CFSi
Center for
Financial Services Innovation

Big Data, Big Potential:
Harnessing Data Technology for the Underserved Market
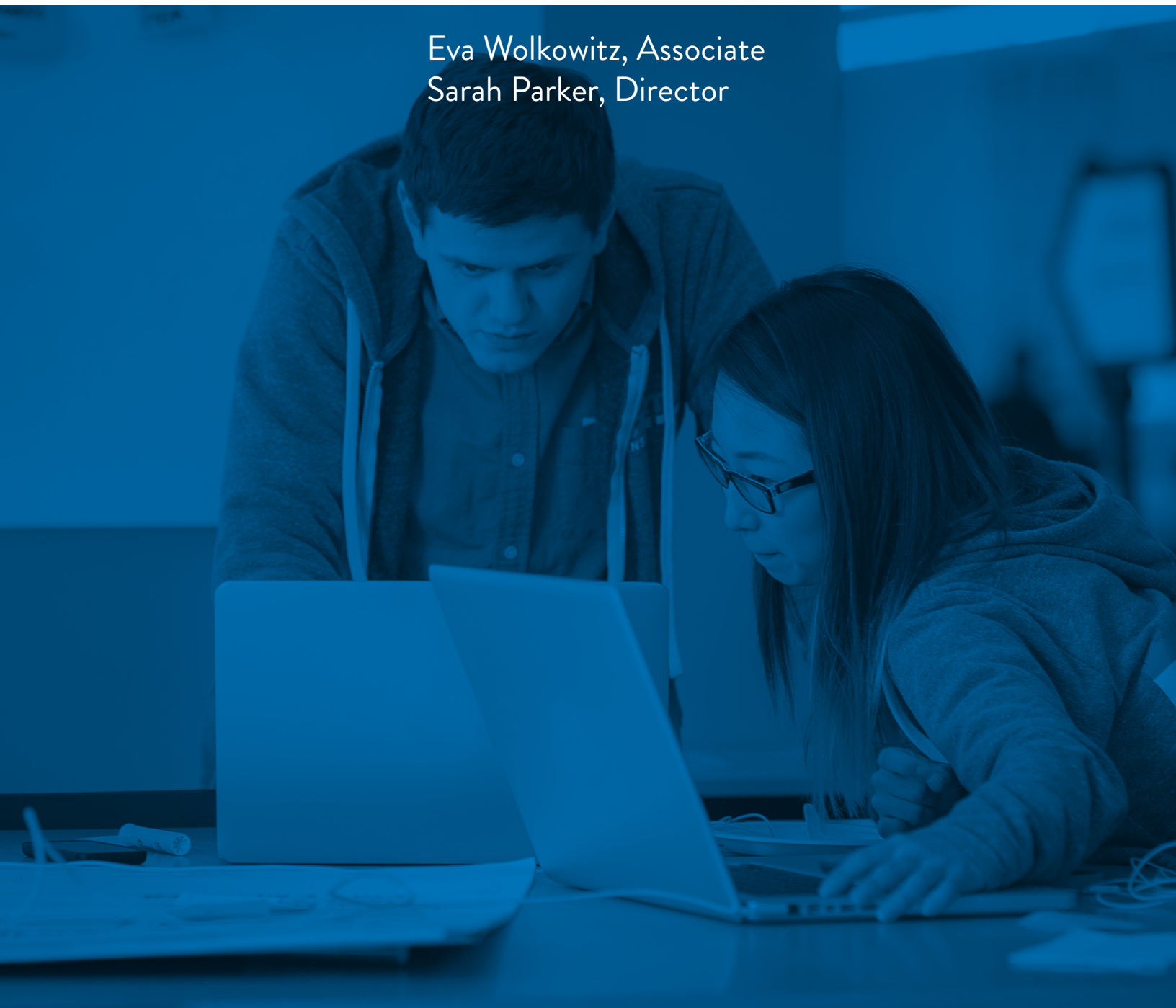
Eva Wolkowitz, Associate
Sarah Parker, Director

Table of Contents

CFSI
Center for
Financial Services Innovation

The Center for Financial Services Innovation (CFSI) is the nation's authority on consumer financial health. CFSI leads a network of financial services innovators committed to building a more robust financial services marketplace with higher quality products and services. Through its Compass Principles and a lineup of proprietary research, insights and events, CFSI informs, advises, and connects members of its network to seed the innovation that will transform the financial services landscape.

Center for Financial Services Innovation
20 N. Clark, Suite 1950
Chicago, IL 60602
312.881.5856

For more on CFSI, visit our website and join the conversation online:

www.cfsinnovation.com

@CFSInnovation

/CFSInnovation

Morgan Stanley   FORDFOUNDATION

This paper benefits from strategic support and financial input from Morgan Stanley. As with all CFSI research, this report is additionally supported by a grant from the Ford Foundation. The opinions expressed in this report are those of CFSI only and do not necessarily represent those of our sponsors.

About Morgan Stanley
Morgan Stanley (NYSE: MS) is a leading global financial services firm providing investment banking, securities, investment management and wealth management services. With offices in more than 43 countries, the Firm's employees serve clients worldwide including corporations, governments, institutions and individuals.
For more information about Morgan Stanley, please visit www.morganstanley.com.

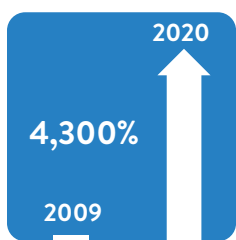**About the Morgan Stanley Institute for Sustainable Investing**
The Morgan Stanley Institute for Sustainable Investing builds scalable finance solutions that seek to deliver competitive financial returns while driving positive environmental and social impact. The Institute creates innovative financial products, industry leading insights and capacity building programs that help maximize capital for a more sustainable future.

For more information about the Morgan Stanley Institute for Sustainable Investing, please visit www.morganstanley.com/sustainableinvesting.

Introduction

The power of Big Data is advancing swiftly. Likewise, technologies, applications, and opinions based on Big Data are quickly changing. Big Data may be framed as the Next Big Thing or utterly dead; a panacea or a menace; the key to all future innovation or simply a hollow branding term. But between these extremes, Big Data is a vital area of focus for consumer finance because it has potential to support consumer financial health efficiently and at scale.

Big Data can be mutually beneficial for providers and consumers because it allows for an individually tailored approach to each consumer while maintaining a high level of efficiency that can support growth for the provider. Nowhere is this phenomenon more relevant than in financial services. The U.S. Bureau of Economic Analysis and McKinsey Global Institute estimate that the finance and insurance industries have the strongest ability to both "use and obtain value from" Big Data. Given these incentives, they also identify these industries as among the sectors with the highest "competitive intensity" to adopt systems that leverage Big Data effectively.[1]

This report takes the first step in identifying the great potential of Big Data technology by looking closely at how early adopters harness it to support consumer financial health and by exploring what the next generation of products that leverage Big Data may include. For the purposes of this report, Big Data can be understood as the collection and use of large data sets that can be broadly combined and distributed to identify patterns and create new data based on these insights—known as metavariables— to increase the effectiveness and efficiency of consumer finance products.

The Center for Financial Services Innovation (CFSI) has identified four trends with distinct momentum and value for developing Big Data capabilities that support increased consumer access, affordability, product quality, provider efficiency, and scale. These trends span companies involved in a variety of activities from lending and credit risk analytics, to personal financial management, to payments activities and beyond.[2] They do not address all companies that are data-driven or integrate data into their models, but rather those that specifically draw on and apply analytics to broad data sets beyond those generated internally. These trends are intended to provide the reader with a set of forward-looking vantage points from which to critically assess Big Data innovations, both those examined in this report and those encountered elsewhere. The four trends are:

Getting Granular Through Getting Huge
Designing Data Mosaics for Personalized Outcomes

Opening the Books
Trading Private Information for Better Value

Connecting the Dots
Communicating Data Elements Clearly

Hitting the Target
Strengthening Models Through Market Adoption

---

[1] Gutierrez, Daniel D., "InsideBIGDATA Guide to Big Data for Finance," Dell & Intel, 2014.

[2] Not all types of financial products hold relevant examples for all trends discussed.

Financial health means an individual's day-to-day financial system functions well and increases the likelihood of financial resilience and opportunity.

Big Data can be understood as the collection and use of large data sets that can be broadly combined and distributed to identify patterns and create new data based on these insights—known as metavariables— to increase the effectiveness and efficiency of consumer finance products.

New data points are expected to increase by 4,300%, to 35 zettabytes between 2009 and 2020.

2020
4,300%
2009

"Big Data Universe Beginning to Explode," Computer Sciences Corporation (CSC), 2012

Scope of this Report

The advent of Big Data raises a range of topics, including privacy, security breaches, data ownership, and advertising practices, which touch many industries.

These aspects, and many more, deserve close attention based on their ethical and practical implications.

This report, however, will only focus on a selection of Big Data's applications in financial technology that hold the potential to positively impact consumer financial health.

Big Data is an inherently technical topic. The focus of this report, however, is on overall trends and illustrative examples of their application in the market. We do not aim to examine the science of algorithms behind specific models of data analysis in this report.

See something new?

Refer to Appendix A for a glossary of select technical terms included in this report.

In the sections that follow, each trend is explained and illustrated with examples of company practices and a case study. Following a discussion of each of these trends, this report also addresses select risks and pitfalls of Big Data usage in financial services that can stem from poor data quality or analytical practices, and the potential for discriminatory or exploitive consumer applications.

The trends, company examples, and risks discussed in this report were selected and developed through a combination of market research and interviews with innovators and industry experts during 2014. Companies referenced and profiled in this report vary in their overall quality and maturity of Big Data usage, and are cited as examples of specific practices or efforts that illustrate trends with potential for success. No company or practice mentioned in this report constitutes an endorsement by the authors or sponsors.

Before turning to these four trends, let's take a brief look at how we have arrived at our current state of Big Data innovation.

Big Data's Evolution in Consumer Finance

Fundamentally, Big Data is nothing more or less than a set of tools that can be applied to creating, refining, and scaling financial solutions for consumers. As a financial technology tool, Big Data is neither inherently good nor bad because its value to providers and consumers depends on the quality of its application. As Big Data is increasingly sewn into the consumer financial services marketplace, in a range of more and less sophisticated ways, it is instructive to examine how Big Data innovations are faring; where the greatest potential lies for further development of Big Data; and how to foster its use in ways that are safe, responsible, and beneficial on a large scale.

Consumer finance applications of Big Data have existed ever since credit bureaus first gathered tradeline information to assign consumer repayment risk; insurance companies utilized applicant histories and demographics to set premiums; and car dealerships used information on average vehicle life expectancy to calculate blue book values. The earliest uses of large data sets to inform financial product offerings did not differ greatly, in theory or aim, from how Big Data usage is conceived today. Rather, its use was limited by rudimentary computing power and the hurdles of gathering and normalizing data from incompatible or non-digitized sources, both of which made the process relatively inefficient.

With this context in mind, Big Data is not an entirely new financial services tool, but has evolved into what it is today. Based on the exponential growth of computing power, data technology, and the unique propensity of Big Data to build upon its own foundations, these technologies will continue to be a work in progress. This progress can be shaped by the nature and quality of attention from established providers, startup innovators, investors, advocacy organizations, and regulators.

While enterprises manage an estimated 80% of all data, 70% is generated by individuals.

All Data

All Data

80%

70%

Enterprises

Individuals
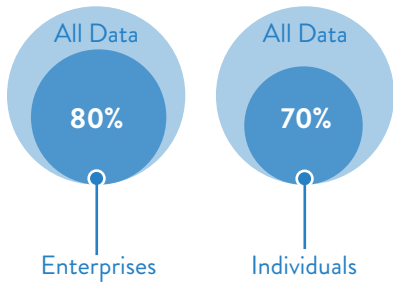
"Big Data Universe Beginning to Explode," Computer Sciences Corporation (CSC), 2012

As a greater share of modern life is encoded into data streams, Big Data tools can enhance the capabilities of a wide range of financial service companies through machine-learning algorithms that get smarter over time, building stronger predictive models as they gather customers and data points. Not all consumers—underserved or otherwise—will benefit from the use of new types of data, but the flow of data between people and financial services providers presents great opportunity for innovations to meet the needs of many consumers with high-quality financial services. It also necessitates greater exploration of how to harness Big Data safely and effectively to provide consumers with value and security. Though many industries, such as advertising and entertainment, are exploring ways to apply Big Data, the potential for consumer impacts are particularly acute in the financial services industry, where the application of new technologies can directly affect financial security.

This is a crucial time to focus on disruptive data technologies that can increase product value, reduce provider risk, and pass on lower costs and greater access to consumers. For some companies developing Big Data capabilities, attaining viable scale and efficiency can be strongly aligned with providing high-quality products that support consumer financial health. We will now take a closer look at how a selection of these companies, both in the financial services market, and in related fields with instructive parallels, are using and planning to use Big Data.

Trends to Watch:
Big Data Developments With Promise for Financial Services

Trend #1: Getting Granular Through Getting Huge
Designing Data Mosaics for Personalized Outcomes

The impersonal becomes personal when overlaying data from multiple sources yields precise, tailored consumer profiles. The data sets may be vast, but successful providers can use them to accurately identify individual needs, better know their customers, and help customers better know their own habits and financial patterns. Creating best-fit products that are responsive to the changing circumstances of each consumer increases efficiencies and cost savings for providers and increases access, value, and quality for consumers.

Trend #2: Connecting the Dots
Communicating Data Elements Clearly

Financial providers can learn a lot from data generated through the daily activities of consumers, but often consumers are not aware they are generating data that impacts analytical models. Some innovative companies are going beyond legally required minimums of disclosure by transparently conveying the types of data sources they use or explaining to consumers how their behavior can drive profile improvements that lead to better rates and offers. Well-informed consumers who are empowered to report erroneous data or shift behaviors to improve their financial standing can enhance data quality and reduce risk for providers while securing better outcomes for themselves.

Trend #3: Opening the Books
Trading Private Information for Better Value

Integrating highly reliable personal data leads to lower risk for Big Data-driven providers, but capturing that data often means venturing into private territory. One effective strategy is to invite consumers to opt in and voluntarily share more personal information and financial data in exchange for more attractive offers and lower rates. Products that allow consumers to control the balance of trade-offs between greater privacy and greater value can allow both customers and providers to reap greater benefits.

Trend #4: Hitting the Target
Strengthening Models Through Market Adoption

Big Data applications become smarter and more efficient as they secure industry partners and access to a richer variety of data sets. But to truly enhance data accuracy and impact end consumers, Big Data companies must not only increase their network footprint, but also scale up their industry adoption rates. Successful companies are finding routes to wide market exposure, often by crafting compelling tools for financial services providers that have a broad market reach and high returns on investment. Others are accessing strategic customer targets through partnerships with existing brands, or through well-travelled rails in other verticals, such as the mobile phone industry.

Trend #1

Getting Granular Through Getting Huge
Designing Data Mosaics for Personalized Outcomes

In our era of supersized banks and nationwide financial companies with millions of customers, a common concern is the lack of personalized service that can effectively respond to an individual's unique combination of needs and preferences. The judicious application of Big Data directly addresses this concern by allowing companies to know each of their customers better by amassing data on a vast scale. Overlaying data sets from hundreds, or even thousands, of sources can yield more precise, tailored consumer profiles. Data mosaics can produce insights that are more valuable than the sum of their parts by cross-referencing a wide range of data points drawn from public records, online habits, spending histories, and more. The results can cut down on inaccuracies, identify meaningful patterns, and predict future behavior on a granular level. That means providers can offer better-fit products and experiences to their customers with greater efficiency.
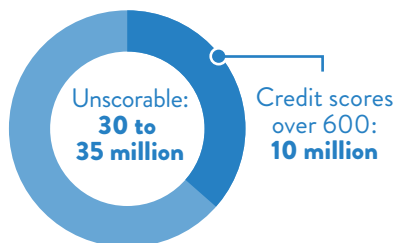
Credit Risk Scoring
During the past decade, as computing power became efficient enough to speedily integrate millions of data points, credit risk scoring was one of the first arenas where financial services companies sought to apply Big Data technology. In many ways, this is a natural extension of the risk-scoring activities of the large credit bureaus, which typically consider several dozen tradelines when forming a traditional credit score. Companies including **LexisNexis**, **L2C**, **CoreLogic**, **Clarity**, **FactorTrust**, **DemystData**, **VisualDNA**, and **Yodlee**[3] combine data from nontraditional sources, such as utility bill and rental payments, with traditional credit scoring sources to provide lenders with real-time underwriting models that reach further into the market of previously unscorable or high-risk borrowers. Their models, often used by nonbank lenders, can address all stages of the consumer credit lifecycle, from approval, to loan repayment, collections, and cross-selling of other services, and arrive at lending recommendations nearly instantaneously.

Some business-to-business (B2B) companies specialize their underwriting technology to fit the needs of particular types of lenders. **Kreditech** currently focuses on markets outside the U.S., particularly in developing economies where few, if any, end consumers have traditional credit scores. **Cignifi** and **First Access** also operate in emerging markets and base their risk models on data from loan applicants' mobile phone use and payments activity linked to their phones. Companies including **Affirm** and **Cognical's** Zibby platform use Big Data analytics to facilitate lending decisions for consumers financing the purchase of large household items such as furniture, appliances, or electronics. Other companies, such as **ZestFinance** and **Think Finance**, began life performing in-house risk scoring for direct consumer lending companies and have since separated into independent B2B operations.

30 to 35 million U.S. consumers are unscorable due to thin or nonexistent credit file information.

Unscorable: **30 to 35 million**

Credit scores over 600: **10 million**

10 million of them have scores over 600 when a wider range of data is used to compute scores.

VantageScore, March 2013
http://www.vantagescore.com/news-story/70

---

[3] See Appendix B for more information on the company names in bold.

While some companies specializing in data analytics, like ZestFinance and Think Finance, started out as internal units of nonbank small-dollar online lenders, others began life as independent Big Data enterprises and later partnered with – or were acquired by – credit bureaus, banks, or credit card companies looking to strengthen their Big Data expertise. **TransUnion** recently purchased L2C; **FICO** and LexisNexis each partner with FactorTrust; and **MasterCard** and **Experian** both partner with VisualDNA. Looking to generate new innovation for future industry partnerships and data technology gains, Yodlee's Next program makes data and funding available to a select cohort of financial industry startups, shepherding them through three stages of incubation, acceleration, and cultivation.

Lending

Some nonbank small-dollar lenders employ the Big Data analytics platforms of outside companies while others are developing their own proprietary platforms for underwriting direct consumer loans. Big Data startups offering small installment loans to consumers include **FairLoan**, a Yodlee incubator alumnus; its affiliate company Trulify, which provides applicant income verification data to FairLoan's lending partners; and the NetCredit installment loan offered by **Enova.**

Small business lenders **OnDeck** and **Kabbage** provide funding to small business owners using a proprietary mix of data that focuses on evidence of business health such as sales volume, customer reviews, and shipping histories.[4] These data points allow each lender to extend credit with more affordable rates and terms than small business owners could access using only their personal credit scores. **Neo** finances loans for first- and second-time car buyers after analyzing applicants' incomes, bank account histories, career and education credentials, and other metrics that sharpen its understanding of repayment risk. **Earnest** examines similar combinations of data to extend financing for higher education or other expenses that target the credit needs of early adulthood.

Personal Financial Management

Just as Big Data mosaics can help companies better know their customers when making loan decisions, similar analytics can help customers to better know themselves. While companies in the physical fitness arena, such as FitBit or JawBone, allow users to track activity and identify personal habits and patterns that can improve their bodily health, innovative financial companies are using Big Data to do the same for consumer financial health.

In addition to its credit risk focus, **Yodlee** has developed PersonalFinance, a set of tools that aggregates and analyzes consumer account information, then provides apps for expense analysis, cash flow management, bill reminders, savings goals, and more. **VisualDNA's** Youniverse product invites consumers to take carefully constructed quizzes that feed its analytics platform with data points to craft personalized profiles of individuals' financial personalities and risk tolerance levels. **HelloWallet**, acquired this year by investment research company Morningstar, works with large employers to advise their employees on ways to maximize their salaries, benefits, and other financial resources through insights found by analyzing large data sets of similar consumers.

---

[4] Kabbage has also recently begun making personal loans not designated for use by an enterprise.

When HelloWallet works through an employer to offer its financial management services to the end consumer, data is sourced at the company level and by individual participation, then tracked over time. As each customer uses its services, HelloWallet can track not only initial individual account information, but also data patterns that appear in the specifics of thousands of users. This allows the company to enhance and fine-tune its predictive models to give better advice and more useful communications to customers in the future.

Another benefit of being able to examine large data sets is the ability of personal financial management companies to share aggregate metrics with individual customers to provide context and comparisons. Thus, a company like **Mint**, bought by Intuit in 2009, which aggregates and tracks bank account activity for over 10 million users, can share with an individual user in Oregon how his spending on gas for the month of October compares with the average amount spent by all Mint users in Oregon that month. Having access to this type of aggregated, targeted information allows consumers to see their behavior in the context of their peers and draw conclusions on their spending choices accordingly.

Other Applications
A wide range of financial behaviors and histories, as well as tangentially related information that records consumers' locations, habits, and other interactions is continually encoded as digitally readable data. Forward-looking banks such as **Capital One**, electronic payments innovators like **Square** and **PayPal**, and prepaid card leaders including **NetSpend**, a division of TSYS, are working to leverage the data at their fingertips to drive smarter consumer engagement tactics and adjust their product offerings and features to best fit their customers. Banks like Capital One and startups like **Larky** are also developing geolocation tools that use real-time data on consumer location to pitch offers, send reminders, or otherwise engage in mobile financial service communication based on finely tuned targeting. The vision of these companies is to turn the wide lens of large, integrated data sets into experiences that are optimized for the individual, and maximize the value of interactions for both parties.

Summary
Combining massive data sets thoughtfully can lead to greater accuracy and granularity. Because financially underserved consumers vary widely and often have unique combinations of needs, data technology tools that allow providers to tailor their services while still maintaining low costs at scale are vital to consumers' and providers' mutual success. The Big Data mosaic effect has often raised concerns about its potential risk to consumer privacy when combinations of data result in overly sensitive insights. But with regulations and practices in place to avoid overstepping privacy boundaries, this technology can also be a powerful tool for returning an element of personalized service to large scale, automated product models. Whether used to assess loan repayment risk, give financial management advice, or make payments and transactions more seamless and intuitive in the future, data mosaics hold the potential for providers to better know their customers and for customers to receive financial services that better fit their needs.

Case Study[5]

Since its founding in 1999, Yodlee has developed Big Data analytics platforms for data-driven financial services, generating 45 patents, 45 million individual users, and B2B customers that include Fidelity Investments, American Express, and five of the 10 largest U.S. banks.[6] This cloud-based data analytics company, which went public in October 2014, anchors its value in profiling consumers and small businesses with a high degree of individualization.

Yodlee's personal financial management platform draws data from 12,500 sources, including banks, credit card companies, billers, and more to populate its end-user FinApps tools which overlay Big Data insights with individual account aggregation. By analyzing patterns in large data sets of financial consumer activity, then applying the results to individual accounts securely uploaded to FinApps, Yodlee helps consumers to budget, save, compare their spending with peers, and set reminders and alerts to stay on track. It uses a proprietary multi-step process to ensure that data is cleaned, normalized, and aligned for maximum usefulness in its vast data mosaic, then augmented with additional data to enhance tailored targeting for consumers.

"There is huge value that can be added to the data—like merchant identification and geolocation—all making it richer, more meaningful, and more actionable information within the experience for which it's being presented," says Eric Connors, SVP of Products at Yodlee.[7]

Yodlee is also notable for acting on its vision that greater openness and availability of data and data platforms can be a boon to the financial services ecosystem, driving innovation and collaboration that benefit all market participants. Its incubation program, Next, which fosters early-stage startups, and its FinApp Store, which provides developers with an open application program interface (API), both support the notion that data analytics strengthened by cross-pollination and collaboration represent the way forward.

---

[5] Note that some data and measurements included in case studies is self-reported by the private companies profiled and cannot be independently verified.

[6] Empson, Rip, "With Top Banks In Tow, Financial Service Provider Yodlee Hits 30 Million Users," TechCrunch, July 6, 2011.

[7] Webwire, "Yodlee Unveils the Formula to Create Meaningful, Secure, and Reliable Data to Power the Financial Industry," November 3, 2014.
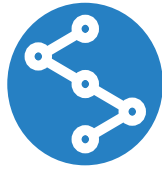
Trend #2

Connecting the Dots
Communicating Data Elements Clearly

Financial providers can learn a lot from what consumers do in their daily lives, turning data from mobile usage, geographic location, biometrics, or shopping habits into actionable insights. Successful companies are thinking creatively about how to indicate to their customers what types of data are drawn into specific analytics platforms, which pieces of data are currently impacting a consumer's overall financial profile, and how consumers can take advantage of two-way interaction with providers to correct faulty data or improve anticipated data points that could hold back their financial health.

Credit Risk Scoring
Federal disclosure law, dictated by the Fair Credit Reporting Act (FCRA) and the Equal Credit Opportunity Act (ECOA), as implemented by the Federal Reserve Board's Regulation B, stipulates that applicants who are turned down for a loan in the U.S. must be given a standard notice that states a legal, nondiscriminatory reason for the rejection along with the sources of information used.[8] This minimum requirement for transparency of data sourcing is crucial, allowing applicants to understand what elements of their profile proved damaging and to dispute the accuracy of this information with the original data source if the information appears incorrect. However, many loan applicants who are successful never learn which details impact their loan decision—either positively or negatively.

The use of thousands of data points in generating a credit decision increases the complexity of such decisions, adding to the reliance of borrowers on scoring platform companies to decipher the key drivers behind the results of their algorithms. Transparent statements of types of data used and how this data affects credit decisions help consumers achieve a basic understanding of the data's impact. In addition, establishing clear avenues of inquiry and communication for consumers to advise on the accuracy of data points or dispute their veracity goes one step further toward encouraging the consumer to be a partner in enhancing data accuracy. Achieving greater accuracy lowers risks for lenders, which can lead to better product access and terms for borrowers.

---

[8] The Federal Reserve advises that, "Adverse action notices under the ECOA and Regulation B are designed to help consumers and businesses by providing transparency to the credit underwriting process and protecting against potential credit discrimination by requiring creditors to explain the reasons adverse action was taken. The FCRA's requirements for adverse action notices apply only to consumer transactions and are designed to alert consumers that negative information was the basis for the adverse action." For more details, see: Ammermann, Sarah "Adverse Action Notice Requirements Under the ECOA and the FCRA," Consumer Compliance Outlook: Second Quarter 2013.

The data analytics firm **Acxiom,** while not directly a part of the consumer finance market, offers a prime example of a company which is experimenting with ways to actively engage the individuals about whom it collects data. By inviting the public to search for their names and profiles on a public interface, and update or correct information that is outmoded or shows errors, Acxiom encourages consumers to be partners in cleaning its data files. The low level of accuracy many website visitors encountered during Acxiom's public interface debut in 2013 means it still has a long way to go in enhancing the accuracy of its data. But the model it is piloting provides a clear avenue for consumers concerned about their financial profiles to scrub problematic data points and participate in enhancing its efficacy.

In other cases, a company may encourage consumers to actively participate in actually enhancing their chance of loan acceptance in real time. **RevolutionCredit** invites bank loan applicants to watch a short tutorial on how to successfully manage a loan. Tracking the applicant's willingness to follow through with the tutorial helps to identify, in combination with other data, who will present the lowest risk for extending credit. By asking prospective borrowers to opt in to these tutorial modules, RevolutionCredit enlists them as partners in building their own creditworthiness through contributing their time and attention to becoming more financially educated.

Lending

Direct lenders that utilize Big Data analytics, such as **LendUp** and **Elevate**, know that certain consumer actions, such as on-time or early repayment and participation in financial education modules, are good indicators of overall low risk. LendUp incorporates borrowers' repayment behavior and demonstrated willingness to engage in financial education into its analytics modeling to produce a sliding scale of loan rates and terms. Elevate, the consumer lending spin-off of **Think Finance**, also provides loans along a sliding scale that depends on its assessment of borrower creditworthiness, then offers borrowers a clear road map for garnering better value over successive loans. By providing guidelines for borrowers to leverage the value of positive behaviors, these lenders create mutual benefits for themselves and their customers. When these companies report activity to major credit bureaus, they also help their customers understand the levers of building credit and managing their scores for the future through positive repayment behavior.

Other lenders, such as **Lenddo**, which currently operates in Latin America and the Philippines, aim to clearly describe the types of source data that feed into their models to prospective borrowers. As Lenddo weaves social media data points into its consumer risk profiles to aid in identifying the most responsible borrowers, it is important for the borrowers to understand that their online social relationships affect their financial profiles. Clear indication of the types of data used in this lending model can help applicants visualize whether their social media data may present an asset or an impediment to receiving a loan.

The subset of online consumer lenders known as marketplace lenders also has recognized the importance of giving borrowers a clear sense of what data points impact their chances of getting a loan, the rates and terms of the loan, and how their financial behavior can shift their scores over time. Companies including **Lending Club**, **Funding Circle**, and **Prosper** that connect individuals seeking small personal or business loans with investors, apply their own proprietary scores to applicants. They create classes of loans with a range of risk assessments and rates of return for the investors to choose among, and maintain a high degree of transparency and personalized communication with customers, both for applicants seeking to understand their status and investors who wish to look at the past performance of loans.

Other Applications

Companies with millions of active credit and debit cards generating real-time usage information, such as **Visa**, **MasterCard**, and **CapitalOne**, are finding that Big Data analytics allows them to detect fraud more accurately by assessing multiple predictive data points in concert to identify unusual activity. **BillGuard**, in partnership with **Experian**, uses the aggregate data of its members to alert users when another user experiences a fraudulent event from a merchant they have in common. The use of predictive data models may also be applied to identifying borrowers at risk for default before they miss payments, which allows lenders to mitigate losses through early intervention. Using Big Data to ferret out fraudulent activity or struggling borrowers can help reduce provider losses and protect consumers, but without a strong consumer interface to communicate about possible fraud or repayment concern, these systems can also produce frustrating impediments for consumers who may find an account disabled or frozen.

By making investments in consumer-facing staff who can engage in two-way communication when analytics raises questions on an account, these large companies can allow Big Data systems to process information efficiently without losing value through bottlenecks in resolving these red flags. Good communication channels for consumers also allow card companies to refine their analytics approaches, getting smarter over time about which possible fraud events or worrying behavior on the part of borrowers are indeed accurate and which are false alarms.

Summary

Since the use of Big Data and machine learning in financial services gets more complex as it matures, innovators face the ongoing challenge of clearly expressing the underpinnings of their analytical insights to consumers, regulators, or others who watch this industry. Yet when companies are able to encapsulate, even at a high level, the data driving their output, this information can be invaluable to consumers. The ability to connect the dots among the ingredients that drive a credit decision, identify and dispute an inaccuracy, or generally understand which data points help or hurt their outcomes, are important elements of empowerment for consumers. At the same time, it is often in the best interest of companies to provide their customers with the information and channels of communication necessary to become a partner in developing accurate data and demonstrating low-risk behavior and habits.

E/evate

Case Study

Elevate's online installment loan *RISE* and similar U.K. product Sunny both operate on the concept that consumers will strive to modify their financial behavior in a positive direction if given clear information about how to conduct themselves as a borrower and how their actions impact the rates and terms of the loan they receive.

Elevate's website explains that new applicants typically qualify for a *RISE* installment loan of between $500 and $5,000 at rates that range from 125% to 365% when calculated on an annual basis. From this starting point, borrowers can access future loans at lower rates by demonstrating responsible behavior through successful repayment, on-time payments, early repayment, and other actions such as completing online financial courses, which translate into discounts for participants in the *RISE* Rewards Program.

Additionally, Elevate reports successful repayment to the national credit bureaus, facilitating higher credit scores for borrowers who act on its clear guidelines for successful repayment. While it provides the bureaus with borrowers' data, Elevate also shares credit score information from the bureaus with its borrowers. By supplying free credit score information and credit alerts flagging changes in borrowers' scores, Elevate completes the circle between consumers and their financial data.

This model encourages borrowers to track their progress and connect the dots between the data they generate and the loan terms they can access, helping them to understand the linkage between their financial choices and their resulting financial profiles. Putting consumers in the driver's seat is good for Elevate's business because proactive borrowers with actionable information reduce the risk of nonpayment or costly loan adjustments. Elevate's borrowers, in turn, have the chance to learn how to generate data points in their favor that merit lower borrowing costs.
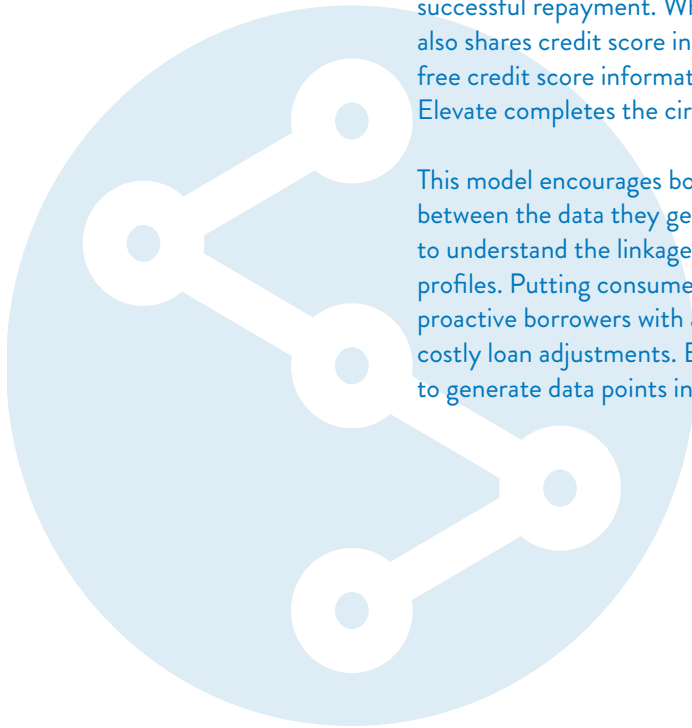
Trend #3

Opening the Books
Trading Private Information for Better Value

Companies that aggregate data sets draw their data from a wide variety of sources, some of it publicly available, some purchased from data brokers, and some newly created metavariables generated by the analytics performed on the initial data. Yet other data remains private, siloed in individual accounts and isolated databases controlled by consumers. When consumers choose to opt in to an invitation from a provider to share this information, they become active participants in supplying data points that feed Big Data algorithms. Sharing private financial information comes with risks which consumers should carefully consider. Providers should actively work to gain their trust by constructing strong safeguards and high standards for maintaining individual privacy.

While consumers should examine how a provider may use their information before opting in, sharing this data can provide great opportunity for consumers to enhance the accuracy of their financial profile and reduce the risk of the financial services offered to them. In effect, the consumer who grants a provider access to her bank account data, list of professional connections in a social media account, or mobile phone usage history is consciously trading an element of her private store of data for a financial service with potentially better rates, better access, or higher quality outputs. If the resulting product offer is of high quality and a good fit, this can be a valuable trade.

Credit Risk Scoring

First Access predicates risk-scoring products for lenders in emerging markets predominately on prepaid mobile phone usage. Though much of its analytics rely on aggregate data provided by telecom carriers that is devoid of personal information markers, the platform matches its analytics to an individual loan applicant by asking for permission to access the applicant's phone records. First Access does not examine the content of phone calls or text messages. Instead, it studies applicants' mobile usage frequency, timing, and other behavioral features. By simply typing "1" into their phones when prompted, applicants for a First Access loan grant access to their mobile records, and then receive a loan decision within a few minutes.

VisualDNA takes a different approach to gathering private data directly from consumers, setting up online quizzes that assess the quiz-taker's financial personality while using his unique combination of answers to add to an aggregated collection of data that identifies patterns correlated with borrower risk.

2.5 billion people—about 1 out of 3 worldwide—have no bank account.

6 billion people—more than 3 out of 4 worldwide—have access to a mobile phone.
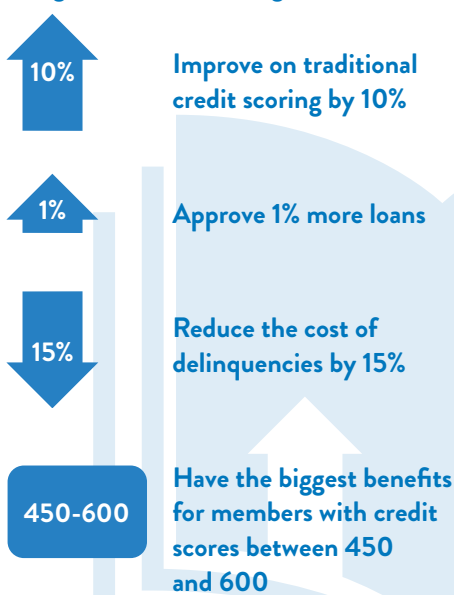
World Bank "Financial Inclusion," Global Financial Development Report, 2014 and "Mobile Phone Access Reaches Three Quarters of Planet's Population," July 17, 2012

Lending

The depth of information provided by consumers when applying for a loan can also factor into lending decisions of direct credit providers. **Lenddo** requests a peek into applicants' social media connections to draw in data points, while **Earnest** frames its installment loans as investments in the future of its borrowers, asking applicants to provide access to their academic and financial records, bank information, and employment histories. The lender then uses these data points in conjunction with traditional credit data and applicants' personal statements of their credit needs and goals to make what it characterizes as a "merit based" loan offer. **Neo** requests similar access to bank account, education, and employment information to accurately predict earnings potential and propensity to repay its auto loans.

Kabbage, which focuses primarily on credit for small online businesses, asks for access to private data that demonstrates healthy business management and cash flow, matching the individual information released by applicants with aggregated analytics performed on data drawn from partners including UPS, Etsy, and Stripe. **OnDeck** de-emphasizes the personal credit scores of small business owners, and instead focuses on data that illustrates business cash flow, past use of credit, and vendor payment history. Often the small business owners these companies lend to would not otherwise qualify for an affordable loan based on their credit scores. By letting companies like these look at their books, small businesses increase their chances of being offered a loan with a lower rate, better terms, or a higher limit in proportion to how much their private data helps these lenders reduce their estimated risk.

For marketplace lenders, encouraging loan applicants to share their personal histories, reasons for their loan request, and plans for the influx of funds is fundamental to attracting investors. **Prosper** and **Lending Club** pair this information with credit scores to create their own scores, while **Funding Circle** requests business tax filings and income statements to better understand the financial specifics of the small business seeking capital. Knowing how the person or small business is faring financially, why they want funds, and how they plan to use and repay the money is integral both to risk decisions and scoring by these marketplace lending platforms, and to successfully attracting investors who will fund the loans.

Personal Financial Management

Companies that help users track and manage their financial lives are also finding that data analysis is enhanced when consumers share nontraditional information to paint a fuller picture of their risk profiles and reputations. **Happy Mango**, a data technology startup that generates alternative credit scores, aims to help consumers track their finances, explore their ability to repay future loans, and find sources of credit that fit their needs. It accomplishes this by incorporating factors such as income, savings, education and career data, and social connections who will vouch for the reliability of the prospective borrower, along with bank account information and borrowing histories.

An experiment by the Jumiya project in conjunction with Filene Research Institute estimates that credit unions using a machine-learning model could:

10% — Improve on traditional credit scoring by 10%

1% — Approve 1% more loans

15% — Reduce the cost of delinquencies by 15%

450-600 — Have the biggest benefits for members with credit scores between 450 and 600

Kallerhoff, Philipp, "Big Data and Credit Unions: Machine Learning in Member Transactions," Filene Research Institute, 2013

Methods that encourage consumers to share a fuller picture of their financial lives can also benefit banks, credit unions, and other financial institutions that are constantly seeking deeper consumer loyalty and engagement. These institutions already possess a wealth of internal financial activity information, but most of them fail to actively request outside data from their customers' other accounts and financial activity that could help them to better understand, segment, and market to their customer base. Giving advice based on aggregated account insights, like Happy Mango, or using surveys and quizzes, like **VisualDNA,** are two models of how financial institutions could grow and maintain customer relationships by maximizing the exchange of value between private information and better services.

Summary

There are fundamental tradeoffs between the accuracy and depth of data and the degree to which private consumer information is revealed to a provider. The right balance between these elements is different for each individual. Companies that transparently ask consumers to share personal information while leaving the choice in their hands respect the boundaries of privacy. At the same time, they provide customers with avenues to turn their personal information into an asset with the potential to garner better access to affordable financial services. The combination of private data shared when consumers open their books and the data available to providers through brokers or public information can drive insights that lower risk and benefit all consumers of a given product by strengthening a provider's aggregate data over time.
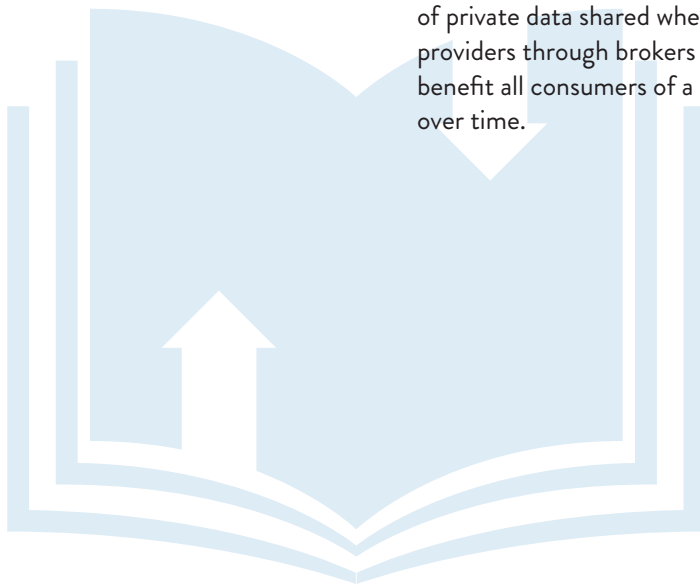
Case Study

When approximately 30 million subprime consumers and 6 million millennials seek an auto loan each year, many don't have the credit scores or established credit history to secure good terms. Some resort to borrowing at inordinately high rates. Others cannot secure a loan at all and end up paying cash for a used clunker. Neo offers another solution to consumers who are turned down for a loan. It asks applicants to grant access to their private bank account information and other details of their educational credentials, job status, and earnings trajectory. Neo shares this information with auto lenders, providing a more comprehensive view of each auto buyer than their credit score alone to drive increased approvals and loan performance.

Simultaneously, Neo analyzes borrowers in the context of large data sets that hold aggregated information on the range of financial circumstances for consumers with similar metrics. By applying Big Data analytics to identify risk in this larger population, supplemented with the growing base of direct client data points it collects, Neo looks beyond applicants' credit scores or meager assets to make lending decisions that reflect its predictions of the borrowers' future income levels and circumstances. The loan decision is, "based on where they are in their life right now, rather than what happened to their credit in the past," says Neo's CEO Navin Bathija, a significant consumer opportunity for, "something as crucial as an auto loan which is directly related to mobility and earnings potential."

By granting Neo access to detailed and current sources of data, like records of income deposited in a bank account, car buyers can potentially trade the value of their personal information for the value of a more affordable and accessible auto loan. And every time Neo makes a loan in partnership with a lender, and follows the data stream of repayment activity from start to finish, its insights become a little stronger. This enables Neo to create more reliable risk assessments that can be passed on to the next consumer, with below-market interest rates as low as 6% APR.
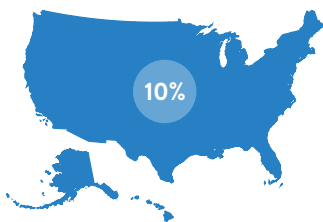
Trend #4

Hitting the Target
Strengthening Models Through Market Adoption

The lack of industry adoption is one of the main barriers to reaching end consumers with Big Data analytics that enhance value and efficiencies. Designing strong predictive models does not ensure they will see wide implementation. Even when companies with a strong Big Data focus form partnerships with industry leaders or garner market interest, their innovations may continue to languish in development behind the scenes. Analytics platforms that rely on machine learning are hobbled if they cannot reach a critical mass of end consumers and hone their models with usage data. It can take heavy lifting to convince entrenched players to adopt the new methods or new infrastructure required to bring an innovation to scale, but strong market implementation that moves past the theoretical stage of data experimentation is key to success.

Credit Risk Scoring

A prime example of this phenomenon is the case of the **FICO** score. Every few years, FICO releases an updated version of its platform—the most current is version 9— that incorporates new and more sophisticated credit risk scoring elements. In recent years, these new versions have begun to include new or differently weighted data streams previously used only by nonbank lenders leveraging Big Data insights. For instance, FICO Score 9 recognizes that medical debt is not a good predictor of other forms of non-repayment, so it largely disregards this metric. Yet this change has nearly no impact on actual consumers struggling to secure credit due to medical debt because most mainstream providers have not switched to using the new FICO model. In fact, the majority of lenders nationwide are still using FICO Score 4, and only a relative handful has gotten so far as FICO Score 8, released six years ago.[9] It seems that investments in upgraded software and training pose a major impediment for lenders in the adoption of a new FICO system.

VantageScore's recent, public efforts to convince Freddie Mac and Fannie Mae to adopt its 3.0 version, which these lending behemoths are currently studying, is one example of the recognition by industry players that wide market adoption is key to support and enhance the effectiveness of risk scoring platforms based on the tracking of consumer financial data.[10]

FICO Score 9 is the first version to differentiate between unpaid medical debt and other types of unpaid debt when determining credit scores.

---

17 million U.S. adults—about 1 in 10 with traditional credit scores—have a lower credit rating due to outstanding medical bills.

10%

"NerdWallet Health finds Medical Bankruptcy accounts for majority of personal bankruptcies," Nerdwallet, March 26, 2014

"FICO® 850 - The Perfect Credit Score is a Reality," SubscriberWise via FICO, July 2010

The CFPB estimates that medical debt, even after it is paid off, can lower credit scores by 10 to 22 points.

"CFPB Study Finds Medical Debt Overly Penalizes Consumer Credit Scores," Consumer Financial Protection Bureau, May 20, 2014

---

[9] Becker, Tracy, "New FICO Scoring Model Won't Matter for Mortgages," Nerdwallet, September 8, 2014.

[10] Wack, Kevin, "Fannie, Freddie to Evaluate Alternative Credit-Scoring Models," American Banker, September 22, 2014.

Other credit risk scoring innovations using Big Data have had more success by taking advantage of systems already in use by a large swath of their target market. **RevolutionCredit** creates a video module product that can be sewn right into the online interfaces banks, lenders, and billers already use with their consumers. **Cignifi** and **First Access** rely on mobile phone usage for both their data streams and their primary channel of communication with local lenders and end consumers. **Cognical's** Zibby integrates its lease-to-own product into online shopping carts and storefront checkout counters, bringing its financing options to consumers through the retailers where they purchase appliances, furniture, and other big-ticket items. **Affirm** offers installment credit and other financing options through online retailers for shoppers who lack other credit options but meet its risk criteria through the use of alternative data.

LendUp, which first launched as a direct loan provider in 2012, opened its application program interface for use by other lenders in 2014. By offering to customize its platform for banks, credit unions, and nontraditional lenders, LendUp aims to expand its reach to more end users and sharpen its platform by gaining a wider pool of current user data.

Lending

In the direct lending arena, **Kabbage** has found a swift route to meaningful levels of end consumer adoption. Its target borrowers are online business owners, and its means of drawing in relevant data are predicated on the existing ecosystem of the online business. Kabbage draws data points from interactions along the online supply chain between vendors and their partners and customers, including shipping companies like UPS and electronic payment companies like Stripe. By interfacing with online business owners in their native environment and drawing information from the relationship history data inherent in operating an online business, Kabbage has found a ready stream of borrowers and relevant borrower information to scale its risk decisioning process.

Personal Financial Management

HelloWallet attempts to close the gap between market adoption and smart data analytics by reaching large groups of consumers through their employers. By offering their service as an employee benefit, and addressing the usage of salaries and other benefits distributed through uniform channels to a large cohort of employees, HelloWallet hopes to capture enough consumer interest to process data points at a scale that provides meaningful insights for their own modeling and for their large employer clients as well.

Summary

There is a circular, sometimes cruel, logic to the successful launch of Big Data analytics solutions that necessitates broad, long-term data streams in order to craft responsive, increasingly accurate algorithms, but also demands demonstrated accuracy and meaningful insights to drive adoption. While it is important for companies to test their assumptions and gather as many data sources and partners as possible before venturing into the market, the strongest models cannot really achieve their goals, or provide an impact for end consumers, until they are widely available. This entails not only launching the product, but creative and sustained efforts to ensure the new model is adopted by intermediate providers or given broad advertisement in, or synergy with, existing channels of product delivery. Big Data financial services that are designed to fit into the software, strategies, and channels of existing B2B clients and consumer habits stand the greatest chance of achieving their goals.

Case Study

RevolutionCredit is a software-as-a-service behavioral data and analytics platform that works with banks, lenders, utilities, and subscription service providers to offer bite-sized gamified personal financial management programs to customers seeking personal loans, credit-based services, or credit card approval. In order to access the final credit offer and preferred rates and terms, the consumer must complete one or more of its 10-minute modules, each divided into one-minute video lessons designed and tracked by RevolutionCredit and offered through their normal bank, nonbank lender, or biller.

Behind the scenes, RevolutionCredit, founded in 2012, uses the data gleaned from these consumer interactions to feed its credit risk scoring model, identifying which participants show signs of engagement that correlate with reliable credit repayment. This information enhances traditional credit score data to drive smarter lending decisions and lower loan losses. During its first years in market, lenders and billers using RevolutionCredit's program realized a 15% increase in approvals, 30% decrease in delinquencies, and 67% increase in retention rates when compared to credit scoring only.[11] The data model emphasizes predictive, forward-looking metrics, rather than the historical focus of traditional credit scores alone, and is able to differentiate between risk levels of consumers who have similar traditional credit scores.

RevolutionCredit can add its gamified modules to the existing consumer interfaces of lenders and service providers already in operation, so it has been able to reach end consumers with the results of its Big Data analytics from the moment it entered the market. The lenders that use its platform, in turn, continue to feed data points following individuals' progress with these modules back to RevolutionCredit to further hone its insights.

Building a Big Data analytics product that can quickly test its premise in the market rather than hang back, absorbing data to tinker with a model in a vacuum, is particularly important for RevolutionCredit because, "the data is created at the point of transaction," founder Zaydoon Munir explains. "RevolutionCredit is about positive selection... Someone who completes the required programs now has demonstrated a stronger 'intent' signal: they are committed."[12]

---

[11] RevolutionCredit pitch deck as shown at CFSI EMERGE Forum, June 5, 2014.

[12] Better Know a Deal: Getting Credit for Creditworthiness," F.B. Heron Foundation blog, October 15, 2014.

Risks & Pitfalls

Risks & Pitfalls

The great potential to achieve efficiencies and high-quality offerings for consumer financial products with Big Data do not come without risks and the potential for misuse. While this report does not provide a comprehensive treatment of all risk topics, an examination of the risks and pitfalls most closely related to the themes of data personalization, transparency, two-way sharing of private information, and market adoption can enhance the appreciation of challenges inherent in pursuing innovation based on these trends.

The Risks of Poor Inputs and Practices

Along with the advent of Big Data analytics companies has come a proliferation of third-party data brokers that collect and aggregate data sets purchased from merchants or gleaned from public records. These data sets can vary greatly in quality and accuracy. No platform based on data sets which depend on erroneous, outdated, or irrelevant information can produce high-quality results. Companies that purchase data can avoid negative results by learning as much as possible about the data's provenance, and by using the very concept of Big Data as a tool to combat inaccuracies. By comparing identical or similar data from multiple sources, analytics companies can identify data points, or even whole data sets, that appear to be outliers that should not be trusted.

Moving past the purchase of data to the analysis stage, financial companies require data from the results of their initial products to feed back into their models. Machine learning and other non-linear methods are designed to adjust and fine-tune their predictive power through such feedback. Yet there is also a human element to crafting the best systems. Companies should strive to question their models and keep a finger on the pulse of macroeconomic developments, monitor evolving consumer habits and attitudes, and scan the horizon for entirely new types of available data to ensure that their platforms do not grow stale or detached from real-world value.

Similarly, the proliferation of large data sets will require increasingly sophisticated data scientists equipped with tools and training to extract accurate insights. Because machine learning and other non-linear data analytics can only be as good as those who analyze them, companies should recruit staff who not only bring superior skills in building data platforms but also have deep knowledge and abiding curiosity about the idiosyncrasies of the financial services marketplace they serve. Data scientists should be proactive in removing significance from data that may appear correlative with consumer results yet have no identifiable relationship. Extraneous results that produce unhelpful "noise" can clutter Big Data outputs and dilute the ability to arrive at clear, actionable consumer insights.

We must also recognize that the analysis of massive data sets will not benefit all consumers. Some may see a negative shift in their overall financial profile when additional data is considered. Others may simply not generate enough meaningful data points to enhance the amount of information that can be obtained about their histories and habits.

The Risks of Discrimination and Exploitation

The use of Big Data in consumer financial services is still in its infancy. While many legal and regulatory requirements already on the books address fundamental rights and protections for consumer lending and the use of individuals' information, these laws predate recent technological developments. The Fair Credit Reporting Act, and the Equal Credit Opportunity Act, as implemented by Regulation B, both address aspects of underwriting transparency, yet only partially address questions raised by the current use of Big Data. The aim of these regulations, respectively, is to ensure the transparency of decisioning by providing an applicant with a reason when a loan is not approved, and to prevent discriminatory results based on analytical inputs.

Yet grey areas remain, especially surrounding the use of nontraditional data drawn from social media or metadata predicated on associations between discrete individuals. Some Big Data platforms operating in international markets cannot be offered domestically without further clarity on legal guidelines for the use of such data. In particular, the use of Big Data to set financial product pricing and determine access presents a significant risk for unintended discrimination. The confluence of thousands of data points means there is an ever-present possibility of results that unwittingly discriminate against legally protected classes by resulting in disparate impacts, even if no algorithm is intentionally constructed to produce such an outcome. This risk makes it incumbent upon data scientists to not only construct models with the intention of avoiding disparate impacts, but to continually test their outputs against national and regional demographic data to ensure that their models are not straying into discriminatory territory.

Regulations that have not kept pace with modern Big Data capabilities present a challenge for government oversight, as well as for companies looking for firm legal guidelines as they build new platforms.[13] As data algorithms and machine-learning platforms become more complex, it becomes more difficult for providers to fulfill their requirements of ensuring that the use of decisioning data does not become opaque to consumers or produce results that leave consumers unable to identify inaccurate or harmful data inputs that impact results. Ambiguities in the law also leave the door open to companies using data in an exploitive fashion to target underserved consumers with financial products that have poor quality or models designed to ensnare consumers in unhealthy relationships using data methods that are not explicitly against current regulations.

As regulators turn attention toward the results of advances in Big Data, open communication between their offices and companies at this new frontier should be as common and frequent as possible. Each has much to learn from the other. Providers seek to understand how aspects of their Big Data platforms undefined in current law are viewed by those responsible for protecting consumers. Likewise, federal regulatory bodies can strengthen their ability to track and guide these new elements of consumer finance by gaining a nuanced understanding of how Big Data tools achieve insights and impact consumers.

---

[13] In January 2015, the Supreme Court weighed arguments related to the disparate impact of data collection on tax credit decisions in Texas Department of Housing and Community Affairs v. The Inclusive Communities Project, Inc., which may have implications for data modeling used in lending products.

Conclusion

The incorporation of Big Data analytics in financial services is becoming increasingly ubiquitous as the costs and barriers to manipulating massive data sets grow smaller. Now is the time for a candid and robust dialogue about how best to harness Big Data tools, while they are still in an early stage of market implementation, for the mutual benefit of providers and consumers.

It is our hope that the identification of key trends illustrating this potential will aid future innovation by encouraging companies already using Big Data, or exploring its use, to incorporate these practices into the products and services they create.

By tailoring products to fit consumer needs and inviting consumers to become partners in data validation, creation, and aggregation—while maintaining standards of safety and responsibility—Big Data can potentially improve the financial health of consumers and the success of providers on a large scale. When providers can utilize opportunities to build these values into their models, and then find routes to robust market adoption that reach consumers and feed back into their models, Big Data can serve as a financial technology tool that enhances the value of products and services by unlocking the data assets providers and consumers can offer each other.

Appendices

Appendix A: Glossary of Terms
Defining Big Data Buzzwords in a Financial Services Context

Term	Definition	Financial Services Relevance or Application
Alternative Data	Sources of data, such as utility bill payments, that inform assessments of debt repayment risk but are not used to inform traditional consumer credit scores.	Use of a broader set of tradelines enables risk assessment of some traditionally unscorable consumers and more accurate risk assessment of some scorable consumers.
Big Data	The collection and use of large data sets compatible with other data sets, both public and private, that can be broadly combined and distributed to identify patterns and expand insights.	As an increasing share of individual behavior is digitally encoded, Big Data can be used to better understand consumer financial profiles, habits, preferences, and predictions of future needs and behaviors.
Disparate Impact	Practices in lending, housing, and employment that have a disproportionately adverse impact on persons included in classes protected by statute under federal civil rights laws including race, color, religion, national origin, gender, and some disabilities.	If the application of data analytics results in discriminatory outcomes in the provision of rates, fees, or access to financial services, these practices are illegal, regardless of intent.
Machine Learning	Computer science capabilities in which nonlinear data analytics are used to build new data points and create iterative knowledge that is refined over time.	Machine learning technology holds potential to increase the accuracy and efficiency of consumer financial products with respect to design, targeting, and costs. Machine learning outcomes benefit from larger scale and longer term sources of data that aid in the identification of patterns and correlative insights.
Metavariable	An original data point generated through the input of existing data points into an algorithm or similar model of analysis.	The creation of metavariable data through the use of proprietary analysis tools allows data scientists to manipulate and assign values to patterns and insights identified in large data sets.
Mosaic Effect	The ability to identify specific and individualized data through the combination of discrete data sources which contain lower levels of specificity when viewed independently of each other.	The overlay of data sets from multiple sources enhances the ability of data analysis tools to achieve individualized insights with greater accuracy.
Zettabyte	1 zettabyte is equivilant to $1,000^{18}$ bytes, or 1 trillion gigabytes.	Between 2009 and 2013, the world wide web is estimated to have grown from one half a zettabyte to four zettabytes, vastly increasing the scale of data points available for analysis.

Appendix B: Featured Companies

Company	Consumer Finance Sector(s)	Big Data Focus	Financing	Big Data Trends	Location	Website
Acxiom	B2B analytics	Business analytics to guide B2B customer marketing strategies and consumer engagement.	Publicly Traded	Connecting Dots	Little Rock, AR	www.acxiom.com and https://aboutthedata.com
Affirm	Alternative installment lending	Risk decisioning analytics targeted at financing purchases by consumers identified as low risk but with low credit scores or thin/no credit files.	Venture	Getting Granular, Hitting the Target	San Francisco, CA	www.affirm.com
BillGuard	Financial management platform with fraud detection analytics	Provides personal financial security by analyzing payments activity in aggregated accounts to flag fraud from erroneous or malicious transactions via crowdsourced detection.	Series B	Connecting Dots	New York, NY and Tel Aviv, Israel	www.billguard.com
Capital One	Retail and investment bank	Banking, credit card, and auto loan provider for individual, small business, and commercial interests. Uses big data analytics to increase accuracy of consumer profiles and value of offers.	Publicly Traded	Getting Granular, Connecting Dots	McLean, VA	www.capitalone.com
Cignifi	Underwriting analytics platform	Analytic platform provides local lenders with consumer credit and marketing scores based on mobile phone behavior data in developing markets where credit scores are not typically available.	Series B	Getting Granular, Hitting the Target	Cambridge, MA	www.cignifi.com
Clarity	Consumer profiling, verification, underwriting and collections analytics	Credit reporting agency that analyzes non-traditional data sources with national coverage to help providers reduce lending risk to underserved, subprime, and thin/no file consumers.	n/a	Getting Granular	Clearwater, FL	www.clarityservices.com
Cognical	Underwriting analytics platform	Risk decisioning analytics targeted at financing online purchases by consumers identified as low risk but with low credit scores or thin/no credit files through Zibby product.	Seed	Getting Granular, Hitting the Target	New York, NY	www.cognical.com
CoreLogic	Consumer profiling, verification, underwriting, and collections analytics	Risk decisioning analytics for short-term loan providers, installment lenders, rental purchase companies, and other consumer finance businesses serving consumers with spotty or thin/no file.	Acquired by Cognizant Technology Solutions in 2011.	Getting Granular	Santa Ana, CA	www.corelogic.com
DemystData	Consumer profiling, verification, fraud protection, underwriting analytics	Brings together traditional and nontraditional data to optimize consumer profiles and predictions for financial institutions with the goal of greater financial inclusion.	Series A	Getting Granular	New York, NY	https://demystdata.com

Appendix B: Featured Companies

Company	Consumer Finance Sector(s)	Big Data Focus	Financing	Big Data Trends	Location	Website
Earnest	Underwriting analytics platform, alternative installment lending	Small-dollar lender offering installment loans with rates based on alternative analytics drawn in part from holistic borrower profile including education, work, and money management history.	Venture	Getting Granular, Opening Books	San Francisco, CA	www.meetearnest.com
Elevate	Alternative installment lending	Small-dollar lender offering installment loans with rates based on alternative analytics and demonstrated repayment behavior of borrowers.	Spun off from Think Finance in 2014.	Connecting Dots	Fort Worth, TX	www.elevate.com
Enova	Alternative installment lending	Enova's NetCredit small-dollar installment loan is based in part on alternative analytics and reports to credit agencies to help borrowers build credit scores.	Spun off from Cash America in 2014. Now publicly traded.	Getting Granular	Chicago, IL	www.enova.com
Experian	Consumer credit reporting agency	Gathers, maintains, and analyzes data on credit holders worldwide, creates credit scores and issues credit reports.	Publicly Traded	Getting Granular, Connecting Dots	Costa Mesa, CA	www.experian.com
FactorTrust	Underwriting analytics platform	Risk decisioning analytics for short-term or installment loans, credit and prepaid cards, automotive loans, retailers, wireless services, and other consumer finance businesses serving consumers with spotty or thin/no file.	n/a	Getting Granular	Roswell, GA	http://ws.factortrust.com
FairLoan	Alternative installment lending	Small-dollar installment lender providing loans through employers based in part on alternative analytics. Only available in California.	n/a	Getting Granular	San Francisco, CA	www.fairloanfinancial.com
FICO	Credit rating product	Assigns credit scores as a measure of consumer borrowing risk. Newest versions 8 and 9 seek to incorporate alternative analytics insights to rebalance the weight given to some traditional tradeline data in order to better identify risk in subprime credit holders.	Publicly Traded	Getting Granular, Hitting the Target	Minneapolis, MN	www.fico.com/en/
First Access	Underwriting analytics platform	Analytic platform provides local lenders with consumer credit and marketing scores based on mobile phone behavior data in developing markets where credit scores are not typically available.	n/a	Getting Granular, Opening Books, Hitting the Target	New York, NY	www.firstaccessmarket.com
Funding Circle	Underwriting analytics platform, small business marketplace lending	P2P small business lender that analyzes and scores risk for secured small business loans provided by peer investors. Rates and terms based in part on alternative analytics derived from applicant disclosure of financials. Entered U.S. market in 2013.	Series D	Connecting Dots, Opening Books	London, UK	www.fundingcircle.com

Appendix B: Featured Companies

Company	Consumer Finance Sector(s)	Big Data Focus	Financing	Big Data Trends	Location	Website
Happy Mango	Personal financial management analytics	Analytics platform that combines account information with alternative data to create personalized tracking tools, evaluate credit worthiness, and connect users with sources of credit that fit their profile and needs.	n/a	Opening Books	New York, NY	www.happymangocredit.com
HelloWallet	Personal financial management analytics	Software provided through employers using analytics and advice to enable employees to maximize their salaries and benefits for financial security and gain.	Acquired by Morningstar in 2014.	Getting Granular, Hitting the Target	Washington, DC	www.hellowallet.com
Kabbage	Underwriting analytics platform, alternative small business lending	Online lender providing small business capital with rates based on alternative analytics derived in part from sales volume and product review data.	Series D	Getting Granular, Opening Books, Hitting the Target	Atlanta, GA	www.kabbage.com
Kreditech	Underwriting analytics platform	Risk decisioning analytics for subsidiaries providing short-term loans to consumers with spotty or thin/no file. Not yet available in the U.S.	Series B	Getting Granular	Hamburg, Germany	www.kreditech.com
L2C	Underwriting analytics platform	Risk decisioning analytics for short-term loan providers, installment lenders, rental purchase companies, and other consumer finance businesses serving consumers with spotty or thin/no file.	Acquired by TransUnion in 2014.	Getting Granular	Atlanta, GA	www.l2c.com
Larky	Geolocation analytics and consumer solicitation platform	Mobile and web platform that relies on geolocation data to provide targeted deals and discount offers to customers through their bank or credit union cards.	Series A	Getting Granular	Ann Arbor, MI	http://get.larky.com
Lenddo	Underwriting analytics platform, alternative installment lending	Small-dollar lender offering installment loans with rates based on alternative analytics and demonstrated repayment behavior of borrowers. Not yet available in the U.S.	Series A	Connecting Dots, Opening Books	Hong Kong, China	www.lenddo.com
Lending Club	Underwriting analytics platform, personal and small business marketplace lending	P2P online lender that analyzes and scores risk for personal and small business loan applicants provided by peer investors. Rates and terms based in part on alternative analytics derived from non-traditional credit data.	Publicly Traded	Connecting Dots, Opening Books	San Francisco, CA	https://lendingclub.com
LendUp	Underwriting analytics platform, alternative installment lending	Small-dollar lender offering installment loans with rates based on alternative analytics and demonstrated repayment behavior of borrowers.	Series A	Connecting Dots, Hitting the Target	San Francisco, CA	www.lendup.com

Appendix B: Featured Companies

Company	Consumer Finance Sector(s)	Big Data Focus	Financing	Big Data Trends	Location	Website
LexisNexis	Consumer profiling, verification, underwriting analytics	LexisNexis Risk Solutions RiskView product suite provides prescreening, verification, and risk decisioning analytics, including to lenders targeting subprime and thin/no file consumers and those who are fully served.	Privately Owned	Getting Granular	Dayton, OH	www.lexisnexis.com/risk
MasterCard	Payments processor	Multinational financial services corporation facilitating electronic funds transfer for credit, debit, and prepaid cards and programs.	Publicly Traded	Getting Granular, Connecting Dots	Purchase, NY	www.mastercard.com
Mint	Personal financial management analytics	Account aggregation service that allows users to see credit card, bank, investment, and other accounts in an integrated format that analyzes spending, manages budgets, and generates leads to other financial services.	A product of Intuit, acquired in 2009.	Getting Granular	Mountain View, CA	www.mint.com
Neo	Alternative auto lending	Auto loan platform with rates based on alternative analytics targeting first time car buyers with limited or subprime credit history.	n/a	Getting Granular, Opening Books	Palo Alto, CA	https://neoverify.com
NetSpend	Prepaid card provider	Exploring opportunities to leverage data on consumer payments history to gain product insights and enhance positive consumer engagement.	A TSYS company, acquired in 2013.	Getting Granular	Austin, TX	www.netspend.com
OnDeck	Underwriting analytics platform, alternative small business lending	Online lender providing small business capital with rates based on alternative analytics derived in part from sales volume and product review data.	Publicly Traded	Getting Granular, Opening Books	Arlington, VA	www.ondeck.com
PayPal	Payments processor	Alternative electronic processor for online payments.	Recently spun off from eBay as a private company; IPO planned for 2015.	Getting Granular	San Jose, CA	www.paypal.com
Prosper Marketplace	Underwriting analytics platform, marketplace lending	P2P online lender that analyzes and scores risk for loan applicants provided by peer investors. Rates and terms based in part on alternative analytics derived from personal financial histories, social endorsements, and community affiliations.	Series H (10 rounds)	Connecting Dots, Opening Books	San Francisco, CA	www.prosper.com
Revolution-Credit	Underwriting analytics platform	Risk decisioning analytics to make automated lending decisions for small-dollar credit installment loans, in part through a platform offering financial education modules to prospective borrowers.	Seed	Connecting Dots, Hitting the Target	Irvine, CA	http://revolutioncredit.com
Square	Mobile payments and merchant services	Provides payments analytics to guide B2B customer marketing strategies and consumer engagement.	Series E	Getting Granular	San Francisco, CA	https://squareup.com

Appendix B: Featured Companies

Company	Consumer Finance Sector(s)	Big Data Focus	Financing	Big Data Trends	Location	Website
Think Finance	Underwriting analytics platform	Risk decisioning analytics to make automated lending decisions for small-dollar credit installment loans.	Series A	Getting Granular, Connecting Dots	Fort Worth, TX	www.thinkfinance.com
TransUnion	Consumer credit reporting agency	Gathers, maintains, and analyzes data on credit holders worldwide; creates credit scores; and issues credit reports. Acquired L2C in 2014.	Privately Owned	Getting Granular	Chicago, IL	www.transunion.com
VantageScore	Credit rating product	VantageScore 3.0 is a joint product of Equifax, Experian, and TransUnion that scores credit holders based on a combination of traditional tradeline and alternative data to better identify risk in subprime or thin file credit holders.	A joint product of Equifax, Experian, and TransUnion	Hitting the Target	Stamford, CT	www.vantagescore.com
Visa	Payments processor	Multinational financial services corporation facilitating electronic funds transfer for credit, debit, and prepaid cards and programs.	Publicly Traded	Connecting Dots	San Francisco, CA	http://usa.visa.com/index.jsp
VisualDNA	Consumer profiling, verification, underwriting and collections analytics	Combines big data with behavioral psychology insights to allow businesses to understand their customers financial risk and people to understand their own financial habits, both for subprime and thin/no file consumers and those who are fully served.	Series C Funding	Getting Granular, Opening Books	London, UK	www.visualdna.com
Yodlee	Personal financial management analytics	Account aggregation service that allows users to see credit card, bank, investment, and other accounts in an integrated format and manage bill payment, expense tracking, and other financial needs.	Publicly Traded	Getting Granular	Redwood City, CA	www.yodlee.com
ZestFinance	Underwriting analytics platform	Risk decisioning analytics to make automated lending decisions for small-dollar credit installment loans.	Series B Funding	Getting Granular	Los Angeles, CA	www.zestfinance.com

Morgan Stanley Disclaimer

This document is not a product of Morgan Stanley & Co. LLC, including the Morgan Stanley Research Department, and its affiliates (collectively "Morgan Stanley"). The information and opinions in these materials were prepared by the Center for Financial Services Innovation (CFSI) with strategic support from Morgan Stanley. These materials are solely for informational and discussion purposes and do not reflect the views of Morgan Stanley. Morgan Stanley does not undertake to update these materials and the conclusions discussed may change without notice. Morgan Stanley shall not in any way be liable for claims relating to these materials and makes no express or implied representations or warranties as to their accuracy or completeness or for statements or errors contained in, or omissions from, them. Past performance is not necessarily a guide to future performance. This material is not a solicitation of any offer to buy or sell any security, commodity or other financial instrument or to participate in any investment strategy. This material was not intended or written to be used, and it cannot be used by any taxpayer, for the purpose of avoiding penalties that may be imposed on the taxpayer under U.S. federal tax laws. Each taxpayer should seek advice based on the taxpayer's particular circumstances from an independent tax advisor. The trademarks and service marks contained herein are the property of their respective owners. Morgan Stanley is a global financial services firm that has pursued and may continue to pursue investment banking, investment management and proprietary investment positions related to the underbanked marketplace. Facts and views presented in these materials have not been reviewed by, and may not reflect information known to, professionals in Morgan Stanley business areas.