

EXCERPT

Westpac's Journey into Big Data: From Transactional Data to Big Data Analytics

Daniel-Zoe Jimenez Sharon Tan
Michael Araneta

IN THIS EXCERPT

The content for this excerpt was taken directly from Buyer Conversations: Westpac's Journey into Big Data — From Transactional Data to Big Data Analytics, IDC Doc#AP2670118T, September 2011

IDC OPINION

Forward-thinking banks clearly understand that financial services organizations should not only sell products and services but should also become true customer-centric organizations in the way they manage customer data and serve their clients across different channels (i.e., providing a seamless multichannel experience). With the advent of Big Data, and in particular in the context of social media networks, it is becoming increasingly clear that channel strategies should not be only restricted to traditional banking channels but also incorporate new customer touch points (i.e., social media sites). While accomplishing this is not an easy task, organizations should start the process by designing an information management strategy, which will help them navigate through the Big Data journey.

Westpac New Zealand Limited (Westpac) is among the banks that have recognized the need to gain deeper visibility into its data assets, and Informatica's solution was, according to Torrance Mayberry — Senior System Manager for business intelligence (BI), data warehouse (DW), and advanced analytics — critical as a way to help the bank break down silos and develop a climate of innovation and knowledge sharing across the organization. While customer centricity is still in the horizon, Westpac is currently exploring the opportunities presented by Big Data, especially in the context of social media, which has further intensified the bank's need to gain deeper customer understanding. In that respect, while Westpac is well aware that social media data is just one of the many data sources available today, the bank is currently focused on sentiment analysis as part of its journey into Big Data analytics.

SITUATION OVERVIEW

Organization Overview

Westpac is one of the country's largest full-service banks, serving more than 1.2 million customers. Westpac provides a full range of banking and financial services for personal, business, and agribusiness customers.

Westpac operates a network of 200 branches and 500 ATMs nationwide, and has 900,000 registered online banking customers, and employs 5,000 staff.

Challenges and Solutions

The Beginning: Managing Transactional Data

Westpac's prelude to its Big Data journey began with the establishment of a DW in 2001. The establishment of the DW was kick-started by the marketing department's initiative to implement customer analytics for a component of the Customer Management Programme. In 2007, Westpac realized that it did not, at that time, have a complete, single-view of its customers and products. Around the same time, data governance was also gaining in priority with Westpac's top management because of compliance, risk mitigation, and how it could help improve operational efficiencies. Similar to other large organizations with data-intensive operations, Westpac had a number of disconnected applications running that were not designed to share data with each other. There was also no data strategy at that point in time. All these led to the following challenges:

- Inability to obtain a complete and single version of Westpac's customers and products as :
 - Critical information was collected and stored in information silos.
 - Disparate data definitions and information usage was the norm.
- Response time for report generation or ad hoc requests was often slow, as data generated from different applications had to be manually cleansed, reconciled, and hand coded. In addition, there were inefficiencies, due to duplications of effort and data.
- Inconsistent techniques were used for data integration (e.g., hand code, push of data into flat files). Moreover, these are not future-proof as there was no concept or techniques used for data services, service-oriented-architecture (SOA), or in order to address new Big Data opportunities.
- Data treatment and data security classifications were incorrectly done, which resulted in unnecessary risks and potential complications.

The establishment of a DW was the first step toward Westpac's key goal of having a single version of the truth to make better use of its data assets, in order to support fact based and consistent decision making across the many business units. During the period November 2001 through February 2002, Westpac also evaluated a number of extract, transform, and load (ETL) tools to determine a standard and preferred supplier for Westpac Banking Group.

The vendors evaluated for ETL were:

- Ascential – Datastage (now part of IBM)
- CA — Decisionbase

- Datajunction (now part of Pervasive Software)
- Hummingbird — Genio RTI (now part of Opentext)
- Informatica

Eventually, Westpac made the decision to standardize on the Informatica platform. In 2007, Westpac recognized that an enabler of its data strategy required it to evolve from traditional ETL to enterprise data integration. The Informatica platform ensured Westpac could take a holistic approach toward data integration and it addressed a range of data delivery opportunities. The following were Westpac's reasons for its choice:

- Simplicity.** Informatica's pricing approach was easy to understand. This was a factor in the decision-making process. For instance, it was easy to correlate its value back to the data strategy. This factor was important because, according to the bank, it helped promote collaborative discussions between IT and the key business stakeholders about the potential value to be achieved from the platform.
- Integrated solution with reusable data quality rules.** Informatica's product components appeared to be well integrated, with common, reusable data quality rules across its data integration, data quality, and MDM products.
- Good infrastructure and support.** Informatica's infrastructure was evaluated by Westpac and was considered to be the most suitable with open repository and server engine operation. In addition, Informatica 9, which was launched in 4Q09, included enhancements to enable data services based on SOA. Informatica also provided reliable technical assistance, which was a crucial area of consideration since Westpac intended to standardize and grow its data integration requirements with one vendor.
- Availability of audit trails.** Informatica's Metadata Manager provides audit trails for all user-definable metadata, as part of the advanced edition of PowerCenter and PowerExchange, which includes Metadata Manager and provides the availability to view the history of changes to business glossary categories, business terms, and custom objects.
- Vendor independence.** According to the bank, Informatica's independent third-party position was one of their key strengths, and it has ensured that Westpac keeps pace with emerging technology.

Data Warehouse Solution Implemented

- IBM Banking DW (BDW).** A customized, hybrid version of the BDW model was implemented in an Oracle RDBMS. Industry best practices were incorporated into its DW by customizing the BDW that Westpac has purchased from IBM. IBM's BDW is positioned as a fully normalized and complete representation of the financial services industries' data needs.

- ☒ **Informatica.** The Informatica platform for data integration and data access was selected for the reasons outlined earlier. Informatica products that were implemented consist of PowerCenter and PowerExchange, including Informatica Metadata Manager. Westpac is currently running version 8.6.1 of the Informatica platform and is in the process of upgrading the platform to version 9.1.
- ☒ **Hand coding** through a combination of PLSQL, COBOL, and BASE SAS.
- ☒ **Hardware used.** IBM p570 AIX 5.3 dual-core 64-bit processors 2.2GHz (4 CPU).

Currently, Westpac's DW size is approximately 6TB; IDC's *Asia/Pacific Software Survey 2011* indicates that 25.18% of financial institutions in this region have DWs that are below 10TB size. While Westpac's DW is currently below the average size of DWs of financial institutions surveyed, the bank expects its DW size to grow by 25% in a year's time. It is also worthwhile to note that certain regulatory requirements also contribute to Westpac's DW growth expectations. Although there are very close ties between Westpac New Zealand and Westpac Australia (which is typical of most financial institutions in the ANZ region), the Reserve Bank of New Zealand requires that banks in this country maintain its sovereignty by physically locating its data in New Zealand. This is to ensure that business continuity in this country would not be impacted should a catastrophe impact Australia. Finally, Westpac's DW has also expanded, over time, in terms of data source/data types stored — and the bank expects more data sources to be gradually added into the DW, which will again, contribute to data volume growth going forward.

The primary's DW architectural approach is based on a centralized EDW as the core hub, which services data to a few dependent data marts. These data marts were established as part of delivery of analytical applications to enable and optimize pricing, marketing, and economic profit. For instance, the history of the financial products' pricing is stored and managed in the DW and feeds into the data mart to meet the pricing optimization solutions analysis requirements. Other similar usage includes the bank's Proactive Risk Management (PRM) system, which Informatica's platform sources data from, allowing the bank's risk practice to rapidly refresh knowledge on fraud and reduce single person risk. Usage of the trusted information generated from the DW for a variety of decision-making purposes has clearly expanded, with business linkages being formed with a growing number of business units within the Westpac organization.

Results

Lessons Learned (So Far)

- ☒ **Keep in mind that early wins are critical.** Westpac's DW team (the NZ DW team) has scored a significant internal success by establishing a complete and accurate view of New Zealand MDS/MIDAS products and transactions within its operational data store (ODS) in record time and budget. This team was then able to use this victory to engage key stakeholders and the line of businesses (LOBs), thereby creating awareness and support for its data strategy.

According to Westpac, Informatica's data integration solutions were instrumental in the success of this project, and this has led to Informatica being adopted as the data integration platform for Westpac New Zealand.

☒ **Provide measurable goals and results to win top management support.**

Westpac quite rightly sees that quantifying data-centric project's goals and results in productivity and profit terms increases the probability of securing both board approval and funding for such projects — especially in the wake of the global financial crisis, where according to Mayberry, "...a 4–6% productivity/margin increase would make an enormous difference to a board."

☒ **Win key business stakeholders' buy-in by getting their skin in the game.**

Several LOBs are currently able to utilize reports or information from Westpac's DW. These include the finance departments in both Australia and New Zealand, the marketing insights team, risk analytics, insurance, business and retail banking units, and customer and technology services divisions. According to Westpac, its key business stakeholders have now recognized the strategic importance of the banks' data assets and have supported the acceleration of Westpac's DW in recent years.

☒ **Focus on improving IT and business collaboration.**

While IT knows what the data is and how to manage it, the business users are the ones that know how to accurately interpret the data in real-life situation contexts. Good collaboration between IT and business users has been proven crucial in minimizing cases of "lost in translation." Going forward, Westpac is also keen to further enhance this collaboration through using the collaborative features that would be available in its forthcoming Informatica platform upgrade.

☒ **Adopt a holistic data strategy.** Westpac has invested in fleshing out a comprehensive data strategy, with the vision to "Inform business decision making through delivery of accurate, relevant, comprehensive, and timely managed information." Westpac's DW was built around four themes:

- ☐ **Information completeness:** Enabling analysis and insight drawn from single source of product, channel, and customer information.
- ☐ **Business performance management:** Enabling measurement, reporting, planning, incentive setting, and area scorecards.
- ☐ **Customer intimacy:** Enabling new contact strategies, targeted marketing, customer segmentation, and analysis of customer experience.
- ☐ **Sustainable infrastructure:** Enabling wider use, greater frequency, quicker response time, reduced outages, and greater staff productivity.

Westpac's Journey into Social Analytics

In a time where the explosion of data has become apparent to every organization and across all industries, real-time analytics is becoming a necessity. Today, there is a deluge of information being generated by social media (e.g., blogs, wikis, social networks), tracking systems (e.g., mobile devices, clickstreams), and transactional data (e.g., enterprise resource planning [ERP], CRM), and this is rising by the second. In fact, IDC has recently estimated, as part of the annual *Digital Universe* study, that in 2011 the amount of information created and replicated will surpass 1.8 zettabytes (1.8 trillion gigabytes) — growing by a factor of 9 in just five years. With that in mind, and given that in the "information society" information is money, organizations are realizing that this data can be tapped, analyzed, and utilized in order to gain deeper customer insights and improve decision making. However, coping with the aggressive growth of internal and external unstructured and semi-structured data sources, mainly boosted by the widespread use of collaborative technologies, is introducing a number of challenges for organizations. In addition, not all data generated by users is relevant, and organizations are becoming more aware of the importance of understanding what data they should analyze in order to gain true customer insights from external and internal Big Data sources, to support effective decision making.

While for hyper-connected enterprises, the definition of internal and external data is getting blurred, some banks prefer to start off their Big Data journey by looking at their internal-unstructured data sources (e.g., call center data logs, relationship-management data, etc.) before they venture into analyzing external-unstructured Big Data. On the other hand, often, Big Data analytics use cases involve querying a large variety of data from multiple sources at the same time. This data would normally be a combination of data already available internally (within the enterprise) and data obtained externally (usually from social media sites).

Westpac's first venture into analyzing semi-structured and unstructured data began with some of the external data sources available today: social media data from the major social media networks. The bank has started its social media activities like most banks in the region: it launched a Facebook page, created a twitter account, as well as its LinkedIn profile. However, it soon realized that social media was not only about presence but also about engaging with customers. In the past, Westpac had used Facebook to run and promote a number of marketing campaigns, but it did not set up its own Facebook page. Clearly, at this stage, the bank was only using social media as a unidirectional marketing channel — in a similar way to how traditional marketing channels were normally used. However, Westpac recognized that social media presented great opportunities for the organization, since millions of conversations are constantly taking place, and some of those were about their bank.

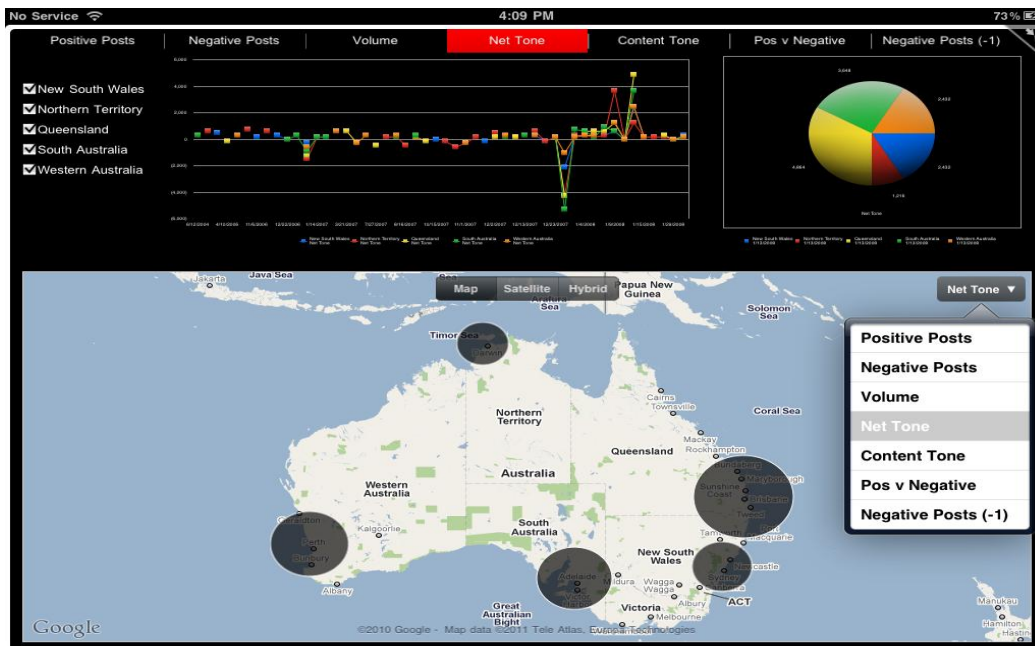
At the unfortunate time of the Christchurch earthquake, Westpac realized that social media was a good way to measure customers and prospects' sentiment and response to the actions undertaken by the financial institution. In parallel, the bank realized that customer data gathered from social media was at risk of creating another silo of information that could have had to be manually reconciled back to on-premises for data analysis.

As a result, in order to stay true to the corporate data strategy and inhibit that from happening, the bank reutilized the existing assets of Informatica, which had previously been used by the bank.

At this point, according to Mayberry, the assets of Informatica were critical, since they help the bank with its new social media data governance. In addition, the bank evaluated a number of tools that could be used in order to analyze unstructured and semi-structured data. Expert Systems (an Italian text analytics software vendor) was finally selected for its existing banking references and its rich product functionalities that help grasp the "meaning" of content — thanks to search methodologies based on deep linguistic analysis (semantic engines). Additionally, Google Maps is being used by Westpac in order to deliver geospatial data, and Microstrategy is being utilized to provide mobile intelligence on iPads, automating analysis of sentiment, emotion, mood, and opinions (see Figure 2). In addition, Westpac is experimenting with an open source distribution of Hadoop. According to the bank, the main challenge lies in picturing the best way to unleash the most from Big Data, so that it tells a story. In addition, being creative is important, since telling a compelling story may involve making connections, as well as going back to other data sources to verify them, which also requires keeping a keen eye on the big picture. Similarly, Westpac recognizes that the main benefit of using Hadoop is its ability to enable agile data analysis. Its near real-time data has enabled faster computations to test a number of data sets, assumptions, and algorithms.

FIGURE 2

Westpac: Social Analytics Pilot Project — iPad View



Source: Westpac, IDC, 2011

The whole project has been challenging, from the process of applying complex machine learning techniques as part of the data integration process, to dealing with the large collections of data. Today, it is clear that the problem is not finding data, but understanding what data is relevant to the organization and figuring out what to do with it. In addition, data needs to be converted into a tractable form, so it can be presented to the business executives not as a collection of data sets, but in the way that tells the story, so actionable insights can be derived from it for faster decision making.

Mayberry has put together a road map with a set of milestones and metrics so the ongoing progress of the project could be easily evaluated. While initially it was not an easy task, since social analysis is still relatively new and not everyone in the organization was aware of its benefits, most of the success was attributed to the increased collaboration between the IT department and the business executives. In fact, it seems that the social media plan has gathered a good level of attention across the whole organization, creating renewed interest around all data-related issues. The process started by taking a macro perspective of the entire social Web to obtain a succinct understanding of customers and prospects. Westpac has recognized that customer persona is much more than just data from social networking majors (e.g., Facebook, Twitter, Google+, Foursquare, YouTube, etc.). This gives the bank an opportunity to reach different types of customers, adapt internal messaging and behaviors to match with customer interest at a micro level to deliver through each channel a unique experience. In addition, the approach ensures the bank can learn proactively the best place to interact with customers or prospects, should a new social start-up arise.

While the pilot was initially restricted to New Zealand, as the bank progressed in its social media project, it was quickly recognized that by harnessing social data for the entire brand it would position it better to uncover future threats and opportunities. The fact that an organization can start listening to customer conversations in real time, and this can also be presented with dashboards and graphs across Australia and New Zealand (ANZ), is something that clearly generates attention. As a result, given that Westpac is a global brand with activities not restricted to a single country, the project has also been extended to cover Australia. By doing this, business executives can now gain real-time insights into the bank's brand health and its product performance across different geographies, correlating specific branch performance to customer social data.

Where the Journey Ends

An organization's journey into Big Data never ends; on the contrary, in a world where the only thing that is certain is change, the need to improve decision making by gaining deeper understanding of larger collections of datasets in real time will increase over time. Mayberry stated that the bank is still at early stages, and that they are now moving from the experimentation phase to working through the future opportunities that can be gained from Big Data analysis. Westpac expects that this journey will help the bank learn from what customers are saying, what are they looking for, or the challenges they are having. In the process, this will enable the bank to build, sell, and market better products, programs, and services.

Westpac is well aware that its social media journey has just started, and that this should be put in the context of its enterprisewide Big Data journey. Along the way, it is clear that Westpac will find numerous challenges but also big opportunities. For example, today, the bank is experiencing an increased level of collaboration between the different groups and countries, as well as between IT and business executives, whom are working together to formulate a blueprint, so the project can be propelled to the next stage.

While the final road map is currently being formulated, it has become increasingly clear for business executives and IT alike that successful results will only be achieved through effective collaboration between the two groups. In addition, Westpac has determined that the next waves of experimentation will focus on expanding the initial prototype to the areas of sentiment trading and MDM. This will be done by exploiting the capabilities of the newly released Informatica 9.1.

On the other hand, the main goal will be to use social media as a sales tool. Like other organizations in the region, it is clear that Westpac is looking to monetize its social media experiments by moving from basic sentiment analysis to customer engagement. In these changing times, business success does not only depend on reacting to competitor moves or customer requirements but also on anticipating trends and customer behavior. In that respect, we believe that social media analysis should transcend basic sentiment analysis and move to actively engaging with customers. While most organizations do this by initiating a dialogue with customers through social media networks and actively pushing products in a reactive way — as a response to their needs — IDC believes that the most mature organizations will start leveraging predictive analytics as part of the social business strategies. This will enable organizations to proactively engage with customers, anticipating their needs, based on the rich mines of data available through social networks.

Copyright Notice

This IDC research document was published as part of an IDC continuous intelligence service, providing written research, analyst interactions, telebriefings, and conferences. Visit www.idc.com to learn more about IDC subscription and consulting services. To view a list of IDC offices worldwide, visit www.idc.com/offices. Please contact the IDC Hotline at 800.343.4952, ext. 7988 (or +1.508.988.7988) or sales@idc.com for information on applying the price of this document toward the purchase of an IDC service or for information on additional copies or Web rights.

Copyright 2011 IDC. Reproduction is forbidden unless authorized. All rights reserved.