CUSTOMER NEEDS AND STRATEGIES

Bridging the Structured/Unstructured Data Gap at BlueCross BlueShield of Tennessee

Henry D. Morris

IDC OPINION

Organizations need to utilize all available information to identify the best way to manage relationships with customers and partners. It is challenging to develop a single view of customer or partner information because of the diversity of sources of structured data. Overcoming the divide between the structured and unstructured data sources to build a single view poses additional issues. Yet, the experience at BlueCross BlueShield of Tennessee shows that it is possible to build bridges across these two information sources with a variety of strategies and techniques. The organization's experience points to the following as critical success factors:

- Clarifying lines of responsibility to bridge the organizational disconnects that impede the formation of a single view
- Defining and documenting a vision for an integrated information architecture and revising it on an ongoing basis to keep it current with evolving technology
- Prioritizing projects that focus on business operations with a significant financial impact to the firm while advancing the implementation of the architecture

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

Filing Information: December 2006, IDC #204813, Volume: 1, Tab: Users

IN THIS STUDY

This study examines the provider reporting project at BlueCross BlueShield of Tennessee, which can be seen as a best practice for bridging the gap between structured and unstructured data. The work was guided by a carefully developed information architecture and used a mixture of well-established and emerging techniques to achieve a productive result.

SITUATION OVERVIEW

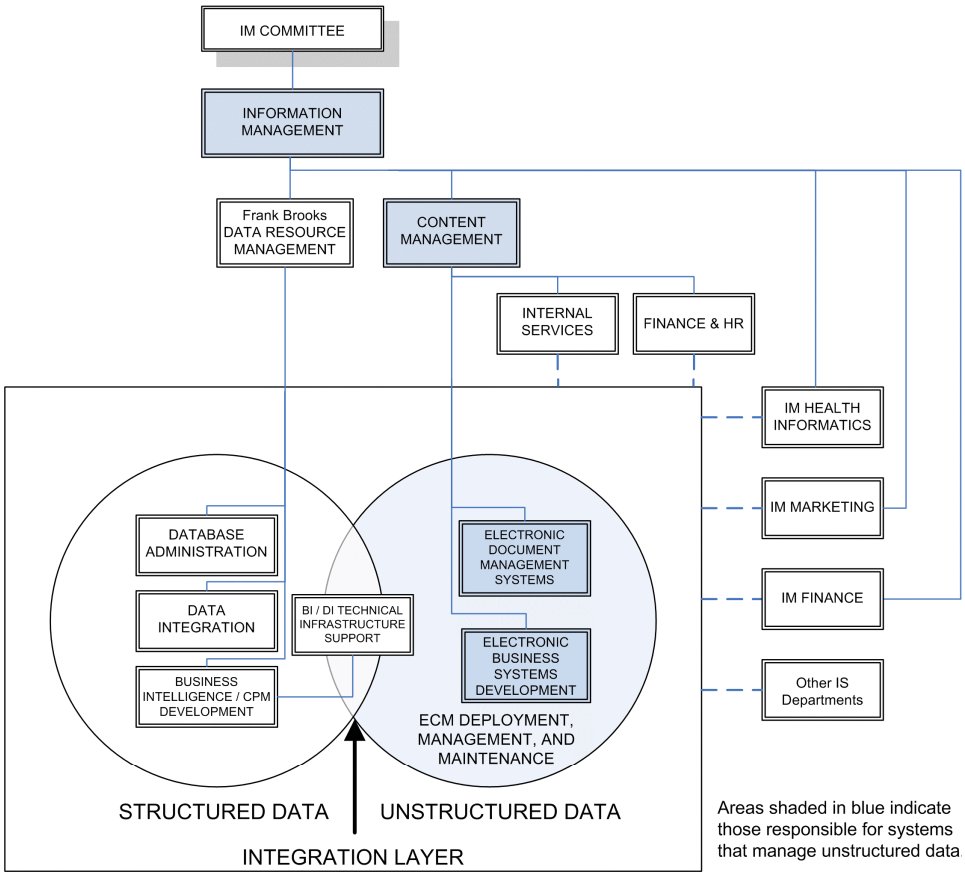
Organization Overview

BlueCross BlueShield of Tennessee (BCBST) is the 11th-largest Blue plan in the United States and the largest health benefits company in the state of Tennessee. BCBST has nearly 3 million members who are served by a large network of thousands of healthcare providers. Ensuring accurate and up-to-date information on members and providers is a primary goal of the IT function at BCBST. Figure 1 shows how the company is organized in support of information management. There are clear responsibilities established for structured data, for unstructured data, and for building an integration layer between the two.

Frank Brooks, interviewed for this study, is the senior manager of data resource management (DRM) and the chief data architect. Figure 1 shows organizational responsibilities at BCBST for enterprise content management (the management of the unstructured data) as shaded areas on the chart. Frank's DRM group is responsible for all of the unshaded areas in the structured data circle in the diagram. But his group is also responsible for the building of an integration layer that reaches out to the enterprise content management (ECM) area. Frank wrote the Data Resource Management Charter in 1986. In 2000, his area began authoring the IM Master Plan, a document that is updated on an annual basis to reflect the current operational, tactical, and strategic objectives. The most recent version covers topics such as business process management, performance management, geospatial information, and master data management.

FIGURE 1

Information Management Organization at BCBST



Source: BCBST, 2006

Challenges and Solution

BCBST's relationships with its providers and members (i.e., customers) are at the heart of its business. BCBST, like some of the Blue plans, is chartered as a not-for-profit organization. Cost management is critical to its financial viability. The company must ensure that the subscription revenue from its members suffices to meet its obligations to pay providers for healthcare services rendered on behalf of the members. Annual rate negotiations with providers is a business-critical activity because it sets the cost structure upon which BCBST must operate throughout the year.

The provider reporting project set out to improve the level of information on providers to be used by BCBST staff in rate negotiations with healthcare providers. It would extend the provider data integrated in the existing enterprise data warehouse, an asset that had grown in value to the company over the past 10 years. The data sources about providers include paid claims data and call center records. The goals

set out were to improve the delivery of information reported from the warehouse and to enhance the available data on providers that could be accessed in building the reports. In a proof-of-concept project, BCBST employed two complementary strategies to bridge the worlds of structured and unstructured data:

- **Turning data into content.** This strategy is a content-centric approach in which the structured data itself or related metadata and reports is incorporated into a search index, enabling access via a search interface. This increases the speed and ease of accessing and delivering reports to the broad population that uses search.

- **Turning content into data.** This strategy is a data-centric approach in which elements or concepts are extracted from the content (using text analytics or other technologies) and then transformed and loaded into the fields of a database record, enabling access via standard SQL queries or SQL-based tools. This expands the data set available for reporting because the supplemental data can be joined with existing data and then accessed via SQL.

These two bridging strategies are described in *Unified Access to Content and Data: Delivering a 360-Degree View of the Enterprise* (IDC #34836, February 2006). The two strategies are mutually complementary in the following ways:

- Enhancing the data available for reporting (by turning content into data) makes the reports more valuable.

- Making these reports and related metadata accessible via search (by turning data into content) increases ease of access for this valuable information.

The following sections consider these strategies in turn.

Turning Data into Content

Account managers select and view reports or create new reports based on current and historical data about their accounts. In addition, over 80 clients (external to BCBST) receive two cubes each with four years of summarized data on their activities. Providing customized information unique to each client is a competitive differentiator for BCBST in its efforts to acquire and retain accounts.

In support of this implementation, BCBST used Cognos business intelligence (BI) technology. BCBST had established Cognos as its standard for reporting and analysis when the enterprise data warehouse project was initiated in 1996. The organization adopted Cognos PowerPlay Web and Impromptu Web Reports in 2003 to create and access reports over the Internet. Then, in 2006, BCBST began migrating to the Cognos 8 BI suite.

In its proof-of-concept project, BCBST implemented Cognos Go! Search to improve access to reports. The report delivery system enables search for relevant reports by filling in a search box. Cognos Go! Search evaluates the request and searches the content of the reports and the related metadata, returning a ranked list of results that meets the search criteria and for which the requester has the security to access. Clicking on any of the results launches a report.

This search interface is presented to users in a portal environment. They can then run the Cognos application within a dedicated portlet. The content managed by Cognos also can become part of a broader search index managed by IBM's OmniFind. This enables the search within the portal to range over any of a number of reports and several external provider-related Web sites. It will be expanded in the future to include unstructured content from a number of content management repositories. This capability leverages Cognos' support for IBM's Unstructured Information Management Architecture (UIMA) framework and is intended to provide comprehensive access to any available content.

This strategy of turning data into content to become accessible via search is an incremental extension of BCBST's content management strategy. The structured data, which has previously been accessible via SQL or a SQL-based toolset (e.g. Cognos), is turned into content. It is then accessible via a search interface like any other type of content.

Though this is a powerful result, the strategy taken by itself has inherent limitations. The reports can access the integrated structured data in the data warehouse but cannot access information from the vast stores of unstructured data. Leveraging the unstructured data asset for structured data analysis was the next step in the work of the data resource management group at BCBST. It would bring them closer to their stated goal of achieving a single view of healthcare providers.

Turning Content into Data: A Proof-of-Concept Pilot

The next phase in provider reporting investigated the practicality of text analytics for extracting facts from unstructured text and transforming this information into structured data. The end goal was to enhance the scope of structured data on providers that is available for reporting and analysis. Roughly 65% of the call center volume handled by BCBST comes from its network of providers. Significant portions of the information is contained in call center notes, stored in comment fields within call center records. These comments were identified as a fruitful source of enriched information on providers.

BCBST turned to IBM for help on this project to leverage its expertise and partner network established through the UIMA initiative. IBM brought in ClearForest, a partner that supports the UIMA framework, to assist on the project. ClearForest's Issues Analyzer has been used in the automotive industry to analyze and annotate comments from repair records to identify product issues. However, the ClearForest technology includes a vocabulary of issues (e.g., financial issues of nonpayment or late payment) that can be applied across industries.

Figure 2 shows the architecture and flow used in the proof-of-concept pilot.

IBM's OmniFind was utilized as a search engine to crawl, parse, and index the text sources (as noted in the blue boxes labeled "Search Portal"). The information is then passed to ClearForest for text analytics to identify, annotate, and extract facts from the text sources. The output of the text analytics can be used to load fielded information into structured records stored in a relational database table. These records can be joined with records from other tables dealing with providers. In effect, an extract, transform, load (ETL) process was established for text sources, expanding

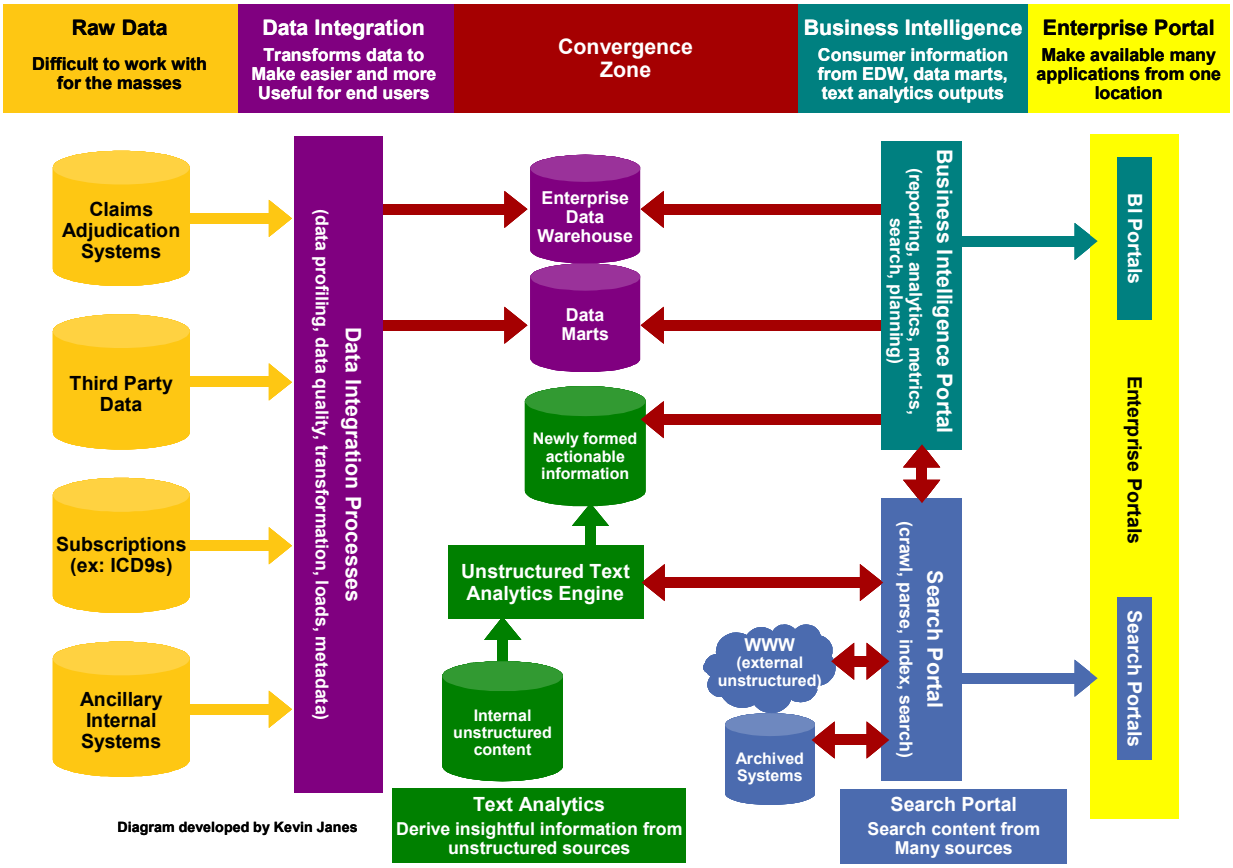
the range of sources that were used to form the integrated view of provider data in the data warehouse.

Once the transformation to structured data is accomplished, the data warehouse records are accessible via SQL-based tools — in this instance, the Cognos BI suite (as depicted in the box labeled "Business Intelligence Portal"). The resulting reports are available by launching applications in the portal. These reports and charts are more valuable because of the deeper insights gained:

- Across the network of providers (e.g., how often particular types of issues are raised on a claim)
- On a particular provider (e.g., how costly it is to service and support this provider versus others in the network)

FIGURE 2

Architecture for Proof-of-Concept Text Analytics Pilot



Source: BCBST, 2006

FUTURE OUTLOOK

The dual bridging strategies applied to provider reporting have resulted in the following benefits to BCBST, with the promise of additional benefits in the future:

- Providing a search interface to BI assets, both metadata and content (i.e., data values) improves the speed and access users have to business-critical information about BCBST providers.
- The proof-of-concept pilot showed BCBST that text analytics could become part of a viable strategy to enrich the data on providers. For example, a chart that showed the frequency of issues encountered by providers concerning payments opened the eyes of business managers at BCBST. The decision was made to move forward on a production implementation.
- Both of these information management benefits enable BCBST to improve its provider management process and ensure optimized payment strategy that minimizes cost while maximizing availability of healthcare resources.

ESSENTIAL GUIDANCE

BCBST shows the value of a carefully defined and executed strategy to improve information access. We see the following as success factors:

- **Clarifying lines of responsibility.** The DRM group is responsible for structured data access and building bridges to unstructured data — the work of the enterprise content management organization.
- **Defining and documenting a vision for an integrated information architecture.** Frank Brooks and the DRM group have maintained a document on strategic and tactical initiatives that has guided progressive phases of implementing the architecture. They have moved in a methodical, organized way from expanding the sources incorporated in the data warehouse to improved BI and search access to information via a comprehensive portal.
- **Prioritizing projects that focus on business operations with a significant financial impact to the firm.** The focus on provider reporting focuses on enhancing the information available to BCBST account reps involved in rate negotiations with their providers. This establishes the cost structure that is a key determinant of the financial health of the firm.

LEARN MORE

Related Research

- *Unified Access to Information: Content Vendors Heed the Urge to Converge* (IDC #202942, August 2006)

- *Unified Access to Content and Data: Delivering a 360-Degree View of the Enterprise* (IDC #34836, February 2006)

Published Under Services: Unified Access to Content and Data