to communicate can be a powerful starting point to understand how information flows across the organization, as well as to generate ideas for how to improve those flows. Ask through what channels does information flow up, down, and across the organization?

Apart from regular communication with employees, organizations should also assess the structures in place to facilitate formal sharing of documents and data. A common first step is to upgrade from ad hoc and decentralized document-sharing systems to some form of company-wide, cloud-based knowledge-sharing platform. In many large, traditional FSPs, internal compliance standards limit data sharing. These standards should be rationalized as far as possible to ensure there is a risk-appropriate attitude towards sharing information.

Organizations looking to encourage innovation could consider using the agile development framework, which has become particularly popular in the software industry as a methodology to fast-track innovative approaches.⁶⁰ Researchers from the Harvard Kennedy School suggest several structural changes to build an 'agile' organization, i.e., one with the ability to problem-solve, respond to data, and adapt.⁶¹ Two of their key principles are: 1) empower field managers to take decisions and experiment; and 2) introduce quicker feedback loops to understand the implications of pilot initiatives early on during the implementation phase, rather than wait to perform analysis afterward when it is harder to adjust for problems. In practical terms, this means managers and end-users of new systems should be active contributors to product development, so that organizations can leverage the collective wisdom of their staff.

Organizations can also organize periodic events or opportunities to actively collaborate and share knowledge. For example, one organization crowdsourced ideas by holding an innovation challenge for employees. According to an employee who was part of a team responsible for identifying innovative partners to work with the organization, the quality of ideas that the organization's own employees generated during the competition surpassed expectations. This experience made him realize there was considerable internal aptitude for innovation; the challenge was to institutionalize mechanisms that could tap into those ideas.

57. Interview with Nektarios Liolios, August 9, 2016.

58. Schein, Edgar H. *Organizational culture and leadership*. Vol. 2. John Wiley & Sons, 2010.

59. Interview with Raju Shetty, August 24, 2016.

60. <http://agilemanifesto.org/>

61. http://www.cgdev.org/sites/default/files/1426292_file_Andrews_Pritchett_Woolcock_traps_FINAL_0.pdf

FSPs need to build checks and balances into data-based decision-making processes to ensure results are backed by their real-world experience, business knowledge, and understanding of their customers.

STEP 7: DON'T GO BIG BEFORE YOU ARE READY

Tempting though it is to launch rapidly to see results, it is key to pilot the end-to-end solution in a controlled manner to learn as much as possible from a smaller sample and further revise the processes and technology. Test your assumptions: does the pilot meet the objectives you set out to achieve? Does the basic cost-benefit analysis support the business case to proceed? Figure 8 outlines several cost-benefit levers to assess when evaluating whether to scale up a pilot solution.

FOCUS ON THE FUNDAMENTALS

As the volume and variety of customer data available for analysis grows, so too does the temptation to build new algorithms. This tendency toward novelty, when left unchecked by common business sense, can lead to unreasonable conclusions. For example, Aneesh Varma, Aire's founder, shared a story of another startup trying to convince FSPs that a loan applicant's astrological sign predicted their likelihood to repay.⁶² Even if this correlation was found to be true in a particular dataset, it is unlikely to have strong predictive power.⁶³ Before basing algorithms on unproven correlations, FSPs need to build checks and balances into data-based decision-making processes to ensure results are backed by their real-world experience, business knowledge, and understanding of their customers. As Varma describes his company's central thesis, "The human mind has the best potential ability to discern creditworthiness, as it understands context really well. Aire is attempting to therefore [use data and technology to] emulate the same fundamentals that loan officers have been using for decades."

62. Interview with Aneesh Verma, July 20, 2016.

63. A tongue-in-cheek website aptly illustrates that spurious correlations abound in the real world, <http://tylervigen.com/spurious-correlations>.

FIGURE 8. KEY COST-BENEFIT DRIVERS



According to Quona Capital's Ganesh Rengaswamy, a vast majority of the early work performed by data analytics teams is focused on the labor-intensive tasks of collecting and cleaning data. He recommends putting together a team that is willing to chase data, rather than a team that expects to receive pre-packaged data.

HIRE THE RIGHT TALENT

Organizations need to develop both the analytical capacity to monitor and interpret the trends evident in their data and the operational capacity to respond to these real-time business insights. Ask who is handling data queries in your organization today, and are they equipped to handle the data queries your organization hopes to make in the future? This is a core business function, rather than an IT function.

One common piece of advice that emerged during interviews was that it is not necessary to hire highly-credentialed analysts in the initial stages of building data capabilities. Candidates with advanced statistical degrees are trained to leverage their technical knowledge with mature data systems that offer them rich, structured data. According to Quona Capital's Ganesh Rengaswamy, a vast majority of the early work performed by data analytics teams is focused on the labor-intensive tasks of collecting and cleaning data. He recommends putting together a team that is willing to chase data, rather than a team that expects to receive pre-packaged data.⁶⁴

One tip shared by Ekow Duker of Ixio Analytics was to structure hiring interviews in a manner that highlights individuals' problem-solving skills. Hiring managers could ask applicants to solve a quantitative problem in a group, and watch for individuals who not only have a structured problem-solving approach but who also work well with others. Other characteristics to look for are young individuals who have self-taught skills and who display a high level of curiosity.

64. Interview with Ganesh Rengaswamy, August 9, 2016.

65. Interview with Johann Bezoudienhoudt, August 9, 2016.

66. Interview with Mike Ogbalu, August 17, 2016.

At a minimum, FSPs should build the internal capacity to develop business hypotheses to be tested by data, and interpret whether the results make sense. If the project involves working with a third party (e.g., to build or provide an analytics platform), organizations need to evaluate the providers, once the required services have been defined and a shortlist of suitable providers has been identified. Most industry experts assert that at this point experience is everything. FSPs with limited resources cannot afford to experiment with service providers who have not successfully delivered their product requirements before. According to Johann Bezoudienhoudt, it is important to look for a mix of technical expertise and relevant field experience.⁶⁵ Mike Ogbalu further emphasized the need to ensure that FSPs have access to a competent local support system, as it is much more difficult to troubleshoot with a team operating in a different time zone and work environment.⁶⁶

STEP 8: PLAN A DATA GOVERNANCE STRATEGY AND STEWARDSHIP

It is important to create a data stewardship team or process to ensure regular monitoring of outputs. You will need to continually test the validity and veracity of the results. This can be done by checking for large outliers or by ensuring the data is internally consistent; i.e. there are no records that logically contradict each other. As data from various sources (including internal, third-party, and "softer," unstructured data) are integrated, ask if the resulting analysis "makes sense"? Terms of payment for third party data providers should be contingent on the data passing these internal quality checks. Responsibility for downtime, complaints resolution, compliance reporting, or other requirements should also be defined upfront.

It is a strategic decision to adhere to high standards of customer protection, because that is how the market is going to evolve, and it makes business sense to consider what constitutes ethical use of data.

Governance is another key part of the stewardship role, as it ensures that the data being collected is in line with local regulations, which may change from time to time. This also includes the necessary controls that will need to be in place to prevent unauthorized access to data sources or aggregated data sets. Third-party data accessed through open APIs or other sources will often come with terms and conditions of use. It is important to assess the legal implications of using any data that does not originate from your own systems. Understanding what you can and cannot do with third-party data can help determine whether the investment is worth the cost.

As organizations begin to expand their data use, they will need to respond to new regulatory frameworks. Proactively implementing standardized frameworks and protocols can help increase the confidence of partners and clients. Many emerging market countries do not offer clear legal frameworks for data security, particularly in the context of new business models. The organizations we spoke to that were operating in uncertain legal environments were adopting best practices from other countries with more mature regulatory environments.

For example, Mike Ogbalu explained that, at Interswitch, Africa's largest payment network, data security lies at the heart of its customer value proposition. Despite operating in a wide range of regulatory environments, Interswitch complies with the most stringent data transaction

standards and follows the global Payment Card Industry Data Security Standards.⁶⁷ Similarly, Selorm Adavedoh, the CEO of Digicel Haiti, believes that regulations in Haiti have yet to catch up to the realities of digital data collection. Digicel Haiti has therefore built its own internal checks for customer data use.⁶⁸ Adopting such a forward-looking approach can prevent later headaches. The board and senior management have a role to play in ensuring high data security standards at FSPs.

There are ongoing debates around data ethics,⁶⁹ effective disclosure, and informed consent. A recent CGAP study in Tanzania found that many consumers are willing to share private information in exchange for improved services.⁷⁰ However, building in systematic checks that respect your customer's private data is an important foundation for ethical business in an environment where it is easy to take advantage of vulnerable clients. Such proactive attitudes help raise the standard for customer protection. In India for example, an industry consortium is collaborating to build a framework that leverages India's biometric identity card to allow uniquely-identified customers to choose how much data they wish to share, with whom, and for how long.⁷¹ Jonathan Hakim of Cignifi believes it is a strategic decision to adhere to high standards of customer protection, because that is how the market is going to evolve, and it makes business sense to consider what constitutes ethical use of data.⁷²

67. Interview with Mike Ogbalu, August 17, 2016.

68. Interview with Selorm Adadevoh, August 12, 2016.

69. <http://sloanreview.mit.edu/article/achieving-trust-through-data-ethics/>

70. <http://www.cgap.org/publications/informed-consent-how-do-we-make-it-work-mobile-credit-scoring>

71. Interview with Sanjay Swamy, August 23, 2016.

72. Interview with Jonathan Hakim, August 25, 2016.

Managing new partnerships is like the early days of a marriage: you need to be sensitive to your partner's needs and adjust your behavior accordingly.

As exciting as the opportunities in a data-rich world may seem, they deserve the same degree of caution one would reserve for any new business venture. Many of the successes cited in this study are relatively young business models that have yet to prove their long-term viability. Monitoring economic cycles remains important for all FSPs, which should continue to adjust their risk parameters based on macroeconomic trends. A well-designed, data-driven approach can also help organizations track such metrics and incorporate macroeconomic trends into their portfolio strategy. Data-driven decision-making should remain grounded in economic reality, and basic risk mitigation strategies continue to be relevant.

STEP 9: ROLL OUT ACROSS THE ORGANIZATION

As with internal change management, fintech service providers repeatedly emphasized the importance of regular and clear communication for successful project management when working with vendors or other third-party partners, for example when rolling out a new data platform, product, or service. Ensure lines of communication extend through to the team responsible for product delivery. One major African bank struggled to develop a viable mobile money platform in a key market in part because its technical service provider had sub-contracted aspects of product development to another third-party vendor whom they had difficulty accessing. These intermediate layers led to delays in communication, which in turn slowed implementation.

73. Interview with Munir Duri, August 19, 2016.

74. Interview with Ben Knelman, August 3, 2016.

75. Ibid.

Munir Duri of Kifiya, an Ethiopian digital financial services provider, shared some experiences of how Kifiya learned the importance of an effective communication strategy the hard way. In its early days, Kifiya entered an agreement to provide a local MFI with branchless banking capabilities. While the agreement was approved at a high level by senior management, technical staff's fears of data security nearly stopped implementation. Kifiya put the project on hold and brought in various experts to conduct a workshop for the MFI's staff on implementation requirements and protocols. This helped establish comfort within the MFI's operational team and confidence in the process. Since then, Kifiya has built an in-house team to lead similar workshops for partner organizations as part of every engagement.⁷³

Of course, not every fintech partner will have the resources needed to provide such support. FSPs share equal responsibility in managing partnerships in good faith. Ben Knelman, Co-Founder and CEO of Juntos Finanzas, which has worked with a wide range of FSPs on customer engagement, likened managing new partnerships to the early days of a marriage: you need to be sensitive to your partner's needs and adjust your behavior accordingly.⁷⁴ By way of example, Knelman mentioned one frequent challenge startups face: cash flow. He commented that small adjustments on the part of the FSP, such as timely processing of vendor payment, can go a long way in ensuring that both organizations can maximize the value proposition of their partnership.⁷⁵

Conclusion

Data can be a powerful force for financial inclusion. The promise of big data is especially alluring. However, there remains a gap between the vision and the reality of FSPs’ ability to leverage big data for financial inclusion. That said, there are many useful early steps an FSP can take to lay the groundwork for big data and advanced analytics by improving its ability to generate, analyze, and respond to data of any size.

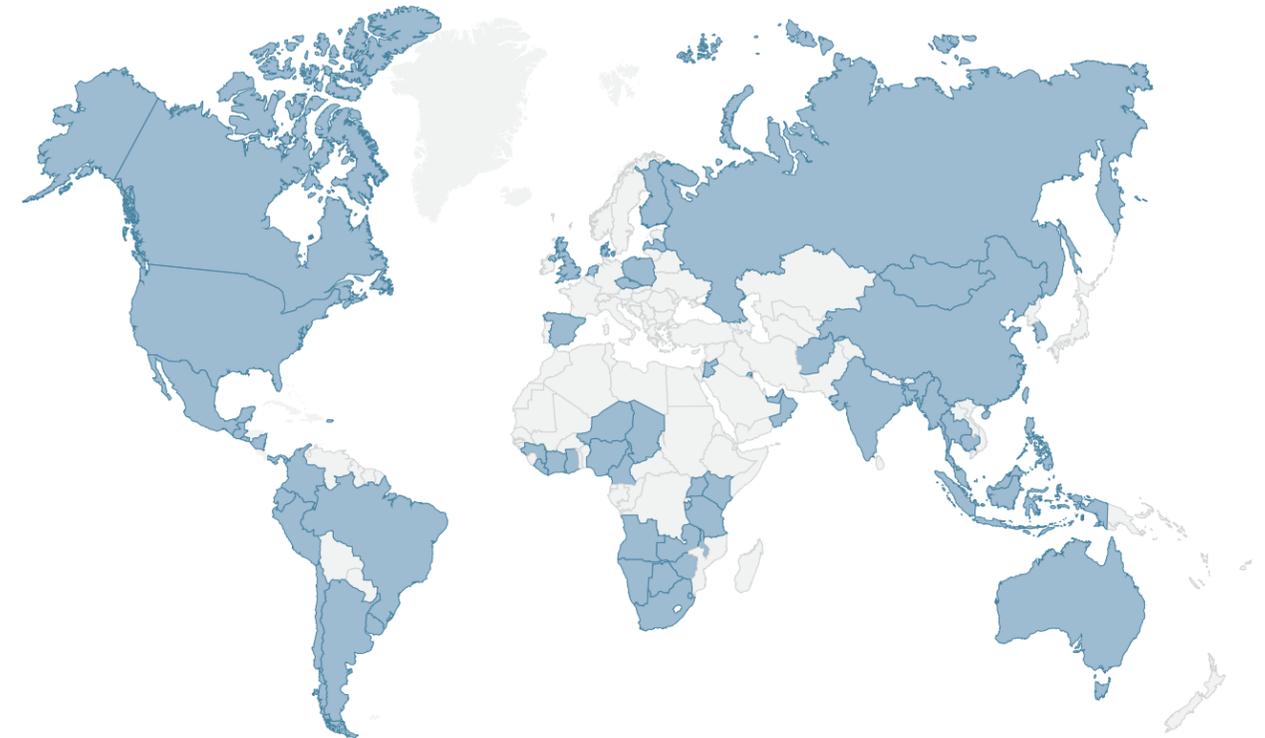
Most of these steps are simpler and more broadly applicable than deploying machine learning algorithms to generate alternative credit scores for automated lending decisions. Instead, they are intended to help FSPs navigate the new global landscape of data and plan strategically for the future. Ben Knelman of Juntos Finanzas captured the essence of the favored approach conveyed by this guide when he noted, “Real innovation sits where it looks the least glamorous and is the most painful.”

We are aware that not all the actions described in this guide will be relevant for every FSP. While the recommendations from Accion’s research are outlined here in the most broadly applicable manner, each organization in the financial inclusion space faces a unique set of challenges. However, regardless of whether an FSP decides to develop its own data capability, explore partnership options, or some combination of these, the components described in this paper need to be in place in order to better integrate data analytics into the organizational ‘DNA’.

“Real innovation sits where it looks the least glamorous and is the most painful.”

Many of the truly challenging decisions facing senior management can only be made through experimentation. **The goal here is to encourage FSPs to start adopting the right mindset and processes that will position them for success.** As Sanjay Swamy of Prime Venture Partners put it, “Small organizations need to be able to place multiple, small bets.” This guide was developed with the intent of helping those organizations willing to place such bets become better prepared to win.

Annex 1



GLOBAL MAPPING

S NO	NAME	DATA USES	DATA SOURCES	BRIEF DESCRIPTION OF MODEL
1	Advanced Merchant Payments	SME Finance	Hybrid	Traditional microfinance underwriting principles along with social media, payment processing, and other 'big data' sources to manage loan portfolio
2	AFB-Airtel	Consumer Finance, SME Finance, and Insurance	Mobile	Mobile repayment behavior and airtime purchases to qualify users for an initial loan
3	Aire	Credit Rating	Hybrid	Virtual interview in the middle of a loan application. Individual and industry data used to understand customer value and score creditworthiness
4	AliFinance	Consumer Finance	Hybrid	Digital footprint on various Alibaba platforms to assess creditworthiness

S NO	NAME	DATA USES	DATA SOURCES	BRIEF DESCRIPTION OF MODEL
5	Arifu	Customer Management	Mobile	Adaptive learning system that measures individual user needs and capabilities over Web and SMS in order to offer customized experience and financial education
6	Artoo	Customer Management	Hybrid	Digitizes operations for small financial institutions. Pulls credit bureau data, and can develop credit risk metrics based on field data
7	Avant	Consumer Finance	Hybrid	Develops customized rates for customers based on traditional credit scores and other data sources
8	Backed	Consumer Finance	Loan Application Process	Social Media for identity verification
9	Big Data Scoring	Credit Rating, Fraud Prevention	Hybrid	Digital footprint of customers to verify ID and credit score
10	Branch	Consumer Finance	Mobile	Smart phone application scans phone for loan applicant data to assess creditworthiness
11	Capital Float	SME Finance	Loan Application Process	Leverages data from psychometric tests, loan application documents, and e-commerce partners of retailers
12	Cignifi	Credit Rating	Mobile	Call data records to assess customer value and creditworthiness
13	Codensa	Consumer Finance	Utility Payments	Analyzes electricity bill to grant loans for the purchase of electric appliances. Repayment linked to electricity bill
14	Creamfinance	Consumer Finance	Hybrid	Online loan application process that automatically assesses applicants' creditworthiness
15	Credit Mantri	Credit Rating	Hybrid	Uses traditional credit score and non-traditional data to profile thin-file clients and match them to possible lenders with an EMI rate they can afford
16	CreditVidya	Credit Rating	Hybrid	Non-traditional data sources to assess creditworthiness of thin-file customers
17	DemystData	Credit Rating	Hybrid	Provides data services to financial institutions
18	Destacame	Consumer Finance	Utility Payments	Utility bills to assess individual's ability to pay back loans

S NO	NAME	DATA USES	DATA SOURCES	BRIEF DESCRIPTION OF MODEL
19	Early Salary	Consumer Finance	Hybrid	Uses national ID, bank statements, and social media to verify identity and assess creditworthiness
20	Earnest	Student Loan Refinancing	Hybrid	Collects data on education, savings, and career trajectory to score students without extensive formal credit history
21	EFL	Credit Rating	Loan Application Process	Proprietary psychometric tests to evaluate creditworthiness of applicants
22	Experian Microanalytics	Credit Rating	Hybrid	Alternative data sources to create credit risk profiles for individuals without access to credit
23	First Access	Credit Rating	Hybrid	Call data records to assess creditworthiness
24	Go Finance	SME Finance, Insurance, and Financial Record-Keeping	Other	Leverage data of consumer good companies to provide uncollateralized loans to suppliers down the value chain
25	Grow Financial	Consumer Finance	Hybrid	Machine learning algorithms on customer repayment behavior to provide personalized interest rate
26	Jumo	Consumer Finance	Mobile	Call data records to assess creditworthiness
27	Juntos Finanzas	Customer Management	Banks	Customer behavior data to customize a two-way messaging customer management platform
28	Kabbage	SME Finance	Business Transactions	Assess business activity to instantly process loan applications online
29	Konfio	SME Finance	Business Transactions	Sales transaction records to assess creditworthiness of small businesses
30	Kopo Kopo	SME Financial Services and SME Finance	Mobile	Business transaction activity and mobile money records to assess value of clients and upsell products
31	Krazy Bee	Student Loans	Repayment	Credit limits are revised on the basis of repayment behavior
32	Kreditech	Financial Services, Consumer Credit	Repayment	Loan repayment behavior to decide future credit limits. Also acquired a proprietary API platform to access bank transaction data
33	Lenddo	Credit Rating	Social Media	Social media activity to score creditworthiness and confirm identification

S NO	NAME	DATA USES	DATA SOURCES	BRIEF DESCRIPTION OF MODEL
34	Lendingkart	SME Finance	Business Transactions	Online loan application for small businesses. Focuses on current year transactions and other alternative data to appraise applicants
35	LendUp	Consumer Finance	Repayment	Provides better loan deals based on repayment behavior. Customers can also earn better deals by viewing educational videos
36	Mahindra Comviva	Customer Management	Mobile	Provides data insights to telecom industry and optimizes value added services
37	Mili	Consumer Finance	Social Media	Assesses social media activity/social capital of applicants to grant loans
38	MiMoni	Consumer Finance	Hybrid	Alternative data sources to provide instantaneous credit decisions online
39	Mjara-MFS Africa	Consumer Finance	Mobile	Mobile money remittance service that uses customer transaction data to set credit limits
40	M-Kopa	Asset Financing	Mobile	Solar products directly linked to mobile money accounts. Big data used to track customers and inform customer acquisition/distribution network
41	Mode	Consumer Finance, Mobile Airtime, and Remittances	Mobile	Allows mobile operators to calculate emergency air time limits based on past user behavior
42	M-Pawa	Consumer Finance	Mobile	Credit limits based on mobile money usage
43	M-Shwari	Consumer Finance	Mobile	Credit limits based on mobile money usage
44	Mybank	SME Finance	Hybrid	Alibaba's internet bank which uses alternative sources of Big Data to credit-score customers
45	myBucks	Consumer Finance, Insurance	Hybrid	Use propriety algorithm to qualify applicants for certain products/rates
46	NeoGrowth	SME Finance	Business Transactions	Alternative data sources to assess creditworthiness and loan amount
47	Oportun	Consumer Finance, SME Finance	Hybrid	Uses alternative data sources to lend to thin-file customers
48	Personetic	Customer Management	Banks	Predictive algorithms to customize client experience for financial institutions
49	Platiza	Consumer Finance	Loan Application Process	Online customer ID verification and credit scoring

S NO	NAME	DATA USES	DATA SOURCES	BRIEF DESCRIPTION OF MODEL
50	Progressa	Consumer Finance	Hybrid	Looks at bank statements, employment history, and earning potential to score loan applicants
51	RevolutionCredit	Consumer Finance	Loan Application Process	Behavioral analysis of customers during financial education videos
52	Scanntech	Credit Rating	Other	Payment transaction data from network of POS machines to assess creditworthiness
53	Sesame Credit	Credit Rating	Hybrid	Uses Alipay, online activity, and payment behavior with other digital vendors to provide a credit score
54	Spotcap	SME Finance	Business Transactions	Algorithms check real-time business performance and creditworthiness of applicants, providing instant feedback. The credit score determines the payable interest rate and fees
55	Tala	Consumer Finance	Mobile	Smart phone application scans loan applicants' data and assesses creditworthiness
56	Tiixa	Consumer Finance	Mobile	Provides nano-loans in mobile money or airtime based on usage patterns
57	TypeScore	Credit Rating	Loan Application Process	Natural language processing analysis on data provided by the borrower as part of the application process
58	Wecash	Consumer Finance	Hybrid	Machine learning algorithms to provide credit assessments in less than 15 minutes
59	Wonga	Consumer Finance	Hybrid	Online loan application that gathers data on applicants from the moment they log into the system
60	Yodlee	Credit Rating	Hybrid	Integrates financial data API with banks for better customer segmentation, product development and personalized products. Allows banks to expand customer base beyond traditional creditworthy applicants
61	Zaimi Online	Consumer Finance	Repayment	Decides repeat loans and interest rates on the basis of repayment behavior
62	ZestFinance	Consumer Finance	Hybrid	Online loan application that gathers data on applicants from the moment they log into the system

Annex 2

INTERVIEW LIST

Abhishek Agrawal, Accion	Kamran Hameed, Prudential Financial
Alan Martinez, EFL	Maria Fernandez Vidal, CGAP
Alex Holderness, Accion in the U.S.	Miguel Herrera, Quona Capital
Ameo Parbhoo, Accion Venture Lab	Mike Ogbalu, Interswitch
Andreas Rindler, Bearing Point	Munir Duri, Kifiya
Andres Calderon, Accion Global Advisory Solutions	Nancy Widjaja, Accion Venture Lab
Aneesh Varma, Aire	Nektarios Liolios, Fintech Bootcamp
Arunjay Katakam, GSMA	Nitin Bhardwaj, Aye Finance
Ben Knelman, Juntos Finanzas	Paul Makin, Consult Hyperion
Byoung-Hwa Hwang, CGAP	Paulo Marques, Capgemini
Ganesh Rengaswamy, Quona Capital	Raju Shetty, NeoGrowth
Hutch Carpenter, RevolutionCredit	Ricardo Calvo, Accion Global Advisory Solutions
Johann Bezoudienhoudt, Bankable Frontier Associates	Sanjay Swamy, Prime Venture Partners
John Plaisted, Accion Venture Lab	Selorm Adadevoh, Digicel Haiti
Jonathan Whittle, Quona Capital	William Derban, Fidelity
Jonathan Hakim, Cignifi	Xavier Enrique Cajiao Flores, Banco Pichincha, Ecuador
Jose Picoaga, Banco Financiero, Peru	

Citi Foundation



Accion would like to thank the Citi Foundation for supporting this paper. The Citi Foundation works to promote economic progress and improve the lives of people in low-income communities around the world. The Citi Foundation invests in efforts that increase financial inclusion, catalyze job opportunities for youth, and reimagine approaches to building economically vibrant cities.

This paper was written by Accion's Global Advisory Solutions team, which provides experienced operational and management support to strengthen our partners and maximize their impact. We leverage innovations to increase the quality and lower the cost of financial services, and thus help to build sustainable and scalable institutions focused on serving the financial needs of underserved individuals and small businesses.

Accion
10 Fawcett Street, Suite 204
Cambridge, MA 02138
Tel: +1 617.625.7080
Toll Free: 800.931.9951
Fax: +1 617.625.7020

accion.org

ACCION

Accion is a global nonprofit dedicated to building a financially inclusive world with economic opportunity for all by giving people the financial tools they need to improve their lives. We are building the next generation of top-tier microfinance institutions and in more than 50 years have helped build 65 such institutions in 32 countries on four continents that today reach millions of clients. Accion's impact investing initiatives are pushing beyond microfinance to catalyze more than

30 innovative start-ups dedicated to improving the efficiency, reach, and scope of financial services for the poor and our Center for Financial Inclusion is helping to build a stronger industry with high standards and broad engagement. Currently, our work spans nearly two dozen countries, including the U.S., where we are the nation's largest nonprofit microfinance network. For further information, visit <http://www.accion.org>



The suggested citation for this paper is: Kathleen Yaworsky, Dwijo Goswami, and Prateek Shrivastava: "Unlocking the Promise of (Big) Data to Promote Financial Inclusion," Accion. Cambridge, MA: March, 2017.

Please share this with your colleagues.

Accion welcomes your comments on this paper.

All Accion publications are available on the Accion website: <http://www.accion.org>