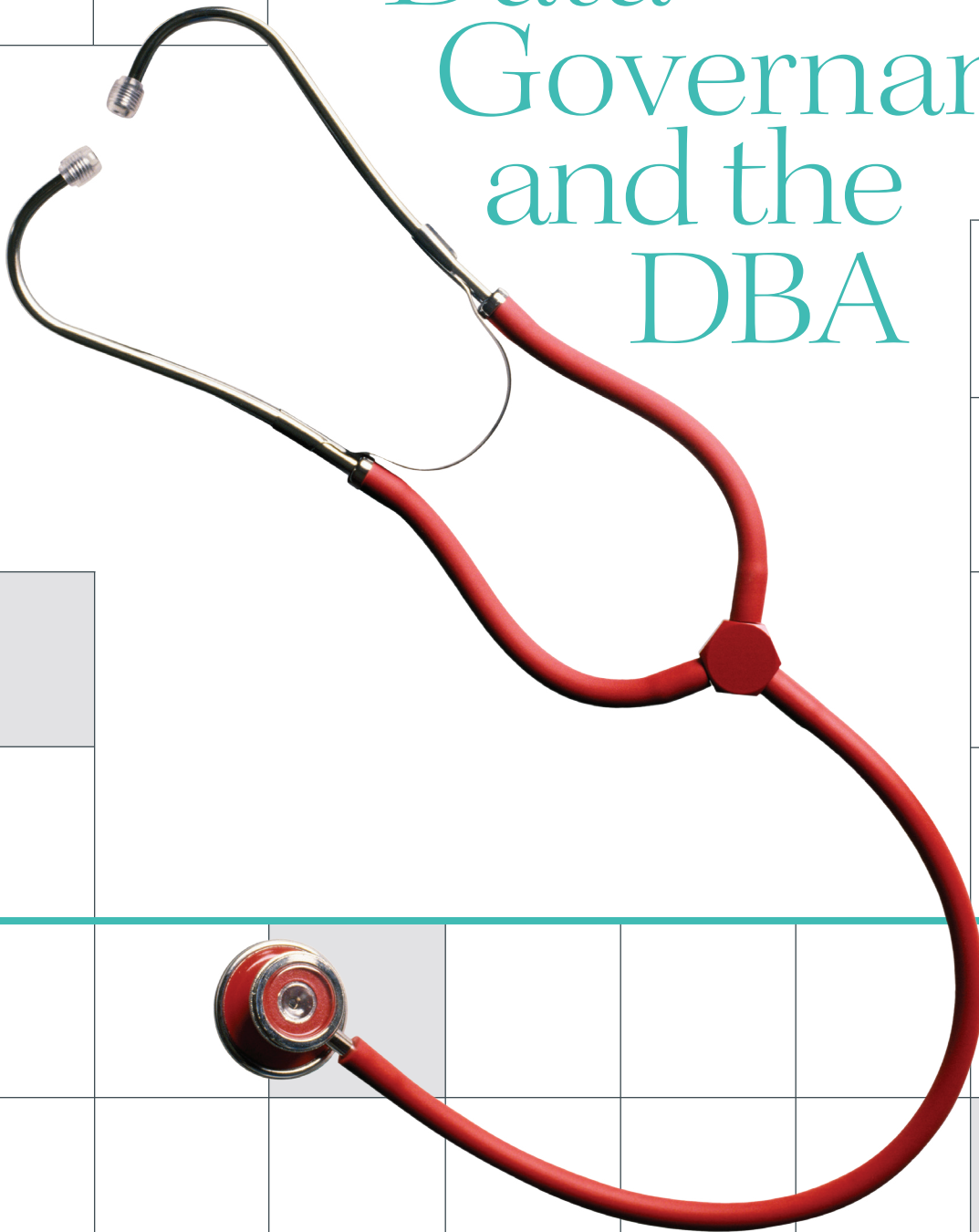


NEW WORLD, NEW WORK:

Data Governance and the DBA



Healthcare and insurance industries offer a glimpse of the DBA's future role

By Ives Brant

BACK IN 1986, HEALTHCARE GIANT BLUECROSS BlueShield of Tennessee (BCBST) did something unorthodox. At the time, most large companies organizationally separated their DBAs who did physical database design from the data administrators who created logical design. “We invited DBAs and data analysts to the same table, to decide together on the logical and physical design of our corporate database environment,” says Frank Brooks, director of data management and information delivery at BCBST. “We got into data governance without realizing it, just by defining commonsense practices related to our corporate data.”

That decision signaled a change in how BCBST—which today has more than 3 million member-customers and 16,000 company customers—would govern data and do business in the future. “Data governance is integral to better

healthcare delivery and reducing costs,” says Brooks. “We currently use a large repository of data from our operational systems to help us stratify, or classify, members. Then we can steer the member into a support group for their asthma or cardiac condition. Reliable classification lets us guide them to better health practices, and helps BCBST manage costs.”

The meeting back in 1986 also signaled a small but subtle change in the role that DBAs play at the company. Today, DBAs at BCBST still focus on changing data structures and adding new ones, but they are aware that SAS-70 rules and the Model Audit Rule give new importance to their work. The concepts behind data governance are changing business operations for organizations in healthcare and beyond. And as it turns out, the rise of data governance is also changing the roles and responsibilities of DBAs in important ways.

“We invited DBAs and data analysts to the same table, to decide together on the logical and physical design of our corporate database environment.”

— Frank Brooks

*Director of Data Management and Information Delivery
BlueCross BlueShield of Tennessee*

The beginnings of data governance at BCBST

Just about everyone depends to some degree on the healthcare and insurance industries, and few are as essential to our well-being. They hold private, vital data on nearly all of us, millions of businesses, and thousands of healthcare providers. That data must be captured accurately, governed, protected, and held confidential. By law, it has to be retrievable upon request.

Healthcare and insurance have been leaders in data governance, though not exactly by choice. They are among the most regulated industries, because they handle so much sensitive and critical data. The best-known regulation of data: the Health Insurance Portability and Accountability Act (HIPAA), which requires healthcare providers to keep patient data confidential, yet accessible.

BCBST dove into data governance before the term existed. In 1994, a BCBST white paper led to the creation of a simple data warehouse architecture. Structured source data collected by operational systems—on patients, medical claims, and providers—was fed through extract, transform, load (ETL) processes into the data warehouse. From there, the data was delivered to business intelligence (BI) tools for analysis, and through more ETL filters to various datamarts.

By 1996, BCBST clearly understood that its data was a valuable resource and worked hard on ways to measure and improve its overall quality. One of the earliest metrics that the organization developed is still in use today: a metrics scorecard that tracks data quality over time. For example, BCBST tracks for what percentage of the company's 3 million members the database holds a Social Security number. That unique ID is the best way to identify each member. The scorecard keeps evolving, and users give their input on which data quality metrics they need most.

From there, says Brooks, "Our data warehouse architecture drove our org chart." In 1996, BCBST aligned IT staff with database architecture, segmenting three groups from IT: database administration to store and manage data, data integration for ETL processes, and BI/performance management to support analytics.

The organization's need for ongoing data quality, and to extract more value from data, drove how it approached its IT projects. BCBST established a process flow for data governance initiatives, from defining the problem and getting sponsorship, to evaluating current maturity levels, all the way to metrics and results (see Figure 1). Each project includes maintaining a metadata repository, creating a data repository, and designating a data steward. DBAs work on data structures as usual and "apply good data governance principles as they do their jobs," according to Brooks. Some projects focus on specific governance issues, and DBAs may be tasked to find creative solutions, such as customized security for the data pertaining to a particular customer.

The DBAs' understanding of what a piece of data means in the different departments and subsidiaries, as well as an awareness of semantic discrepancies, is valuable knowledge. Data that is captured correctly and that follows a uniform structure can help a liability insurer to reduce risks. For example, vehicles and other assets at a single site might belong to different corporations, but analyzing risk by location would alert the insurer to excessive exposure to a disaster striking that single site. A business glossary (that is, what does "location" mean?), master data management, and data validation can all play a role in recording the location of insured assets perfectly.

Project step	Tool applied
Define problem, obtain sponsorship, evaluate maturity	Data governance maturity questionnaire
Build road map, sample constructs, build data dictionary	InfoSphere Business Glossary
Understand data	InfoSphere Information Analyzer
Create data repository	InfoSphere Metadata Workbench
Define metrics	Cognos data governance scorecard
Appoint data stewards, manage data quality, implement master data management	InfoSphere QualityStage
Measure results	Cognos scorecard

Figure 1: How BCBST carries out data governance projects: actions and tools

Datamarts for specific groups of users

BCBST built a data governance application to balance and control updates to the data warehouse. The IT group runs about 1,000 processes monthly with this application, as data moves from BCBST’s operational systems and external sources to the data warehouse, then to datamarts and BI analysis, building data cubes. As the data moves, it is checked intensively. Once in the data warehouse, the financial data is reconciled back to the general ledger. Then, to keep the financial datamart in sync with the general ledger, once data arrives there it is reconciled back to the data warehouse, down to the penny. “Our tools monitor everything that we do,” notes Brooks.

BCBST’s data-driven perspective has brought cost savings. “If you have the right information, and it’s timely and stored right, then you can build anything on top—any logical process,” says Brooks. “We developed a financial datamart, then showed Finance how to logically order and use their data elements for financial analysis.” With its closed loop of common ETL processes and a robust BI infrastructure in place, BCBST enabled financial performance management by adding just one column to a single table. “It saved \$5 million versus the approach recommended by outside consultants, and we delivered it faster,” Brooks says. A coup for the home team, and an illustration of why data governance has a silver lining for DBAs. Uniformity and clarity, and alignment with the direction of the business, make it easier to complete new initiatives.

BCBST has found it effective to organize its people to address overall enterprise architecture organization in ways that reflect the teams working on its data governance (see Figure 2). Both draw upon executive, middle management, and practitioner staff. The relationship between IT and the business has evolved as well. Executives on a data governance committee provide oversight, setting policy and standards. They look for new ways to leverage BCBST data. Then, an information management committee decides how to align IT with the business priorities, and it balances demands from constituents. Finally, an execution group that includes IT and business users gets it done; this group includes the core integration, administration, and BI/analytics teams described earlier. In this latter group, DBAs contribute frequently on security to protect data confidentiality, as well as on backup and recovery considerations. “The DBAs serve an important role as our last line of defense in protecting our data assets,” says Brooks.

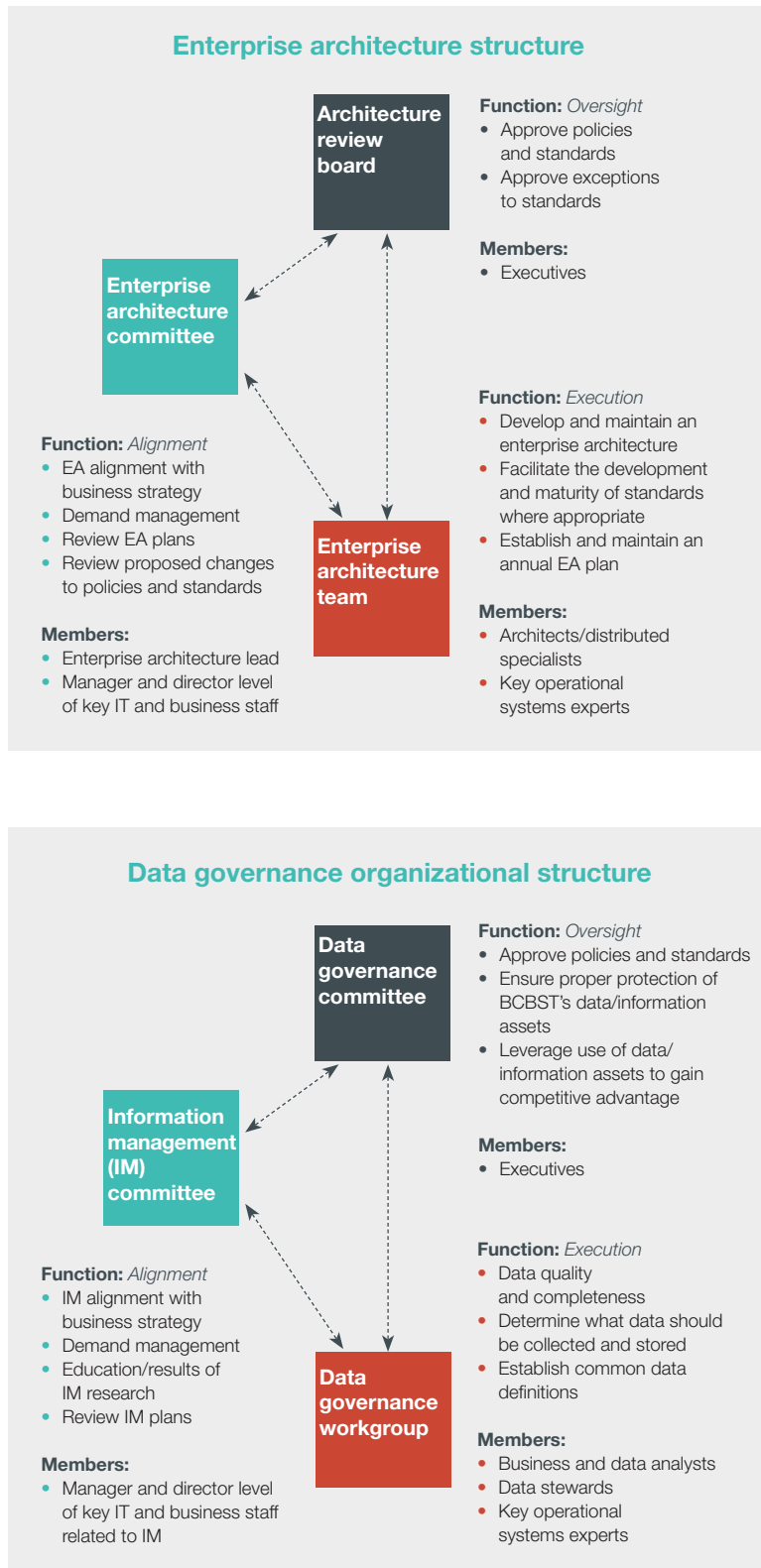


Figure 2: BCBST aligns its people in similar ways to address data governance architecture and enterprise architecture effectively.

Is the data ready to go out the front door?

Six years ago, BCBST started giving analytic access—in the form of IBM Cognos cubes—to its largest customers on the health statistics of their enrolled employees. “These customers are self-insured, so they are at risk. To improve the healthcare of their employees but limit costs at the same time, we enabled them to analyze their data 24x7 and store their reports,” Brooks says.

“We got into data governance without realizing it, just by defining common-sense practices related to our corporate data.”

— Frank Brooks

*Director of Data Management and Information Delivery
BlueCross BlueShield of Tennessee*

Security was a challenge that involved the DBAs. “We had to architect this and be careful how we loaded the cubes, so each customer can see only the information on its employees. We protect the data cubes, firewall them, and let each customer analyze their own,” Brooks says.

“Insurers are becoming very involved in the healthcare of their members, on a close to real-time basis,” Brooks continues. “The master data will give us a 360-degree view on everything about a member, account, or provider. The right data architecture lets us show all of it in near real time.” That still requires a consistent, accurate combination of all data pertaining to a customer. Semantic discrepancies must be ironed out, data errors avoided, and duplicates eliminated.

DBAs at BCBST contribute to this evolution in four governance-related aspects. They provide effective logical and physical models, appropriate indexes, and ongoing tuning to ensure efficient loading and access. Second, they

provide views and consulting to address ease of use and usefulness of the data. Next, they work with the data integration team and production application staff to ensure the overall integrity and quality of the data, so information derived from the data can be trusted. Finally, they are involved in applying records management policies related to archiving and purging data.

What's ahead: Text analytics and in-database predictive models

“In the future, we will add text analytics to gain new insights from our unstructured data, such as comments fields and certain fields embedded in electronic records. The results of text analysis combined with existing structured data will enable more powerful, accurate predictive models,” says Brooks. BCBST is designing a new data mining and predictive analytics environment. “We already develop predictive models with SAS, but a more structured architecture will let us add in-database, near-real-time, predictive analytics.”

Brooks sketched out one possible way to combine real data with what-if scenarios: “BCBST’s patient-facing staff can gather enough data for us to offer the member a forecast—on the phone call—of his life span. ‘Sir, you’re 50 and based on your profile, you have an additional 12 years of expected life.’ Then the kicker: ‘Sir, would you like to know what changes in weight loss, exercise, and smoking will add 10 years to your expected life?’”

An architecture that reflects the imperatives of data governance

Concludes Frank Brooks, “To proactively engage with its customer base, a business needs data governance as a foundation of its IT structure.” The overall database and data warehouse architecture—not just policies, roles, and tasks—should reflect the imperatives of data governance. With that consistency, IT can confidently move to a data hub environment like BCBST, and can have the ability to serve up analytics to major customers and individual history and predictions to members, give near-real-time performance management, and capture and leverage both structured and unstructured information. And as the story at BCBST shows, DBAs bring knowledge critical to the practical implementation of data governance initiatives. ✱

Ives Brant has experience in the database and analytics industries, and was formerly editor-in-chief at Tornado Insider magazine, a European competitor to Red Herring. These days, he writes technical content for a wide range of companies.