Conquering big data challenges

Big data is here for financial services

Don't get left in a "cloud" of dust

Financial institutions have invested in Big Data for many years. Regulatory requirements for record retention often have been the main driver, while development and access for advanced analytics have been key challenges. New advances in technology infrastructure have opened the door for leveraging these data caches for better customer insights, models and strategy development in the past few years. While the technology exists, adoption is challenged by factors such as a soft economy, costs associated with infrastructure and lack of internal know-how.

Where we've been

Financial institutions are flooded with a massive amount of information, including:

- Customer information and their performance over time
- Transactional information and marketing campaigns
- Different customer views in various databases (e.g., CRM and application systems)
- Prospects, partnerships, market demographics and even competition
- Consumer and business credit data

One of the challenges with executing analytics on these data sets is the fact that they are spread across many different platforms. Historical customer data and transaction data may reside in different warehouses, treatments remain in the Customer Relationship Management (CRM) or application systems, and some data could even be stored on a single analyst's local desktop. This creates a problem because executing effective analytics on widespread data not only is inefficient, but it also opens the door for missed opportunities and compliance challenges. Not knowing where data exists or not having access to these key elements can change the outcome of the best model or strategy. The key to staying competitive is not only what your analysts know and apply to models, but also leveraging that information in an efficient and increasingly scalable manner to drive strategies across array and variety of data.

Navigating the business challenges with Big Data

The four strategic challenges most financial institutions face when defining their end-state Big Data analytical infrastructure are development, data, integration and regulatory compliance.

The challenges

Development

- Costs for system development can be daunting, ranging from $5 million to $100 million

- Lack of internal experts on new systems and technology

- Difficulty around accessing legacy systems and incorporating ongoing traditional processes

Data

- Internal data flows for movement of large files and across silos

- Data governance and data linkage

- Ongoing sourcing and integration of new data from third parties

Integration

- Integrating new systems with existing tools and processes
  Implementation across business units and divisions

Regulatory compliance

- Ensuring compliance with regulatory requirements regarding data retention for model building

- Adherence to depersonalization and security of sensitive information

- Planning for future regulations

Financial institutions are starting to look at Big Data as an investment rather than a technology expense. Proving return on investment (ROI) on Big Data infrastructure is a common concern, and businesses must evaluate multiple benefits across business lines, e.g., marketing strategy improvements, risk management, cross-sell strategies and operating efficiencies, as they try to overcome initial investments on internal system development. The key to an initial Big Data ROI analysis often is about what costs can be eliminated once a Big Data platform is in place.

The biggest advantage in moving to Big Data is giving your analysts the ability to do their work faster and with greater flexibility than before. This benefit translates into a real competitive advantage as the business gains greater insight and intelligence into customers, their activities and their interactions, thus leading to the development of relevant analytics and strategies.

The biggest advantage in moving to Big Data is giving your analysts the ability to do their work faster and with greater flexibility than before

The time to invest is now

Your business is likely investing in Big Data already. However, the transformation of data into actionable information is the next threshold to cross. With the entry cost into Big Data technology shrinking, now is the time to invest. This investment is about not only getting ahead, but also remaining competitive. In a recent Forbes article, the International Data Group (IDG) revealed:

- Seventy percent of enterprise organizations are either deploying or are planning to deploy Big Data–related projects or programs

- Nine percent of these businesses plan to spend from $10 million to $100 million on Big Data this year, with an average investment of $8 million

- Quality and speed of decisions are the main drivers for investment

While the investment dollars from IDG seem staggering, it's important to note that there are various options for entry. This includes outsourcing into cloud-based hosted environments that not only minimize costs, but also offset internal development timelines and resource constraints. Big Data doesn't have to be a budget-busting initiative.

The cornerstone of a successful big data infrastructure

Big Data has become an all too common phrase in the industry, arguably to the point of overuse. While Big Data can have different connotations for different business entities, in general it is the only fitting way to describe the availability of large data sets that may not have been accessible before. Gartner defines Big Data as "high-volume, high-velocity and high-variety (3Vs) information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making."1 While this definition may seem quite intimidating, the 3Vs are the building blocks for most Big Data technologies. For many businesses, the true definition of Big Data might be "It depends." The truth is that the phrase "Big Data" is relevant not only based on the size and scale of the business, but also on the data sources available to it. Big Data is not just about increased data and storage. It's also about finding opportunity in your existing data sources and scaling for the future.

The 3Vs of data management

The 3Vs of data management, originally introduced by Doug Laney of Gartner, speak to a three-dimensional view of data-management platforms. They (or some variation of them) have largely been adopted by the industry to define Big Data. Thirteen years later, the same 3Vs — Volume, Velocity and Variety — now are the cornerstone dimensions for any Big Data service.
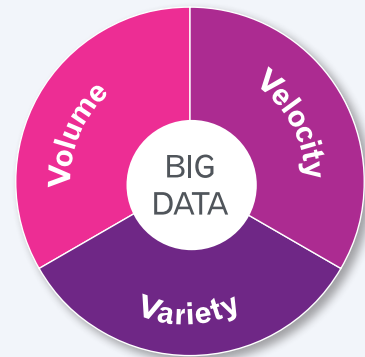
VOLUME

Volume of data is nothing new for financial institutions. For many years, banks have been receiving and processing millions of transactions, payments, applications and automated customer treatments on a daily basis. Over time, these volumes have increased as a result of gathering additional data points not captured traditionally and an increase in the frequency of transactions. Increased volume also can be attributed to the need to service customers better during automated phone and online interactions by retaining data, such as log events, as customers interact with the business. There is not only more data being captured and stored, but it also is broader, deeper and more complex.

VELOCITY

Velocity also is nothing new for financial institutions such as credit card issuers and processors. Along with daily volume, transactions need to be processed quickly. Most banks can make a decision on a credit application in seconds or determine if a transaction meets conditions for potential fraud in near real time. Velocity in Big Data analytics speaks to a business's ability to take the volume of data quickly — transactions, payments, applications and customer-interaction events — and position it for use by analysts so they can improve strategies and customer insight. Velocity also speaks to the speed at which the technology used to house Big Data can process these vast chunks of overflowing data points. While simple in concept, the velocity of data coming into a financial institution can often slow the process of transforming data into insight.

VARIETY

Enterprise data arrives in a variety of types and formats. While in most cases data gathered internally from databases — such as transactions, Customer Relationship Management (CRM) and application systems — is somewhat structured, making these data sets from across various systems available to analysts and business users in a consistent format is a hurdle that must be cleared. Inconsistent formats from varying tools across business silos further complicate the process of turning Big Data into actionable insight. If you factor in the influx of unstructured data — such as data from social media, customer or industry blogs, call centers and product reviews — the task of making the data usable for a wide audience becomes even more difficult. New file-system structures, such as Hadoop, are slowly taking market share away from traditional structured databases because they have proven an ability to handle data of varying formats, both structured and unstructured.

analysis to detect potential independent variables that can identify a swap set. Once "swap-ins" and "swap-outs" are defined, clients can test the new challenger strategy either by running a test campaign or by testing the outcome in an independent strategy simulation within our sandbox. The analysis is focused on the following questions:

Did I make the right decision? Are the customers that I thought were too risky still a risk?

- If not, when did they improve and how do I identify predictive behavior for improvement at the time of making a decision?

- Would I book them today?

- Did a competitor book them with a similar product in a 60-day window?

- Are the customers I booked as good as I thought?

- Are they profitable?

- Did they book competing products in the same window?

- If so, are they using my product or the competition's and why?

Wallet-share analysis

Wallet-share analysis is similar to market-share analysis but with a twist. Market share is geared toward how your portfolios compare with your competition's. Wallet share takes this analysis a step further and asks the question, "How does my product perform versus the competitor's product in my customer's wallet?" The goal is to highlight:

Position in wallet

- Are customers using your product as their first, second or third option?
- Are your customers paying you first?
- What payment stresses do your customers have from other lenders?
- Is there cross-sell potential within your portfolio?

Wallet share then can be taken a step further by leveraging peer grouping to view the same analysis for your top competitors, thus giving complete market-share insight.

Concept in motion —
what our clients are doing
with our hosted data platform

Since the launch of Experian Analytical Sandbox™ our clients are discovering that Big Data provides limitless insight into customer trends and provides them with the tools they need to keep ahead of competition. Beyond traditional segmentation, modeling and strategy development some additional examples of the ways clients are leveraging on-demand historical credit data sets include:

Decline analysis/reject Inference

Decline analysis is all about refining your discreet back-end decisioning criteria on all credit applications that come through your door. It is more easily defined as "Where can I do better?" The concept is applied through an overlay of a financial institution's application history and account performance onto a raw credit snapshot in the Experian Analytical Sandbox and the monthly credit data following the point of application through 12 or 18 months out. The analysis then goes a step further and leverages a Chi-Squared Automatic Interaction Detection (CHAID)

Experian's hosted analytical sandbox solves these challenges by providing a fully hosted platform with a secure Web-based interface

How to get the most out of Big Data hosting

Financial institutions have leveraged their in-house historical data for many years to conduct look-back analysis, and it was the most common technique used to predict future performance. While this technique works marginally well, it presents multiple challenges associated with a sample that is skewed to an institution's customer base and performance within the portfolio. The missing link always has been how the same customers perform elsewhere and what other customers who are not part of an entity's portfolio look like. The only method of gaining a full market perspective was to purchase a large amount of depersonalized historical credit data sets and hope no essential data elements were missed in the requirements phase

Until now, the challenge to acquire vast amounts of historical market sample data was cost-prohibitive and potentially unmanageable, especially at the raw data level due to size constraints. Experian's hosted analytical sandbox solves these challenges by providing a fully hosted platform with a secure Web-based interface that not only gives clients access to vast amounts of on-demand historical credit data both at the raw and aggregated levels, but also provides the analytical tools to allow our clients to join their in-house performance data with the historical snapshots.

"The power of data is good. Couple it with analytics and it becomes great."

— Craig Boundy, CEO, Experian North America

Development
Quicker implementation

- No development costs for clients
- Reduced overhead through hosted access
- Immediate access, to allow you to build stronger models and strategies instantaneously

Data
Integration to key data sources

- No data lag time for backfill of relevant size
- Unbiased samples for stronger models and strategies

Integration
Create departmental synergies

- Consistent nonconflicting insight from all groups/departments
- Ability to create consumer links through a full consumer relationship picture

Regulatory Compliance
Data compliance issues are solved

- Satisfies Fair Credit Reporting Act rules
- No need to devote overhead to manage sampling and depersonalization

The Experian difference

Experian® is widely recognized as a global leader in consumer credit data and analytics. Armed with the right tools and the right resources, Experian is well-positioned to assist your business in conquering Big Data challenges by offering best-in-class:

Custom consulting — Establish your Big Data business plan from mission to completion. Our global consulting team can help with:

- System requirements and design
- Bureau data feeds (defining data elements, frequency and sourcing)
- Data linkage, depersonalization and compliance
- Industry-use cases
- Advanced analytical resources via:

Experian DataLabs — Our data scientist experts in Big Data systems and analytics are versed in handling seemingly impossible data sets and transforming information into actionable knowledge.

Experian Decision Analytics — Our financial-industry experts can build, enhance and expand your current strategies, attributes and models. From acquisition strategies to Comprehensive Capital Analysis and Review compliance, we can help you navigate your way with ease.

Experian Analytical Sandbox — Our remotely hosted Big Data platform gives your business the power to integrate credit data into your analytical framework without systems development, resource acquisition or retraining. Using familiar analytical tools in a compliant user environment, Experian Analytical Sandbox empowers analysts to:

- Link in-house applications data with credit information in a depersonalized manne

- Research and analyze market trends

- Validate and expand acquisition, risk and marketing strategies

- Conduct-wallet share and benchmarking analysis

Experian
475 Anton Blvd.
Costa Mesa, CA 92626
T: 1 888 414 1120
www.experian.com