Aite PARTNER. ADVISOR. CATALYST.

Machine Learning: Fraud Is Now a Competitive Issue

OCTOBER 2017

Julie Conroy

This Aite Group report is provided compliments of: iovation®

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

IMPACT POINTS

- In this research study, sponsored by iovation, Aite Group interviewed 28 senior fraud and data analytics executives at 20 North American financial institutions (FIs) in August and September 2017.
- Sixty-five percent of FIs interviewed say the priority for investment in machine-learning (ML) analytics for fraud mitigation is very high and is a key area of investment. Another 35% say that the investment priority is moderate; while it's on the roadmap, other fraud solutions will take priority.
- Nearly every FI interviewed includes retail account takeover (ATO) among its top pain points. Application fraud comes in second, with 10 FIs saying that was a key pain point, followed by wholesale ATO and the specter of faster payments fraud.
- Ten percent of FIs interviewed are using ML analytics to help orchestrate authentication today, while another 30% are in the process of implementing analytically driven orchestration of authentication.
- Eighty percent of the FIs interviewed have in-house data scientists dedicated to the fraud team, although the quantity of resources available to fraud varies widely.
- Forty percent of the FIs interviewed have an ML-enabling platform deployed in production, while another 10% have one or more proofs of concept (POCs) underway. Twenty percent say that the deployment of an enabling platform is on their one- to two-year roadmap, while one in five of the FIs interviewed have no plans to deploy an enabling platform in the next couple years.
- Effective fraud prevention is increasingly a competitive issue for FIs. Those that are early adopters of advanced analytics will be able to do more than reduce fraud; the associated improvements to the customer experience give them a decided edge over their competitors that lag in these investments. Data is the new currency, and creating intelligence from data at scale requires ML technology.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

INTRODUCTION

Time is money when it comes to fighting fraud. Organized crime rings, fueled with billions of compromised data records, are systematically and methodically targeting the financial services value chain with sophisticated card fraud, application fraud, and ATO attacks. The volume of these attacks continues to increase, since there is very little in the way of adverse consequences (i.e., jail time).

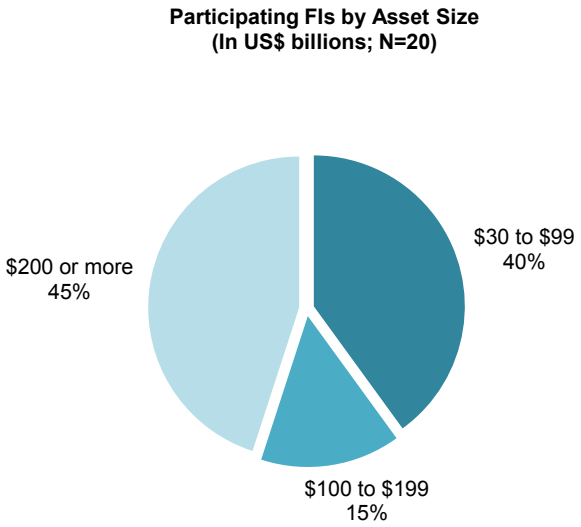
Another key challenge for fraud executives is that even as the threat environment continues to escalate, FIs are under intense competitive pressure to make the banking experience easier and frictionless. In the face of this seemingly contradictory set of mandates, many FIs are turning to ML analytics as part of their solution set. Leveraging the vast amount of customer data at their disposal and applying advanced ML techniques, FIs are able to create insights and intelligence that achieve the dual goals of better fraud mitigation as well as customer experience improvements.

The path is not without its challenges, however, and many FIs are in the early stages of this journey. This Impact Report will help FI executives benchmark their progress in moving toward ML analytics against their peers and better understand the considerations and lessons learned along the way.

METHODOLOGY

In this research study, sponsored by iovation, Aite Group interviewed 28 senior fraud and data analytics executives at 20 North American FIs in August and September 2017 (Figure 1).

Figure 1: Participating FIs by Asset Size



Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

ML FOR FRAUD MITIGATION

To understand the potential of ML, it's important to understand what the technology is and does. All too often, the term "ML" is used interchangeably with "artificial intelligence"; however, ML is actually a subset of artificial intelligence. ML encompasses analytics techniques that can identify patterns of behavior through iterative optimization.[1] These are deployed by FIs in three ways:

- **Analytics toolkits:** Programming languages such as R, SAS, or Python used by in-house data scientists to construct homegrown models
- **Enabling platforms:** An analytics engine that enables businesses to deploy ML models at scale across multiple use cases
- **Embedded analytics:** An embedded part of a point solution used to enhance vendors' scoring algorithms

The use of ML analytics for fraud prevention is rapidly gaining traction in financial services. Fraud is moving too fast for the legacy approaches that rely on rules and annual model refreshes to be effective. FIs need advanced analytics technology that can evolve rapidly and keep pace with the progression of fraud attacks so they can prevent losses while maintaining a positive customer experience.

Detection strategies are shifting from analyzing siloed transactional activity to instead making better use of data and analytics to enable a holistic understanding of the customer's activity. By bringing together cross-product and cross-channel data, and by applying nimble ML analytics, businesses can understand the context of transactions and make better decisions. FIs are bringing these analytics to bear across the enterprise, with use cases ranging from cards to digital banking and from authentication to faster payments.

PAIN POINTS

When asked what types of fraud are garnering the highest priority for investment over the next couple of years, nearly every FI interviewed included retail ATO. This is a sharp contrast to a 2015 Aite Group study, in which the majority of large U.S. FIs interviewed said that ATO losses were largely flat.[2] More than 9 billion data records have been lost or stolen since 2013, containing personally identifiable information (PII), credentials, and/or payment card data.[3] The ready availability of data makes ATO and application fraud much easier for the organized crime rings responsible for the attacks.

1. For a more detailed explanation of ML, see Aite Group's report *Machine Learning for Fraud Mitigation: The Substance Behind the Buzz*, April 2017.
2. See Aite Group's report *Digital-Channel Fraud Mitigation: The Mobile Force Awakens*, June 2015.
3. Breach Level Index, accessed on September 15, 2017, http://breachlevelindex.com.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

Application fraud came in second, with 10 FIs saying that was a key pain point, followed by wholesale ATO and the specter of faster payments fraud, which, as many executives point out, goes hand in hand with ATO (Figure 2). Faster payments is a double whammy from a fraud mitigation perspective. Fraud executives expect rapid volume growth, and the transactions themselves move in real time with no repudiation rights once settled, which emphasizes the need for effective real-time detection capabilities. A few FIs also note that check fraud continues to be problematic, particularly in the wake of the U.S. migration to EMV, although only one named it a key pain point garnering investment.

Figure 2: Fraud Pain Points

Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

Cross-channel fraud was often cited in the interviews as a key attack vector, with the contact center often playing a prominent role. Here are just a few of the actual fraud incidents cited in the interviews that caused recent losses:

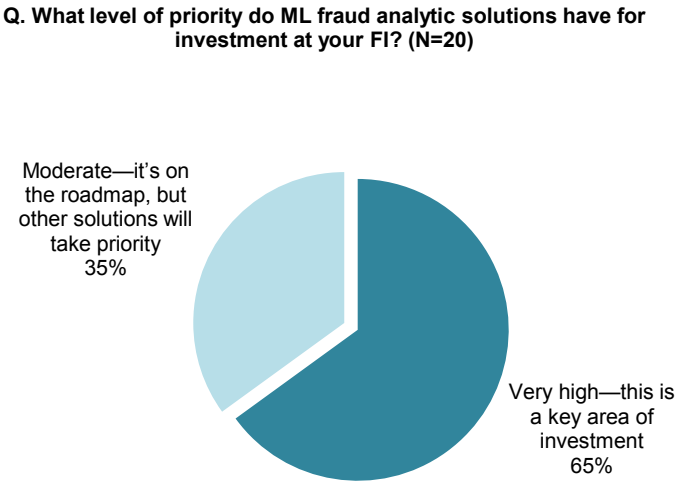
- Criminals logged into a consumer's online-banking account and captured the check images, then used those to perpetrate counterfeit check fraud.
- Compromised cards were used to authenticate to online banking, then the criminals transferred out funds, opened new accounts, and/or added authorized users to secure new cards.
- Fraudsters transferred funds from a home equity line of credit (HELOC) to a new checking account, then clicked over to the check printer interface, ordered new checks, and used those new checks to spend all of the illicitly obtained HELOC funds. This particular fraud entailed four platforms and three account types, so it was nearly impossible to detect using the FI's legacy systems.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

USE OF ML BY NORTH AMERICAN FINANCIAL INSTITUTIONS

Sixty-five percent of FIs interviewed say the investment priority for ML technologies for fraud use cases is very high and a key area of investment. An executive at one large FI says that all of the technologies that the fraud group is deploying must have elements of ML, or they won't get prioritized. Another executive says that his FI is focused on investing in ML to remove friction and "take the bars off the front door." Thirty-five percent of respondents say the investment priority is moderate; while it's on the roadmap, other fraud solutions will take priority. A correlation exists between the size of the FI and the priority for fraud-related ML investments; all of the FIs that indicate that ML is a moderate priority have less than US$200 billion in assets. Notably, none of the FI executives say that the priority is low (Figure 3).

Figure 3: Priority for Investment in ML Fraud Analytics



Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

IN-HOUSE DATA SCIENCE RESOURCES

Eighty percent of the FIs interviewed have in-house data scientists dedicated to their fraud team. The quantity of resources available to fraud groups varies widely, however, as shown in Table A, as does the level of sophistication of the internal fraud analytics programs. Some of the FIs interviewed are in the earliest stages of the journey, having just brought fraud-dedicated data scientists on board within the past year. Others are quite advanced; one executive from a large FI stated that they model almost every strategy they apply to fraud mitigation.

In most cases, only a subset of the team is employing ML techniques. In the larger and/or more analytically progressive FIs, many fraud groups also tap into resources at enterprise analytics departments, in which resources are dedicated to the fraud team. Internal data science resources are working on a wide variety of use cases—everything from check fraud to card fraud to ATO.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

Table A: In-House Data Scientists Dedicated to Fraud

FI asset size (In US$)	Number of resources	Additional planned over next one to two years
$30 billion to $99 billion	0	0
	0	1 to 2
	2 (shared with Infosecurity)	1 to 3
	3	2
	4	0
	8 to 10	0
	Approximately 9	0
	19	1 to 2
$100 billion to $199 billion	0	0
	5	0
	26	0
$200 billion or more	0	Hoping for 5 to 6
	9 (digital channels only)	Will grow, quantity unknown
	3	3 to 4
	6	Will grow, quantity unknown
	20	5
	Approximately 30	0
	Approximately 30	0
	More than 40	Will grow, quantity unknown
	Approximately 50	Will increase substantially

Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

Most of the FIs plan to incrementally add talent to their data science teams, while one plans to make substantial increases. Hiring talent is a challenge, however; the data science skill set is in high demand, and a number of executives say finding the right talent is difficult. One FI executive is concerned that universities are churning out data scientists without the right skill set; this FI is looking for people well-versed in Python and NoSQL, who are quite scarce.

ENABLING PLATFORMS

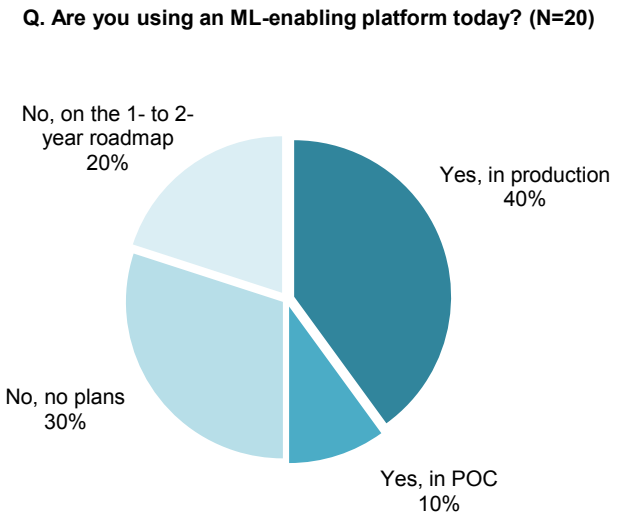
A few of the FI executives interviewed say they are in the process of trying to make the build-versus-buy decision: Do they try to build out homegrown ML models using in-house resources or

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

invest in enabling-platform technology? While many of the FIs with more than US$200 billion in assets are doing both, relatively smaller FIs with more finite resources are having to make a choice. One of the FIs with less than US$100 billion in assets is proceeding aggressively down the build path, with 19 data science resources already dedicated to the fraud team, while the majority of the other FIs in this size range are leaning toward buying enabling platform solutions.

Forty percent of the FIs interviewed have an ML-enabling platform deployed in production, while another 10% have one or more POCs underway. Twenty percent say that the deployment of an enabling platform is on their one- to two-year roadmap, while 30% of the FIs interviewed have no plans to deploy an enabling platform in the next couple years (Figure 4). Enabling platform vendors that are in production with the interviewed FIs include BAE Systems, Feedzai, FICO, Nice Actimize, SAS, and Simility. FIs planning to implement are considering all of these vendors, as well as Featurespace, IBM, and ThetaRay.

Figure 4: Use of ML Enabling Platforms Among Large North American FIs



Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

Many of the FIs plan to leverage the enabling platforms for multiple use cases. Table B discusses the use cases and deployment stages for those FIs engaging with the enabling platforms.

Table B: Enabling Platform Use Cases

FI asset size (In US$)	Use case and deployment stage
$30 billion to $99 billion	Application fraud was the first use case, but the FI is not wholly satisfied with current vendor. The FI plans to do a POC with another enabling platform in 2018 but has not yet selected a POC vendor.
	The FI is finishing a POC with an enabling platform; the three use cases are application fraud, online banking, and check, Automated Clearing House (ACH), and debit transactional analysis.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

	The FI is in production with an enabling platform for online banking and wires. Check fraud is the next use case; the FI will also evaluate using the platform for payment card fraud.
$100 billion to $199 billion	The first use case will be faster payments; the vendor has been selected.
	ACH and wire analysis is in production, ATM and debit analytics is in the implementation phase, and credit card and faster payments will be the next use cases.
$200 billion or more	Check and deposit fraud was the first use case. The executive interviewed said that there was so much low-hanging fruit given the high false positive rates with legacy check fraud solutions that it made a good test case. The platform will be extended to ACH next, and a cross-channel project to evaluate ATM, branch, and call center activity will kick off in early 2018. The FI expects that the extension of the platform to all fraud use cases will be completed in 2021 or 2022.
	The FI is planning to do a POC with an enabling platform for card-not-present fraud. Faster payments is the next use case on the roadmap; the FI wants to have faster payments in production by late 2018.
	The FI will start with an enterprise approach to customer risk assessment; it is currently in the vendor-selection process.
	Digital banking is in production, and the FI is working on orchestration of authentication, although this is not yet implemented.
	The FI has an enabling platform in production for application fraud and demand deposit account transactional analytics.
	Payment card fraud was the first use case; it will expand to digital banking.
	The FI is doing multiple POCs with multiple providers for a variety of use cases.

Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

EMBEDDED ANALYTICS

Aite Group also asked FIs about their use of embedded analytics vendors, i.e., those vendors that use ML to enhance the scoring of their own point solutions. The interviewees indicate that they use a wide variety of these vendors for everything from application fraud detection to device identity. However, they also say that the term “machine learning” is often applied by these vendors inappropriately, and that while the vendors may say that their scoring is derived from ML algorithms, often that’s not technically the case upon close examination. The overuse of this term is leading to a fair amount of confusion in the market.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

ORCHESTRATION OF AUTHENTICATION

The widespread nature of data compromise, combined with consumers' penchant for oversharing on social media, makes it very difficult for FIs to authenticate their customers. Reams of data are readily available on the dark web, and criminals also have social engineering down to a science. During the reconnaissance phase of the attack, criminals often study the consumers' genuine transactional behavior; then when they attack, they try to emulate that behavior so they don't trigger red flags in the FI's monitoring systems. Fraudsters also have the scripting down to a science, so they sound totally natural when they're calling into the contact center with social engineering attacks.

The concept of orchestrating authentication is one that has been gathering momentum for some time in financial services.[4] Today, authentication is typically a one-size-fits-all activity, with stepped-up authenticators applied universally, regardless of the context of the transaction. For example, any time a retail banking customer tries to change his or her address online, the customer must answer a knowledge-based authentication question, or any time a commercial customer tries to send a wire over a certain dollar amount, the user must input a one-time password.

Orchestration of authentication seeks to better analyze the customer's usual behavior patterns as well as the context of the transaction. It does away with the one-size-fits-all approach and instead only inserts the friction of stepped-up authentication when necessary, i.e., when the analytics flag that the context of the transaction is unusual.

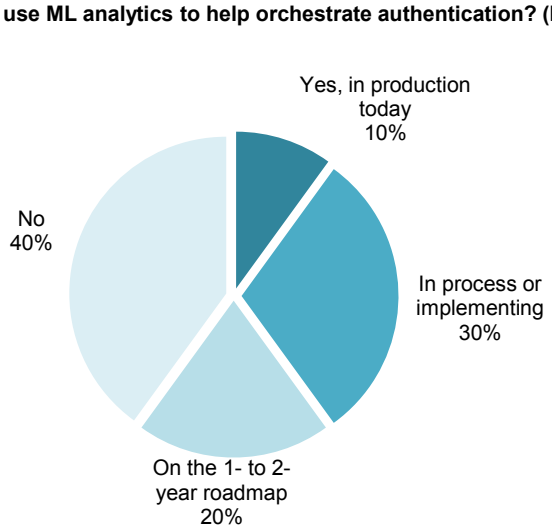
This degree of customization is by no means easy, however, and the practice is still nascent. A broad spectrum of capability and focus exists among the FIs interviewed. While a couple of FIs already have early versions of orchestration in production, other FIs are years from making this a reality. In one FI executive's words, "We're a *long* way away from that. We may end up there, but it'll be years. We're very tactical right now; we're having trouble just rolling out one-time password as stepped-up authentication for a new device."

Ten percent of FIs interviewed are using ML analytics to orchestrate authentication today, while another 30% are in the process of implementing analytically driven orchestration (Figure 5).

4. See Aite Group's report *Top 10 Trends in Retail Banking & Payments, 2016: The Quest to Reduce Friction*, January 2016.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

Figure 5: The Use of ML to Orchestrate Authentication



Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

Both of the FIs enabling orchestration today are using in-house data analytics resources. Of those that are implementing the capability, three FIs are using internal resources, one is leveraging Transmit Security's authentication hub in conjunction with Nice Actimize's risk engine, and two FIs are relying on enabling platforms. While all of the FIs are using the digital channels as their starting place for orchestration of authentication, the vision is to take the approach into other channels as well, such as branch, contact center, and ATM.

While the concept of orchestration and peeling back friction sounds great, it's not without its challenges. When the authentication is one-size-fits-all, customers over time grow accustomed to the stepped-up prompts and are preconditioned to participate. As these triggers become variable and customer authentication becomes more random, FIs are concerned that this will confuse customers, result in increases in call volume, and possibly even make customers feel unprotected for certain types of transactions. In order to make the concept work, it needs to go hand in hand with a strong customer education program.

MODELING TECHNIQUES

There are a few essential ways in which analytics models are trained:

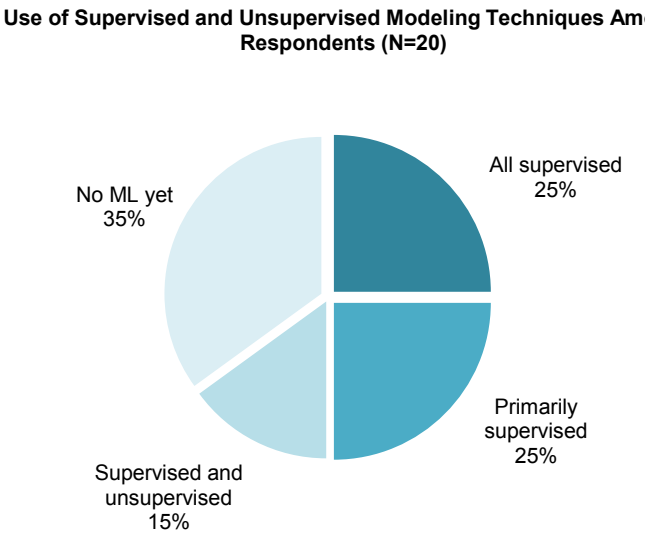
- **Supervised learning:** Supervised models are created using labeled training data, i.e., data that has been specifically identified as fraudulent or good transactions. This approach is ideal to use when a good amount of historical data is available to train the analytics. As a result, supervised models typically have lower false positive rates than do unsupervised ones.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

- **Unsupervised learning:** Unsupervised models do not have the benefit of the labeled training data and are useful when the organization doesn't have a lot of history to use for modeling (e.g., with new payment methods, such as faster payments). The answers are not known in advance, so the system is learning to detect outliers based on their similarity to prior transactions. Unsupervised models are more prone to false positives, since a portion of good customers will inevitably have outlier characteristics.
- **Semisupervised learning:** Semisupervised learning falls somewhere in between. It leverages both labeled and unlabeled training data to inform the models, and, as is to be expected, the false positives rate also tends to fall somewhere in between.

One of the misconceptions about ML is that many think that ML is synonymous with unsupervised learning, which is not at all the case. In fact, the majority of the production installations of ML among the FIs interviewed are supervised, often with dynamic retraining. Half of the respondents are using entirely supervised or primarily supervised techniques. Another 15% are using a combination of supervised and unsupervised models (Figure 6).

Figure 6: Supervised vs. Unsupervised Modeling



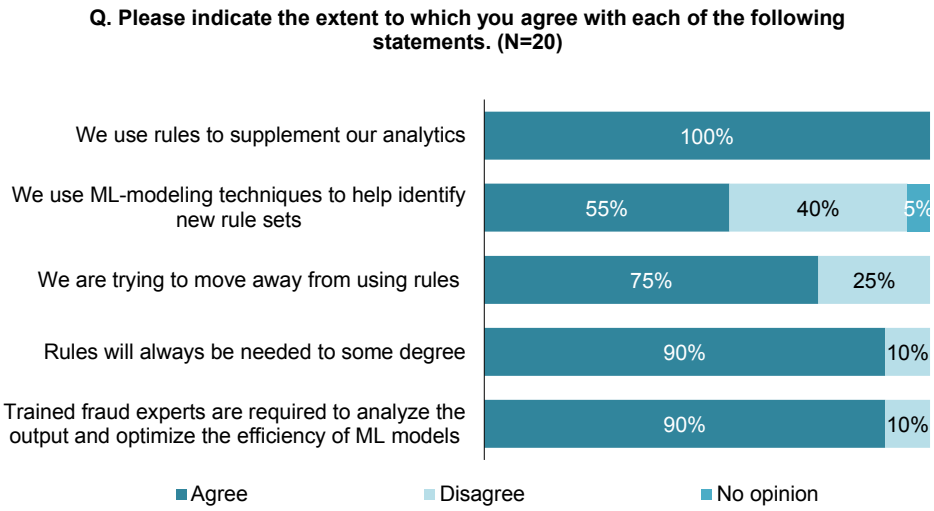
Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

THE ROLE OF RULES AND HUMANS

Another popularly held belief about ML is that it eliminates the need for rules and humans. Aite Group tested this concept with the interviewees. All of the FIs interviewed currently use rules to supplement their analytics. Fifty-five percent of those interviewed find value in using ML techniques to help identify new rule sets. Forty percent of the FIs interviewed do not use ML techniques to identify rule sets, although most of this group also do not have ML-modeling capabilities within the fraud group. One of the FIs interviewed with an advanced ML competency does not believe that using ML to find new rule sets is a good approach (Figure 7).

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

Figure 7: The Role of Rules and Humans



Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

Seventy-five percent of FIs are trying to move away from using rules, while 90% of FIs believe that rules will always be needed to some degree. FIs are trying to reduce the volume of rules and the ad hoc nature of the rules, but most believe the ability to respond to fraud quickly with a new rule is essential. One example that was given is when the FI has a VIP customer worth US$25 million who is about to perform an out-of-pattern transaction, and the FI wants to ensure that individual does not experience a false decline—rules are very useful in that scenario. In the words of one FI executive, "Models have a time-consuming model validation process. Rules can be put in place in five minutes." Another executive further elaborated his belief that it's naive to assume that machines can capture everything, unless we get to a point in which models are updating every second. The two executives who dissented contend that while rules are needed for the foreseeable future, technology will progress to the point in which rules are no longer needed.

While the majority of FIs believe that there will always be a need for rules, 70% of FIs are trying to reduce their reliance on rules through more advanced analytics. In many environments, a layered patchwork of episodic, flash-fraud rules look for specific patterns of bad behavior. In some cases, these rules have been put in place in a piecemeal fashion over a period of years, and it's very difficult to discern those that are still performing well and are valuable versus those that are causing more noise than value.

People are also an important part of the advanced analytics equation. Eighty-five percent of the FIs interviewed believe that trained fraud experts are required to analyze the output and optimize the efficiency of ML systems. "If you don't understand your data and the context for your data, you could make some boneheaded mistakes," is the view of one FI executive. One of the executives who disagrees believes that FIs will need fewer and fewer people over time as the technology continues to mature and as more automation takes hold in the form of algorithms and bots. Another posits that it's not fraud analysts who are required to do this but rather data

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

scientists who are familiar with the fraud use cases. One of the FIs with an emerging fraud data science practice is adopting a hybrid approach. Its data science team is composed of one fraud expert (who is not yet a modeler), one who is a modeler (and is not a fraud expert), and one individual who has a bit of each skill set.

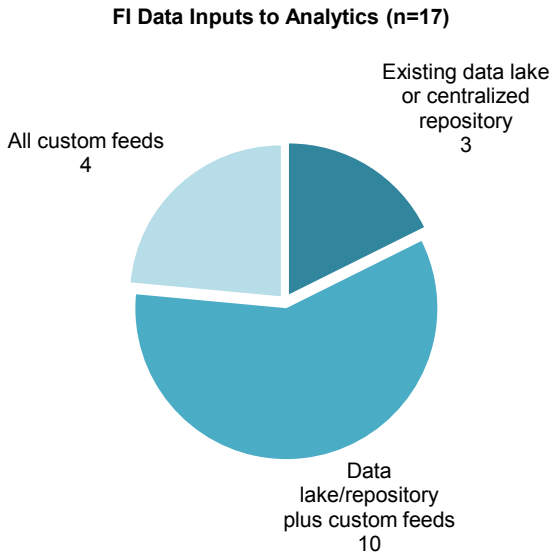
Another human element needs to be considered as FIs are migrating to ML, and adjustment is required to trust in the analytics and automation. One FI interviewed used to place outbound calls to the customer for every wire transfer over a certain dollar threshold. The analytics enabled the FI to eliminate that process, but the FI ran the analytics side by side with the manual process for some time before the FI executives had the confidence to forgo the manual intervention entirely.

DATA INPUTS

The vast amount of data now available to inform ML analytics is a key reason for its success. As the concept of identity has expanded to also encompass consumers' digital identity, the application of ML analytics is essential to making sense of the data and detecting fraud while minimizing false positives.

Corralling that data can be a challenge, however. Of the FIs that are using some form of ML in their fraud shop, three are tapping into an existing data lake or central data repository, 10 are using a combination of an enterprise data lake and custom feeds, and four are feeding their analytics entirely with custom feeds (Figure 8).

Figure 8: Data Inputs



Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

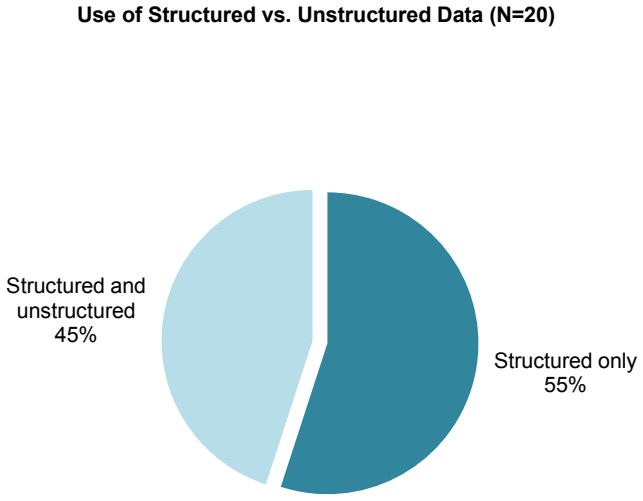
The data journey is slightly different for all of the FIs interviewed, as highlighted by the next steps for six of the FIs:

- Four FIs are in the process of building a bespoke data lake for the fraud team, because the timeliness of the enterprise data lake does not meet the near-real-time needs of the fraud team.
- One FI is working at the enterprise level toward streaming as much real-time data as possible.
- One FI is also looking at how it can leverage its centralized Splunk environments and apply analytics.

STRUCTURED VS. UNSTRUCTURED DATA

Most ML models are capable of ingesting and processing both structured and unstructured data sources (although a couple of the FI executives interviewed said that their enabling platform is incapable of handling unstructured data at this time). Structured data is that which is available in a clearly defined database, whereas unstructured data is that which is extracted from free-form documents and data streams (e.g., a PDF invoice supporting a trade finance transaction or recorded calls in contact centers). Fifty-five percent of the FIs interviewed are only using structured data today, while 45% are using both structured and unstructured data (Figure 9). All but one of the FI executives using both data forms say that the majority of their inputs are structured, while one FI executive says that the majority of what feeds its ML models is unstructured. The vast majority of FI executives indicate that incorporating more unstructured data is one of the goals of their data journey.

Figure 9: The Use of Structured vs. Unstructured Data



Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

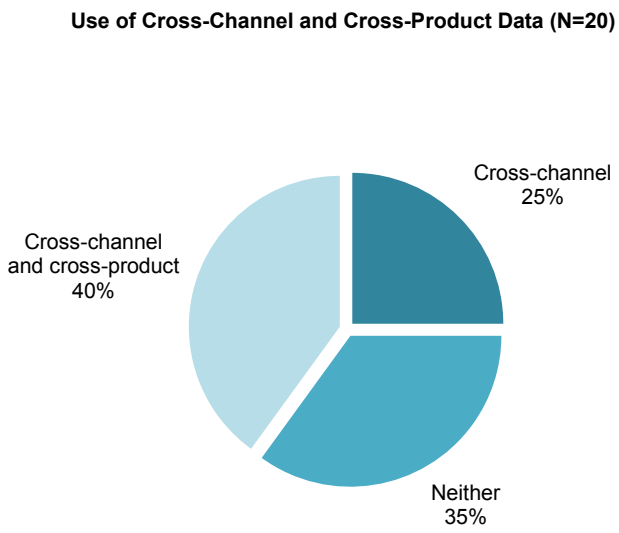
101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

CROSS-CHANNEL AND CROSS-PRODUCT DATA

Since cross-channel ATO is so problematic, it stands to reason that both transactional and non-monetary cross-channel data is helpful in detection. Cross-product data also helps in better understanding the customer's behavior patterns and establishing context for transactions; one of the FI executives interviewed says that some of their best data comes from the insurance side of the house, and they've learned a lot about fraudulent behaviors in the contact center by analyzing fraudulent claims calls. Cross-product data can be difficult to harness, however, since the data architectures and entity structures are often quite heterogeneous and difficult to reconcile.

Forty percent of FIs interviewed have some level of cross-channel and cross-product data informing their analytics today; 25% bring in cross-channel data, and 35% of FIs do neither (Figure 10). Device data attributes and contact center data are some of the key cross-channel data that FIs are bringing together today, or that are in some phase of implementation. The vast majority of FI executives interviewed indicate that gathering and optimizing data inputs will be an ongoing process for quite some time.

Figure 10: The Use of Cross-Channel and Cross-Product Data



Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

KEY PERFORMANCE INDICATORS (KPIS) AND PROOF POINTS

In the words of Peter Drucker, "If you can't measure it, you can't improve it." The FI executives interviewed use a number of KPIs to evaluate the performance of their fraud analytics solutions, as shown in Table C. The KPIs varied somewhat from FI to FI, as did naming conventions, which can make it difficult for FIs to benchmark their performance.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

Table C: KPIs for ML Solutions

KPI	Definition
False positive rate	Number of good accounts flagged for each true fraud detected
True fraud rate/hit rate/detection rate	Detected fraud dollars divided by the total fraud dollars lost
Alert rate	The ratio of the number of true fraud hits over the number of total alerts
Precision	Measures the accuracy of identified instances that are relevant—e.g., if an application fraud solution flags 10 applications and one is a true fraudster while the remaining nine are good customers, it will have a 10% precision ratio
Receiver operating characteristic (ROC)	Graphical curve that plots the true positive rate against the false positive rate
Proactive detection	The percentage of fraud the FI proactively detects versus the percentage of fraud customers alert them to
R values	Also known as coefficient correlation, the more closely the value is to one, the more closely the two variables are related
Challenge rate	Percentage of time that customers get prompted for stepped-up authentication
Queue penetration rate	Makes sure the operations team doesn't get flooded with alerts they don't have the bandwidth to address
Analyst capacity avoidance	Not having to hire a new full-time employee
Prevented loss	Fraud loss dollars prevented by the detection system
Replacement of manual countermeasures	Number of manual processes that the analytics replaces
Fraud to sales ratio/ loss exposure	Fraud dollars lost as a percentage of total exposure
Yield	Loss exposure as percentage of exposure combined with false positive ratio

Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

One of the many benefits of ML is that the models can be iterated and retrained quickly. As one FI started down this path, it first tried big overarching models and found that rules still worked better. Next, it tried using ML for highly targeted models focusing on specific fraud use cases (e.g., ACH ATO) and found these to be much better-performing, especially with dynamic retraining logic. The FI executives interviewed shared a handful of proof points, which illustrate some of the value gleaned from the application of ML (Table D).

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

Table D: ML Proof Points

Fraud use case	Performance improvement
Check	Improvement includes a 10% to 15% decrease in net loss, with substantial decrease in false positives.
Check	It has a 5-to-1 false positive rate, a significant improvement upon the prior state, which was over 20-to-1. Two people created the model, saving US$10,000 per day.
Check	Legacy countermeasures had a hit rate of 13%; ML platform is achieving a 76% hit rate.
ACH ATO	Models are achieving the same prevented loss month over month; rules seldom achieve this.
Investment ATO	Rules had a 14-to-1 false positive rate; model is 3-to-1, with only about eight alerts per day.

Source: Aite Group interviews with 28 executives at 20 North American FIs, August and September 2017

While one FI executive said his deployment was still early enough that he didn't have firm metrics to share, he did say that his FI had already enabled new transactional capability that had previously been deemed too high risk, thanks to the superior detection capabilities enabled by the analytics.

THE FLY IN THE OINTMENT: REGULATORS

ML presents substantial opportunity to improve upon legacy methods of detection and give FIs a fighting chance at stemming rising fraud losses while improving the user experience. The path forward is not without obstacles, however. As if the internal challenges of budget, data access, and IT resource constraints aren't enough, the specter of regulatory scrutiny also looms large. When asked whether they had concerns about potential regulatory issues related to fraud scoring transactions using ML, the response from interviewees was a unanimous yes. The crux of the concern is around two issues: privacy and model risk management (MRM).

PRIVACY

One of the global FIs interviewed classified privacy as one of the most difficult hurdles fraud executives have to clear internally. One of the challenges is that both internal teams (e.g., model governance and fair lending) and external regulators are struggling to understand the expanded concept of digital identity. In the traditional world of PII, models often had just a handful of input parameters. Now, with the wealth of digital identity data available to inform risk analytics, anywhere from 50 to as many as a few hundred inputs is not unusual. While these parameters can be extremely informative from a fraud detection perspective, it can be challenging to educate internal and external regulators about their use.

In overseas markets, the challenge is multiplied, since there are wildly different regulatory regimes—regulators can be either hands off or immensely prescriptive. Another big consideration for global banks is the handling of PII and how the bank can pass data from one data center to another, as well as emerging regulations such as the E.U. General Data Protection

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

Regulation (GDPR). While GDPR was not top of mind among most of the FIs interviewed, it should be, since it applies to any FI that has ex-pat clientele.

MODEL RISK MANAGEMENT

Regulators want to know that banks have a clear understanding of how their models work and how any changes impact detection and false positives. To that end, regulators as well as internal model governance teams are increasingly requiring extensive documentation of how the models function and the impact of any changes made to the models. This has long been a burden that the credit risk and anti-money laundering teams within FIs are accustomed to shouldering, but fraud teams are now feeling the pressure as well.

The genesis of the emphasis on model risk management came in the wake of the 2008 financial crisis, when regulators wanted to ensure that FIs were attentive to the possible adverse consequences of decisions based on models that are incorrect or misused.[5] The initial MRM focus was on credit risk models, which were a trigger for the financial crisis, and anti-money laundering analytics quickly fell under regulatory scrutiny. Within the past few years, regulators have extended their focus to fraud teams as well.

While the goals of MRM are sound, the implementation is costly for FIs. One large bank's internal compliance interpretation is that fraud models have to be validated every three years. This bank currently has 40 to 45 fraud models, which translates to 15 validations per year, each of which takes three to four months. This bank had to dedicate one to two internal resources as well as one to two external resources just to the annual validation process. This doesn't include the models used by external vendors using embedded analytics (which the FI executive hopes don't get added to scope, or that could substantially add to the MRM burden). Another large regional bank has seen its internal model governance team grow from three to 15 people over the past few years. It also had to hire a fraud headcount solely dedicated to the MRM process.

In addition to the labor overhead associated with MRM, FIs are apprehensive about a number of areas in the intersection of regulators and ML analytics:

- **Ambiguity:** North American regulators haven't clarified their expectations with regard to the use of ML. One FI executive believes this uncertainty needs to be resolved quickly, because a lot of innovation is at stake.
- **Unsupervised models:** Unsupervised models can be quite valuable in fraud, since it is such a volatile arena and trends change so quickly. They are often more effective at finding nuance that the human-supervised model won't find. Clearly explaining the causality in these models can be impossible, however, which runs contrary to model governance expectations, and there's not a lot of precedent to work with. One FI executive fears that the industry faces an "unsurmountable barrier" with regulators on this front.

5. "Guidance on Model Risk Management," Federal Reserve, April 2011, accessed on September 15, 2017, https://www.federalreserve.gov/supervisionreg/srletters/sr1107.htm.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

- **Vulnerable variables:** Model variables that include indications of age, class, and race will be highly scrutinized by regulators to ensure they are not resulting in biased outcomes. While there is certainly the potential that those variables could be considered to create adverse outcomes, the unfortunate aspect is that these variables are often highly predictive. One example given by one FI executive is that 30% of all their mobile remote deposit fraud emanates from a finite set of zip codes. Unfortunately, they can't use those zip codes as a variable in their analytical models due to redlining concerns.

The FI executives interviewed are seeing varying levels of scrutiny from their external regulators on this front. As is often the case, the implementation of regulations can vary from region to region based on both the central regulator and the local examiners. One FI executive indicates that they have been under intense MRM pressure from their regulator, the Federal Reserve, for the past two years. The FI not only had to provide complete documentation for all internally created models but also had to provide documentation for all external vendor models. In contrast, another executive, whose large FI is regulated by the Office of the Comptroller of the Currency, says they've seen very little demand for documentation of their fraud models. In almost every FI represented by interviewees, internal model governance teams are requiring extensive documentation of homegrown models and those facilitated by enabling platforms, although the burden of documentation does not tend to be as high for embedded analytics models when dealing with internal model governance teams.

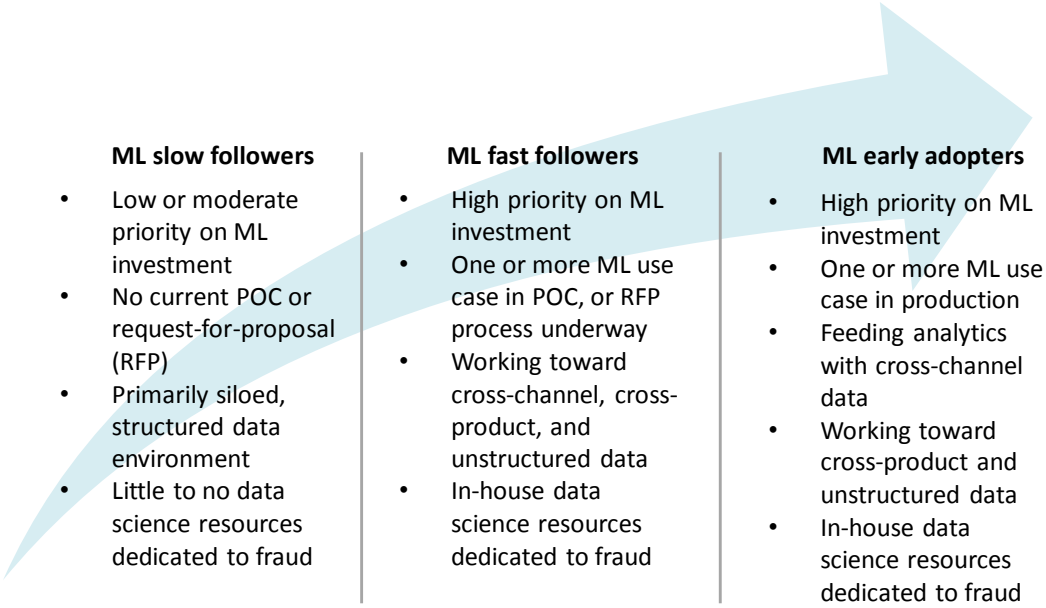
Not all executives are entirely pessimistic about the regulatory landscape with regard to ML. One FI is applying the guiding tenet in the application of ML that if there is a negative action that results from the analytics, it needs to provide a clear and easy opportunity for customers to resolve it. This a best practice from a customer experience standpoint, and the executive believes this will also keep regulators happy. And while MRM certainly slows down model deployment and limits some of the options available to FIs as they deploy ML, a silver lining does exist. One of the FIs interviewed discovered during the model documentation process that two of their models were broken. One had been in place for five years, and nobody had previously figured out that it wasn't working.

WHAT WILL THE FUTURE BRING?

The FIs interviewed are at many different stages in the ML journey. As with many emerging technology trends, there are early adopters, fast followers, and slow followers of ML (Figure 11). Some FIs have made substantial investments already and plan to continue this pace, while others haven't even begun. With ML, however, the stakes are high for FIs that delay. Early adopters are already seeing significant benefits in terms of their ability to detect fraud, improve operational efficiency, enable new transactional activity, and remove friction from the customer experience. The longer the ML slow followers delay, the greater the customer experience divide will be between their FI and the ML early adopters.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

Figure 11: ML Maturity Model



Source: Aite Group

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

CONCLUSION

Effective fraud prevention is increasingly a competitive issue for FIs. Early adopters of advanced analytics are able to increase their fraud detection, and the associated improvements to the customer experience give them a decided edge over their competitors that lag in these investments. Here are a few recommendations for FIs as they are evaluating their ML investments:

- **Understand your inputs.** Good data is essential to good analytics outcomes but is not always easy to come by. Ensure your project includes considerations for getting the depth and timeliness of data needed. Cross-channel and cross-product data is particularly valuable and also can be some of the most challenging to harness.
- **Be mindful of regulators.** A high degree of uncertainty still exists around regulators' expectations. Follow sound model governance practices, and if you leverage vulnerable variables, make sure that you can explain how and why the outcomes are not prejudicial.
- **Don't delay.** Data is the new currency, and creating intelligence from data at scale requires ML technology. As this technology continues to take hold, early adopters will enjoy a competitive advantage, as they are able to improve the customer experience while detecting more fraud.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

RELATED AITE GROUP RESEARCH

Digital Authentication: New Opportunities to Enhance the Customer Journey, September 2017.

Combating False Declines Through Customer Engagement, May 2017.

Machine Learning for Fraud Mitigation: The Substance Behind the Buzz, April 2017.

Moving Beyond the Password: Consumers' Views on Authentication, March 2017.

Global Consumer Survey: Consumer Trust and Security Perceptions, February 2017.

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com

ABOUT AITE GROUP

Aite Group is a global research and advisory firm delivering comprehensive, actionable advice on business, technology, and regulatory issues and their impact on the financial services industry. With expertise in banking, payments, insurance, wealth management, and the capital markets, we guide financial institutions, technology providers, and consulting firms worldwide. We partner with our clients, revealing their blind spots and delivering insights to make their businesses smarter and stronger. Visit us on the web and connect with us on Twitter and LinkedIn.

AUTHOR INFORMATION

Julie Conroy
+1.617.398.5045
jconroy@aitegroup.com

CONTACT

For more information on research and consulting services, please contact:

Aite Group Sales
+1.617.338.6050
sales@aitegroup.com

For all press and conference inquiries, please contact:

Aite Group PR
+1.617.398.5048
pr@aitegroup.com

For all other inquiries, please contact:

info@aitegroup.com

101 Arch Street, Suite 501, Boston, MA 02110 • Tel +1.617.338.6050 • Fax +1.617.338.6078 • info@aitegroup.com • www.aitegroup.com