Putting big data to work in your bank

An article by Mohammad Nasar and Christopher J. Sifter, PMP
Despite its widespread use in the banking industry and elsewhere, the term “big data” sometimes is difficult to define in a consistent manner. Many financial services organizations are unclear about whether the analytic techniques associated with big data really are applicable to their situations. Understanding the advantages and challenges of big data in banks is an essential first step in developing an appropriate and effective analytics platform.

Big data – managing information more effectively

Although the term “big data” is of relatively recent origin, the concept goes back many decades to the beginning of computerized data storage in the 1960s. As database models and data warehouses developed, bank technology evolved to accommodate the growing volumes of data that had to be handled.

According to one widely cited study, U.S. banks and capital markets were holding more than 1 exabyte in 2009 – that is, more than 1 billion gigabytes – of stored data.¹ If global data volumes are doubling every two years, as some analysts believe,² the banking industry could be handling data volumes in the tens of exabytes well before 2020. But the challenges of big data involve more than just volume alone, as the exhibit illustrates.
Big data challenges: Volume, velocity, and variety

In addition to handling massive volumes of data, banks today also must analyze and react to information faster – often in real time or at nearly real-time speeds. Moreover, many of the sources of banking data are tangential to the banking system itself. In addition to data that exists within their core business systems, banks also must accommodate unstructured data from third-party sources such as social media and the internet.

This combination of variables – volume, velocity, and variety – greatly complicates the challenges associated with data storage, access, management, and analytics. As a result, even small to midsize institutions find they can benefit from developing and implementing big data tools that can quickly turn masses of unstructured data into meaningful information.
Understanding big data tools

The underlying challenge of big data is a reflection of processing capacity. Current technology is limited in terms of the amount of data that can be stored efficiently and the speed with which that data can be accessed.

Big data technology is designed to overcome these limitations by using various methods to organize how data is stored and accessed. Two of the most widely used methods include:

**Breaking data sets into smaller chunks.** Using this approach, large data sets are oriented logically and restructured into smaller, separate nodes. This greatly reduces the volume of data required for processing by individual queries and calculations that must be performed when accessing and analyzing data. The benefits of “chunking” include significant time savings and reductions in the network capacity required.

**Structuring data into a columnar or hierarchical format.** This approach organizes and aggregates data according to common attributes. For example, individual banking transactions in a database could be stored according to the physical branches where the transactions occurred, types of transactions, the individuals handling the transactions, or other similar attributes. This approach is beneficial especially when looking for trends or patterns within the data.

In most cases, big data technology involves a combination of chunking and columnar structures, as well as various other technical approaches designed to structure data more efficiently, reduce throughput, and make the most effective use of existing network and processing capacities.
Big data success stories

How can big data technology be applied to add value in financial services organizations? Applications that use social media to address customer interactions and customer service issues are probably the most commonly cited examples, but some other potential uses can be of even greater practical value. Some recent success stories include:

**Customer analytics.** By integrating customer data from a variety of sources (both internal and external), a large regional bank recently developed a large-scale analytics platform that enabled it to improve cross sales and track customer onboarding based on certain key demographic indicators. Beyond this initial objective, the bank also was able to develop a more structured workflow and benefit analysis. This allowed it to increase the use of low-cost digital and mobile banking platforms, find other ways to improve its interactions with customers, and even do a better job of tracking the effectiveness of employee sales incentives.

**Stress testing.** As banks grow and must comply with new stress-testing and reporting requirements, the use of statistical modeling processes and other analytical tools normally associated with big data technology can be helpful. In one recent instance, a $15 billion bank used big data approaches to access and organize large amounts of data from a variety of sources. These data sets included large-scale economic and market data, as well as both structured and unstructured customer portfolio information. By using big data tools, the bank was able to organize data across clusters of computers, eventually becoming self-sufficient in complying with *Dodd-Frank Wall Street Reform and Consumer Protection Act* stress-testing (DFAST) requirements, without having to add full-time, in-house specialists.

**Risk management.** At first glance, the data volumes involved in monitoring customer transactions for the *Bank Secrecy Act of 1970* or anti-money laundering (BSA/AML) violations might not seem large enough to justify the use of big data tools. In fact, the information compiled during the preparation of a BSA/AML suspicious activity report (SAR) is far-reaching and highly valuable, but it remains largely unstructured within the bank’s information systems. This characteristic can be overcome by applying big data analytic techniques to identify trends across various relationships and individual transactions. In one recent case, the application of data mining techniques helped a bank identify additional patterns that were not immediately obvious, which led to the discovery of criminal activity.
Making big data work effectively

The implementation of big data technology requires significant technical expertise. Beyond the technical components, however, the successful development and application of a big data analytic platform also depends on well-planned and well-executed strategic and tactical decisions, based on three foundational elements:

1. Effective data governance
The primary purpose of big data analytics is to help answer critical business questions. This capability depends on access to data that is consistent, accurate, and trustworthy at every step of a business process. Effective data governance can help give users confidence that the data not only is accessible, but is accurate and complete.

A commitment to data quality and strong data governance must be ingrained in an institution, with buy-in from users at every level of the organization. In addition, security and accountability become even greater concerns whenever an institution is accessing, analyzing, and retaining massive volumes of data.

2. Careful, selective implementation
Big data does not mean traditional data warehouses are eliminated; they still have value and play an important role. But traditional data warehouses typically are designed for a very specific reporting or business purpose, such as credit risk analysis.

Big data analytics, on the other hand, typically go beyond a single functional need and, as such, involve access to data further upstream to provide additional insights and explore relationships that are not immediately apparent. This generally requires the use of distributed storage tools that can access data across multiple sources, along with other tools such as data federation/virtualization, and data visualization software.

It is important to apply the right techniques and tools to fit each situation and to distinguish genuine big data opportunities from those that could be more accurately described as “pretty big data” situations. In many instances, business needs involving large data sets still can be met through the application of traditional data reporting technology.

3. Insightful, analytic people in a culture that values them
Big data by itself does not answer business needs. Having access to data is an essential starting point, but the true value of big data is in the analytics it enables. This requires people who are able to use data to answer questions, make better decisions, discern previously unidentified relationships, and recognize new opportunities.

The most successful big data analytic functions typically involve people who have subject matter expertise in the specific business needs being explored, along with an inquisitive nature to dig into the data and ask new questions. Some IT capabilities are needed, but often there is considerable value in involving front-office personnel in these data analytic functions.
Ongoing commitment can add value

Developing and implementing big data techniques initially can appear to be an overwhelming challenge – often compared to “boiling the ocean” due to the immense amounts of data involved from both internal and external sources. It is a large undertaking that not only takes time, but also requires ongoing commitment and follow-up as a bank encounters new uses for data, new sources of data, and new challenges that must be addressed to respond to business needs.

When carefully planned and executed, however, the application of big data tools can add significant value to financial services organizations by empowering them with new analytic capabilities that reveal previously unrecognized opportunities and enable faster, more responsive approaches to business needs.
Contact information

Mo Nasar is with Crowe Horwath LLP and can be reached at +1 630 574 1846 or mohammad.nasar@crowehorwath.com.

Chris Sifter is a director with Crowe and can be reached at +1 312 857 7363 or chris.sifter@crowehorwath.com.
