premium

VOLUME 21 • NUMBER 4

Business Intelligence Journal

THE LEADING PUBLICATION FOR DATA WAREHOUSING AND ANALYTICS PROFESSIONALS

BI Project Success Is in the Eye of the Beholder 5
Hugh J. Watson, Hauke Heier, Hans P. Borgman, and Fabiano G. Neves

The Articulate Architect: A Practical Approach for Communicating With Business Stakeholders 10
David Hendrawirawan, John Lucker, and Courtney Parry

Fighting Fraud with Advanced Analytics 19
Troy Hiltbrand

Drone Safety Calls for Complex Event Processing, Operational Intelligence 25
Linda L. Briggs

Outthink Cognitive Hype: Creating a Business-Driven Cognitive Strategy 28
Steve Williams

BI Experts' Perspective: Avoiding Storms in Your Move to the Cloud 37
Ravi Chandran, Norman C. Nicholl, and Tracy Ring

Using Lean Methods to Advance the Business Intelligence and Analytics Organization 44
Timothy Sullivan; Eric Hixson, Ph.D.; Andrew Proctor; Christopher Kucharik; and Timothy Crone, MD

Winners: TDWI Best Practices Awards 2016 52

tdwi
Advancing all things data.

TDWI is your source for in-depth education and research on all things data.

tdwi
Transforming Data With Intelligence™

TDWI ONSITE EDUCATION

Your Team, Your Location, Our Instructors

TDWI Onsite delivers the highest-quality business intelligence (BI) and data warehousing (DW) education directly to your office so each member of your team learns the same best practices, methodology, and strategy directly from the industry gurus. **Maximize your training budgets today. Schedule a free consultation.**

Core Tracks:

- Data Asset Management
- Core BI Skills
- Data Analysis and Design
- Big Data
- Agile BI and DW Development
- Leadership and Management
- CBIP Certification

Contact:

Yvonne M. Baho
Director, Onsite Education
Phone: 978.582.7105
Email: ybaho@tdwi.org

tdwi.org/onsite

VOLUME 21 • NUMBER 4

Business Intelligence JOURNAL

3 From the Editor

5 BI Project Success Is in the Eye of the Beholder
Hugh J. Watson, Hauke Heier, Hans P. Borgman, and Fabiano G. Neves

10 The Articulate Architect: A Practical Approach for Communicating With Business Stakeholders
David Hendrawirawan, John Lucker, and Courtney Parry

19 Fighting Fraud with Advanced Analytics
Troy Hiltbrand

25 Drone Safety Calls for Complex Event Processing, Operational Intelligence
Linda L. Briggs

28 Outthink Cognitive Hype: Creating a Business-Driven Cognitive Strategy
Steve Williams

37 BI Experts' Perspective: Avoiding Storms in Your Move to the Cloud
Ravi Chandran, Norman C. Nicholl, and Tracy Ring

43 Instructions for Authors

44 Using Lean Methods to Advance the Business Intelligence and Analytics Organization
Timothy Sullivan; Eric Hixson, Ph.D.; Andrew Proctor; Christopher Kucharik; and Timothy Crone, MD

52 Winners: TDWI Best Practices Awards 2016

56 BI StatShots

TDWI CERTIFICATION

Get Recognized as an Industry Leader

Advance your career with CBIP

"Professionals holding a TDWI CBIP certification command an average salary of $120,851—$15,000 higher than the average for noncertified professionals."

2016 TDWI Salary, Roles, and Responsibilities Report

cbip
Certified Business Intelligence Professional

Distinguishing yourself in your career can be difficult yet rewarding. Let your résumé show you have the powerful combination of experience and education that comes from the BI, DW, and analytics industry's most meaningful and credible certification program.

Become a Certified Business Intelligence Professional today! Find out how to advance your career with a BI certification credential from TDWI. Take the first step: visit tdwi.org/cbip.

tdwi.org/cbip

tdwi

From the Editor

As Strother Martin's warden told Paul Newman's troublesome prisoner in the 1967 film *Cool Hand Luke*, "What we've got here is failure to communicate." Communication is vital to BI project success and project management is the subject of several articles in this issue of the *Business Intelligence Journal*.

Authors John Lucker, David Hendrawirawan, and Courtney Parry examine causes of—and solutions for—communication disconnects between data architects and their audiences. They use a case study from a large U.S. bank to explain how their best practices worked.

How do you measure project success? Is it more than just completing a project on time and under budget? Senior editor Hugh J. Watson and coauthors Hauke Heier, Hans P. Borgman, and Fabiano G. Neves examine a new study about what influences perceptions of project success and add insights from their own experiences.

Our 2016 Best Practice Awards winners are further evidence of the range of project success your colleagues have enjoyed in a wide variety of industries.

Security is another focus of this issue of the *Journal*. Troy Hiltbrand explains how advanced analytics can automate the discovery and prevention of fraudulent transactions, saving enterprises millions. He explores the need for creating a training data set to help us understand today's fraud by examining past activity, and how models must change as new fraud methods are employed.

Ravi Chandran, Norman C. Nicholl, and Tracy Ring examine the dynamics of moving to the cloud, including addressing questions of security.

Linda Briggs explores why operational intelligence is key to commercial drone adoption and discusses how improving drone safety helps organizations take drones beyond the "visual line of sight" current regulations require.

What do you do with all the data a drone collects? Steve Williams looks at the "cognitive era," when unstructured data is converted into usable information for decision support. He discusses how to formulate a successful strategy for creating a next-gen decision support system.

Timothy Sullivan, Eric Hixson, Andrew Proctor, Christopher Kucharik, and Timothy Crone explain how the Cleveland Clinic used lean methods to become more effective and efficient in their business intelligence.

We welcome your comments at jpowell@tdwi.org.

James E. Powell

Business Intelligence Journal

VOLUME 21 • NUMBER 4

tdwi.org

[ISSN 1547-2825]

Reaching the Staff

Staff may be reached via email, telephone, fax, or mail.

Email: To email any member of the staff, please use the following form: FirstinitialLastname@1105media.com

Renton office (weekdays, 8:30 a.m.–5:00 p.m. PT)
Telephone 425.277.9126; Fax 425.687.2842
555 S Renton Village Place, Ste. 700
Renton, WA 98057-3295

Corporate office (weekdays, 8:30 a.m.–5:30 p.m. PT)
Telephone 818.814.5200; Fax 818.734.1522
9201 Oakdale Avenue, Suite 101, Chatsworth, CA 91311

Business Intelligence Journal **(article submission inquiries)**
Peter Considine
Email: journal@tdwi.org
tdwi.org/journalsubmissions

TDWI Premium Membership (inquiries and changes of address)
Email: membership@tdwi.org
tdwi.org/Premium-Membership
425.226.3053
Fax: 425.687.2842

BI Project Success Is in the Eye of the Beholder

Hugh J. Watson is a professor of MIS and holds a C. Herman and Mary Virginia Terry Chair of Business Administration in the Terry College of Business at the University of Georgia. He is a TDWI Fellow and senior editor of the *Business Intelligence Journal*. hwatson@uga.edu

Hauke Heier is a managing director at Accenture and an affiliate professor of information systems at the ESC Rennes School of Business. Heier is also an editorial board member of the *International Journal of IT/Business Alignment and Governance*. hauke.heier@accenture.com

Hans P. Borgman is a professor of information systems at the University of Amsterdam and a professor of business administration at Leiden University. h.p.borgman@gmail.com

Fabiano G. Neves is an independent consultant and Ph.D. candidate in business administration at ESC Rennes School of Business. fabiano.n.gomes@gmail.com

Hugh J. Watson, Hauke Heier, Hans P. Borgman, and Fabiano G. Neves

Introduction

Most BI professionals are familiar with project work; it's the essence of what they do. Whether the project is upgrading data warehouse hardware, selecting a BI tool, or developing an application, its success or failure impacts the careers of the people involved and the prospects of the organization.

Project performance is often judged by what is commonly called the Iron Triangle of project management success, shown on the following page in Figure 1 (Atkinson, 1999). It asks, "Was the project finished on time and within budget, and did it meet quality requirements for scope and function?" This is a very logical, normative model for project success, but it is also overly simplistic because it incorrectly assumes that project management performance criteria are set, measured, interpreted objectively, and remain constant over time.

In reality, "the eye of the beholder" has significant influence on what constitutes success and failure. You are probably familiar with projects that met quality, cost, and time criteria but were still considered failures, along with others that fell short in satisfying one or more of those criteria that were regarded as successes.

A recent research study on project management success (Neves, Borgman, and Heier, 2016) reveals a nuanced understanding of the factors beyond the Iron Triangle that influence perceptions of project success, as well as practical insights for increasing the likelihood of BI project success. In this article, we will share the research findings, some of our own experiences, and our concluding insights.

Figure 1. The Iron Triangle of project management success.

The Research Study

There is a large body of literature about the normative aspects of project success, but less is written about perception-based factors. In order to investigate these factors, the literature was reviewed and business, BI, and IT leaders were interviewed. The interviews were recorded and lasted between 60 and 90 minutes. The leaders were asked to think about a specific project and what made it a success or failure.

The study identified five factors that create a potential mismatch between perception and the success measures associated with the Iron Triangle (see Figure 2).

Expectations

Although managing expectations is clearly important, it requires a careful balancing act. Getting a project approved may require promises that set high expectations. When these promises are not fully met, it creates the impression that the project was unsuccessful. Research has found that unrealistically high expectations result in lower levels of perceived benefit than those associated with realistic expectations (Staples, Wong, and Seddon, 2002). As one interviewee said:

> Expectations management (and people management) does affect the perception of success. If you do a poor job in these areas, you will affect the team's and the organization's motivation. I have seen BI projects where we worked hard, with blood, sweat, and tears, and where the end results were positive primarily because a good job was done managing people and their expectations.

One of the authors was involved in a BI project where the maxim "BI is a journey rather than a destination" was successfully used to manage expectations. Because of business need, it was important to roll out the system in 90 days. The problem was that some of the data needed to fully satisfy the system's information requirements would not be available that quickly.

The solution was to emphasize in the project proposal and management presentations that the system would be an ongoing development effort ("the journey") and that additional data, features, applications, and users would be added over time. When the initial version was rolled out, management and users knew that this version would meet some, but not all, of the requirements. As a result, they were not disappointed. When the deliverables and timeline for the later versions were met, everyone continued to feel that the project was a success. Guiding expectations was a key to success.

Client/Contractor Relationship

Some BI projects involve contractors or consultants. In this situation, the project manager's perceptions of success are influenced by the quality of work done by the contractors or consultants, as well as trust, professionalism, predictability, and mutual respect.

One leader said, "The way you (the contractor or consultant) deal with people ... may damage the relationship and the trust, which negatively affects the perception of project results." Another interviewee explained, "A successful project is one that fulfills the expectations of the

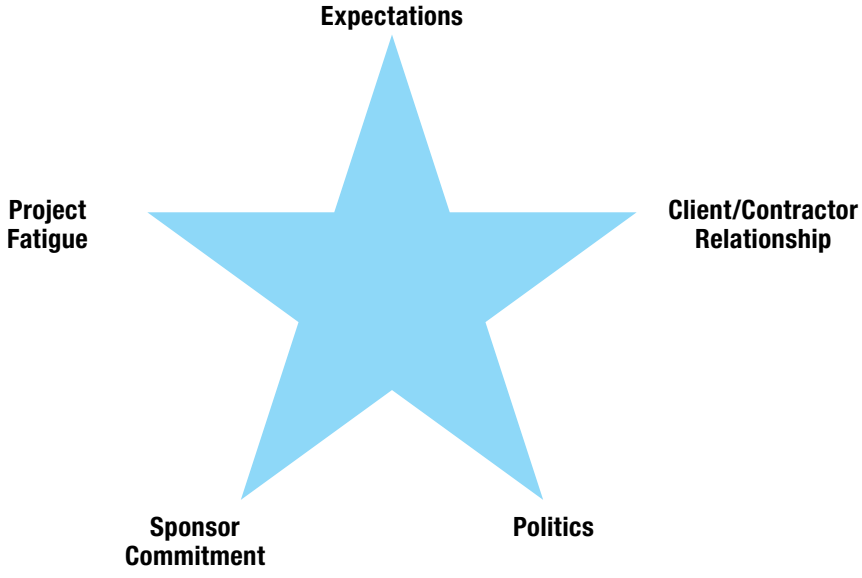


Figure 2. Five factors that affect perceptions of project management success.

client and the implementer and that reaches the financial estimations and deadlines."

When both parties benefit from the successful delivery of projects, they are more likely to engage in perception-building positive marketing.

Contractors and consultants often have an additional consideration when considering project success—whether the current project leads to future work. A consultant revealed:

> The project was delivered on time, on budget, and delivered the planned scope. ... Despite delivering what was agreed, the client just paid for the project and stopped the contractor relationship without further explanation. Continuation of the relationship is something that should have been targeted and was possibly a better indicator of actual success.

Overall, the client/contractor relationship is relevant to project success not only during the project but also afterwards (Bryde and Robinson, 2005).

Sponsor Commitment

It is universally accepted that having a committed sponsor is a key to BI project success. Ideally, the sponsor comes from the business unit where the work is being done, is high level (e.g., CMO), is well regarded in the organization, and is an opinion leader. Strong sponsorship can help both normative and perceived success.

Having continuing sponsor involvement in a project allows for quicker project functionality changes and other adjustments, if required. As one leader observed, "The sponsor can help revise the project objectives and the organizational pact if a change is needed." This can result in higher perceived success once the project is completed.

Senior sponsors can influence and form opinions. One interviewee said, "A senior manager can help show project relevance to the organization and create project respect." It was also observed that senior management commitment increases the overall probability of perceived success since people defend—or at least see in a brighter light—things they are directly involved in and have a stake in.

Good sponsors do not put too much distance between themselves and their projects by intentionally avoiding direct involvement in day-to-day details. Several respondents commented on this, saying that the project leader needs to be able to escalate important issues (such as gaining access to needed data) in order to foster success.

Politics

Organizational behaviors are not always rational; rather, they are influenced by organizational politics. People have reputations and careers at stake, previous experiences (good and bad) with people, resources to gain or protect, power to acquire (sometimes at the expense of another), and so on.

All of these political considerations can affect how BI projects are perceived and acted upon. Politics can influence whether projects are approved such that individuals or organizational units that are out of favor may have difficulty getting approvals. A project can be "contaminated" by virtue of the people involved. One interviewee observed, "The project image is very important. If a project is seen as something negative or problematic—even if it delivers on its promises—the perception seems to be less positive."

Politics can also prevent bad projects from being stopped. Individuals and organizations are reluctant to acknowledge project failure, both to avoid blame and to conceal the waste of resources. One executive illustrated this point:

> Organizations can have a difficult time making the decision to stop a project when they realize that the project is no longer needed or recommended. If you start a project and realize after the initial assessments that it should not proceed, you need to stop it, but frequently projects just continue. There seems to be a difficulty in making this kind of hard decision in an organization and admitting a failure in creating the project (or in the decision to allocate investments to it), so the project just goes on and the problem tends to become worse.

A different study interviewed project leaders of failed data warehousing projects. One of the most interesting findings was how quickly the sponsors "bailed" as soon as the project began to develop difficulties (Watson et al., 1999). They did not want their reputations tarnished by a failed project.

Project Fatigue

The conventional wisdom is that "big bang" projects that are expensive, take a long time to finish, and produce benefits only at the end should be avoided. If the project is large, it should be broken into smaller phases that deliver value with each phase. When projects take a long time—and especially when they take longer than anticipated—project fatigue sets in. This will likely negatively impact perceived success, regardless of how well the project meets the original objectives.

The longer a project takes, the more likely there will be changes in requirements, the development team, and user groups. There is also a decrease in satisfaction as people become tired of the project. One leader said, "Long projects usually interfere with the sense of satisfaction and people feel worn out. It is important to have delivery cycles that eliminate or minimize any dissatisfaction caused by projects taking a long time."

Another leader said, "Smaller projects do tend to present results earlier. ... Frequent checkpoints are surely better than a long BI project that only presents its results after a long time."

Long projects also tend to have a detrimental impact on the development team. As an interviewee observed, "You may deliver the project within schedule, functionality, and cost but have half of the team resigning due to the work environment and stress. ... This cost of losing

resources has to be considered." Another said, "Too much stress is bad. It causes attrition, wears people out, and puts a focus on the 'missing things.'"

What It All Means

It would be great if BI projects were approved and judged based strictly on their merits—cost, delivery time, quality, and benefits. This is seldom the case, however. As the recent research shows, other factors affect how project success is assessed and perceived. Mismatches between perceived and real project management performance occur. They are also influenced by the perceptions of industry peers or other organizational outsiders.

Seasoned BI managers and professionals have learned much of this through hard-earned experience:

- Manage expectations carefully.
- Break large projects into smaller parts that deliver value faster. Smaller projects or phases are also more likely to be perceived as successful (Mieritz, 2012).
- Different stakeholder groups have their own metrics for assessing project success.
- The factors that influence perceptions of project success are often out of the project manager's control.

Keep these and the other factors in mind as you undertake your next BI project. ■

References

Atkinson, R. [1999]. "Project Management: Cost, Time and Quality, Two Best Guesses and a Phenomenon, It's Time to Accept Other Success Criteria," *International Journal of Project Management*, Vol. 17, No. 6, pp. 337–42.

Bryde, D. J., and L. Robinson [2005]. "Client Versus Contractor Perspectives on Project Success Criteria," *International Journal of Project Management*, Vol. 23, No. 8, pp. 622–29.

Mieritz, L. [2012]. "Survey Shows Why Projects Fail," *Gartner*, June 2012.

Neves, F. G., H. P. Borgman, and H. Heier [2016]. "Success Lies in the Eye of the Beholder: The Mismatch Between Perceived and Real IT Project Management Performance," *Hawaii Conference on System Sciences*.

Staples, D. S., I. Wong, and P. B. Seddon [December 2002]. "Having Expectations of Information Systems Benefits That Match Received Benefits: Does It Really Matter?," *Information Management*, Vol. 23, No. 2, pp. 115–31.

Watson, H. J., J. G. Gerard, L. E. Gonzalez, M. E. Haywood, and D. Fenton [Spring 1999]. "Data Warehousing Failures: Case Studies and Findings," *Journal of Data Warehousing*, pp. 44–55.

The Articulate Architect: A Practical Approach to Communicating With Business Stakeholders

By John Lucker, David Hendrawirawan, and Courtney Parry

John Lucker is an advisory principal with Deloitte & Touche LLP, a Global Advanced Analytics and Modeling market leader, and a U.S. leader of Deloitte Analytics.
jlucker@deloitte.com

David Hendrawirawan is an advisory senior manager with Deloitte & Touche LLP with expertise in enterprise risk information architecture, strategy, and governance.
dhendrawirawan@deloitte.com

Courtney Parry is an advisory manager at Deloitte & Touche LLP, providing risk data management and regulatory compliance services to the banking industry.
cparry@deloitte.com

Abstract

A key challenge for data architects is to articulate their work in a way that is intelligible, relevant, resonant, proactive, and actionable to business stakeholders. This article takes a look at some of the root causes of disconnects in communication between data architects and their nontechnical audiences, and proposes some practices to consider for overcoming these disconnects by reframing the contents of discussion and encouraging active engagement.

Using a case study based on a data architecture function at a large U.S. bank, we will apply these practices to the challenges and demanding expectations that data architects often face. We will conclude with practical applications designed to tackle these issues and present illustrative artifacts.

Introduction

In the world of big data, social media, and the Internet of Things, the volume of data is growing exponentially and increasing in complexity and volatility. This acceleration makes the data architecture function even more critical as a foundational capability for managing risk and achieving greater performance. Many nontechnical business leaders understand this but have only vague ideas for how a robust data architecture can deliver value to their companies.

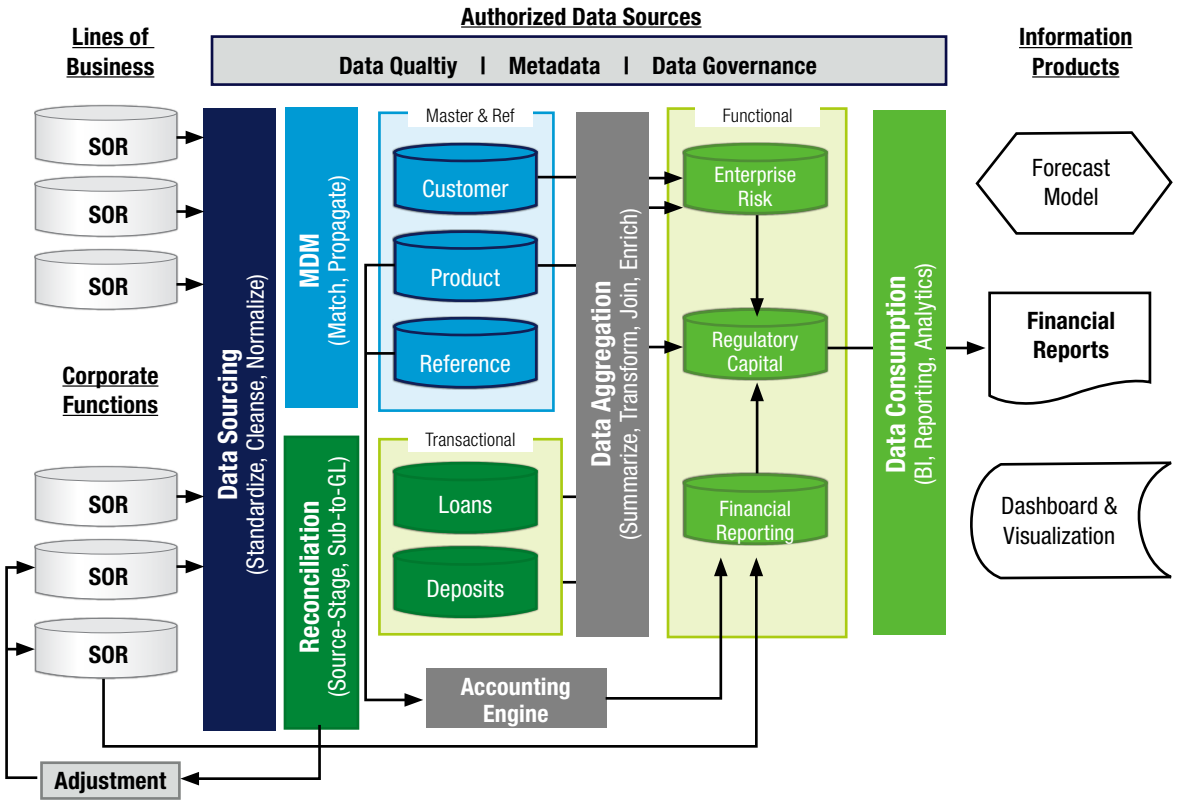


Figure 1. Example information-value chain.

It is no longer adequate for data and IT architects to solely focus on technical domains. Those involved in designing, managing, and executing data transformation must enhance their communication skills to effectively engage nontechnical business partners across functions. We wrote this article to share lessons from our experience providing professional services to leaders of data architecture, data governance, and BI strategy functions in enterprise data and analytics programs.

In defining data architecture, industry standards such as TOGAF, DAMA, and the Zachman framework provide accurate descriptions but are often unintuitive for nontechnical audiences. The following descriptions may not be formally correct but are likely to appeal to a broader business audience. Data architecture can be made up of one or more of the following elements.

Data model: A pictorial representation of business entities (objects, resources, locations, processes, or activities) and relationships, often complemented by a business glossary, taxonomy, and technical definitions and rules pertaining to those entities (often called *metadata*). The aim is to enable consistent use of information according to its intended purpose across an enterprise, and to add richer context to empower analytics use cases. A good data model improves the *effectiveness* of and *insight* derived from data.

Information value chain: A logical or conceptual map of information repositories and interfaces and how data flows from the point of entry to the final provisioning point. This includes capabilities such as data sourcing, aggregation, storage and archiving, advanced analytics, and business intelligence. With an information value chain, organizations are able to deploy the appropriate configuration of capabilities to improve the *efficiency* and *agility* of data processing.

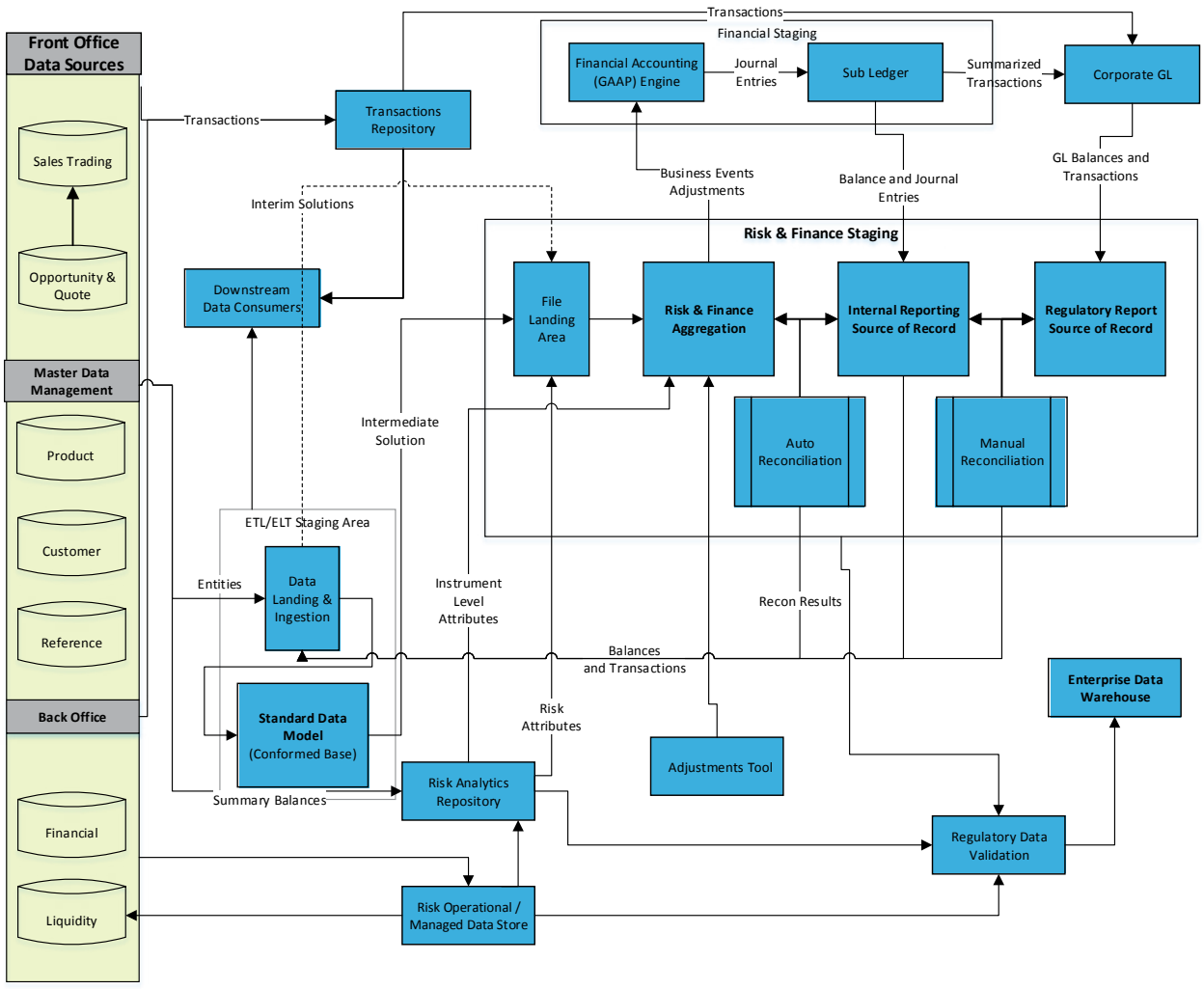


Figure 2. Example of a traditional architecture diagram.

Data strategy: A strategy for governing and executing the target data architecture. This includes clear, time-bound architecture goals, transition road maps, cross-functional program governance, and links between data architecture outcomes and the company's broader strategy. An effective data strategy creates *synergy* and *alignment* across analytics programs, as well as assigns *accountability* for achieving architecture goals.

Root Cause Analysis, Inherent Challenges, and Good Practices

There is a well-known stereotype that data architects (and IT practitioners in general) excel at technical analyses but communicate their work ineffectively. Over time, this can lead to disengaged business stakeholders. Root causes of many architecture design flaws include unclear expectations, poor teamwork, finger-pointing, and unhealthy politics.

In this section, we discuss two major themes to improve communication:

- **Reframing the Contents**—how to articulate data architecture activities, deliverables, benefits, and expected outcomes to business stakeholders

- **Encouraging Active Participation**—how to best define (and redefine) roles, participation channels, and conversation styles to engage stakeholders

Reframe the Contents

In the construction industry, the technical blueprint of a building is usually converted into illustrated marketing collateral for the prospective buyer. Similarly, data architecture artifacts should be converted into a form that speaks to nontechnical stakeholders. In no particular order, here are a few recommendations for creating such materials.

Provide Definitions and Context for Architectures and Governance

The term *architecture* may mean different things to those from different functions or parts of the organization. There are different types and levels of architecture and each is usually owned by a different team.

It is important that the roles of the various architecture owners are aligned and that a common framework is in place to connect, manage, and govern interdependencies. Without such alignment, architectures may be created in silos, creating unnecessary redundancies, incompatible or dysfunctional solutions, or even conflicting processes.

For example, in one large organization with a decentralized operating model, multiple departments had different architecture teams with overlapping responsibilities and differing interests. These teams got into frequent disputes and eventually stopped collaborating. The enterprise architecture (EA) team decided to engage each team and create an enterprise data governance forum across the organization, chaired by the EA leader.

The EA team proposed a reference architecture model and a set of principles to serve as a common language for the various functional teams. Rather than dictating the solution, the EA team used the reference architecture as a platform for collaboration toward a common long-term future. This enabled the architecture teams to cooperate more constructively without fear of losing power or being forced to adopt a specific design.

Make the Abstract More Intuitive to Nonarchitects

Architecture involves abstraction and the creation of a model that represents the real world using notational languages. These abstract models are designed primarily to help the designer or developer, but they are challenging to understand for nontechnical end users. A good approach is to substitute complex diagrams with simple images or to use intuitive terms that resonate with stakeholders.

In one workshop with a business stakeholder, the data architecture team presented a simple hierarchy of business terms by extracting entities and relationships from an entity relationship diagram (ERD), while omitting technicalities such as referential integrity and cardinality. This approach allowed the business subject matter expert to quickly grasp and confirm accuracy and completeness. The original relationship and ERD diagrams were then updated accordingly.

Another team engaged a design studio to create aesthetically appealing contents, typology, and infographics. They used a subway station map to illustrate the intersections and connections among architecture goals and strategic enablers. They also used gamification techniques to create a user-friendly map of data applications by representing data attributes with tokens. Invoking certain business events placed tokens on the correct application on the map.

Choose Dynamic Complexity over Complicated Minutiae

There is often a tendency among architects to solve problems with extensive analyses and documentation. Architecture specifications usually address and separately model each single dimension—process, data, application, organizational unit, and so on. Hence, even a simple model consists of multiple layers and diagrams. The result is often a voluminous and unintentionally complex set of artifacts.

Authors Jeanne Ross and Peter Weill pointed out that the reason "enterprise architecture core diagrams are tough to draw is because they force management to develop a simple vision of a complex organization. Agreeing on what not to include can be a challenging but fruitful exercise [as it] forces a clarification of a workable vision" (Ross, 2006).

The adage "less is more" is usually very appropriate when presenting to a nonarchitect audience. Cognitive psychologist Dr. George A. Miller's research (Miller, 1956) showed that the quality of judgment is dependent on the amount of information presented to the decision makers. According to Miller, it is generally best to discuss only between four to eight ideas at a time for optimum decision quality.

However, business stakeholders do not necessarily want a trivialized view—they want a holistic one. Such stakeholders can process voluminous (even complex) information—as long as there is a pattern or theme to help them understand the big picture. The art of architectural storytelling is about presenting dynamic complexity but not complicated minutiae. Key tips include:

- Start by outlining the four-to-eight takeaways you want the audience to remember
- Classify and summarize facts into logical groups
- Determine the right level of information according to the audience
- Sequence ideas in a progressive plot, from background to "story climax" to conclusion
- Incorporate a healthy dose of thought-provoking questions to stimulate participation
- Conclude with a strong call to action

Prioritize Business Value over Technical Jargon

Many technical professionals have admirably deep subject matter expertise. Specialization often develops with career progression and increasing focus on narrow topics or areas. An occasional side effect is that the professional gravitates towards using technical jargon and esoteric terms instead of commonsense language.

However, establishing oneself as a technical subject matter expert does not have to make one less relatable to business stakeholders. Architects should be wary of their "communication blind spots" and learn to articulate functional features, benefits, and business value simply. As an example, one enterprise BI team maintained a glossary to translate technical terms into commonsense descriptions, along with familiar user scenarios.

The adage "less is more" is usually appropriate when presenting to a nonarchitect audience. However, business stakeholders want a holistic view, not a trivialized one.

In another case, an organization created an add-on application to measure the frequency of "bull" (i.e., irrelevant, unclear, or opaque) words in documents (Fugere, 2005). Codenamed the Bull-Fighter, the premise of the tool was that a well-intentioned attempt to soften a difficult message can create vagueness, which can then produce a skeptical or apathetic response from the message recipient. Although not specifically designed for architecture terms, the same concept can be applied to excessive use of technical jargon, which an audience may perceive as posturing or aloofness.

Data architects should use phrases that reflect key strategic themes, goals, and metrics that are relevant to their business counterparts. For example, instead of "improved system uptime," the outcome of a BI infrastructure project can be restated as "reduced business interruptions." Customer data integration or master data management capabilities can be expressed as a "single source of truth for customer information" or a "360-degree view of the customer."

Encourage Active Participation

A 2005 research study by the MIT Center for Information Systems Research (Westerman, 2005) found that "gaining business value from IT requires more than great technology—it also requires firm-wide participation in targeting IT investment, ensuring IT performance, and IT-enabled process change." Therefore, "an effective CIO

makes a big difference" when he or she improves tasks that require joint effort between IT and business and helps business executives understand how to manage, oversee, and use IT effectively. Although the study focused on CIOs, the lessons are relevant for data officers and architecture leaders too.

Business leaders want data architects to bring technical skills, participate in strategic decisions, and coach others to navigate an unfamiliar discipline. The expectation has shifted from producing artifacts to partnering and guiding business stakeholders through the design and implementation of the architecture. Without getting the business involved in the design process, the proposed architecture may be seen as an "ivory tower" product that is out of touch with business reality. In that spirit, here are some methods that can go a long way toward cultivating trust and partnership between business leaders and data architects.

Teach the Trade-offs, Not the Trade

Nontechnical audiences appreciate receiving some education on data architecture topics, such as an introduction to the principles and patterns of good design and key trade-offs between architectural options.

To involve business partners in the journey, they should be allowed to provide input in choosing among alternatives, assessing pros and cons, and setting priorities. Even when a data architect already has a clear opinion on a matter, it is wiser to give business partners some influence, even if the final decision still rests with the data architect. Though this approach takes more time, it can generate a greater sense of ownership in final decisions.

Data Architect as a Collaborator and Bridge Builder

As previously discussed, there are often multiple teams owning overlapping architecture territories. The enterprise reference architecture can be used as a platform to promote cross-functional networking and collaboration across the various architecture owners.

At one firm, the enterprise data architecture and the data governance teams jointly initiated a cross-functional forum to create and maintain the enterprise reference architecture, including business entity definitions and standards governing data management and quality processes, and infrastructure technology. This enabled the architecture teams across functional groups to cooperate more constructively and to keep each other accountable for adhering to agreed-upon data architecture decisions.

Winning Hearts and Minds

Understanding workgroup formation and interpersonal dynamics is crucial to establishing a healthy environment for long-term collaboration. Soft skills such as storytelling, effective communication, negotiation, and conflict resolution are critical skills in making major architectural decisions because there can be significant budgetary and organizational implications with highly political dynamics.

Soft skills such as storytelling, effective communication, negotiation, and conflict resolution are critical skills in making major architectural decisions.

At one company, the EA team invested in soft skills training courses. They conducted a series of "lunch and learn" sessions on topics such as emotional intelligence, how to deal with difficult conversations, the cycles of group dynamics (forming, norming, storming, and performing), and how to avoid groupthink and maintain creative tension. They worked with management consultants to simulate stakeholder analysis and organize executive off-site meetings.

Break the Rules When Necessary

Sometimes breaking from the traditional systems development life cycle or the usual sequence of agenda topics can avoid deadlocks.

Sometimes breaking from the traditional systems development life cycle or the usual sequence of agenda topics can avoid deadlocks.

For example, the software development life cycle usually starts with current state assessment and then moves to the target-state design. While this may be the right order to conduct analysis, it does not have to be the order that the results are presented in. Presenting the proposed architecture as a "flashback"—that is, by starting with the simpler and more idealized target state and then showing the transition from the current, less-than-ideal state toward the target—can often be more appealing to the audience. This can accelerate the development process significantly.

Another example of effectively breaking with tradition is the example of an organization where a notoriously difficult business executive was criticizing a proposed target state architecture. The data architect asked the executive to draw what he imagined the picture should look like. It turned out what the executive had in mind was a strategic road map instead of a target state architecture.

Although a road map is typically created after a target state architecture is finalized, the data architect broke with tradition by acknowledging the executive's idea and taking it further by drawing connections between the strategic road map and his architecture diagram. By doing this, he was able to point out the obvious gaps in the road map drawing. The architect requested permission to review the draft and come back with a revised design in a follow-up meeting.

From that point on, the business executive became much easier to approach. By adapting the process, the data architect avoided unproductive churn, won the trust of his stakeholders, and provided a sense of co-ownership of the architecture artifacts.

Case Study and Application

The following case study is a composite the experiences of enterprise data architecture teams at several large banks.

As a Global Systemically Important Financial Institution—one of an internationally recognized list of institutions whose importance to the global economy puts them under stricter regulatory scrutiny—ABC Bank must comply with Basel Committee on Banking Supervision Regulation 239, known as BCBS 239.

BCBS 239 is a set of 14 principles meant to ensure accurate, comprehensive, and timely risk data aggregation and reporting. BCBS 239, Principle 2 requires ABC Bank to design, build, and maintain its data architecture and IT infrastructure to support the required data aggregation capabilities and reporting practices—including integrated metadata, master data management, adequate quality controls throughout the data life cycle, and others—both under normal conditions and under times of stress.

From 2013 to 2015, the Basel Committee found that major banks continued to struggle with complex and long-term data architecture projects, which has led to Principle 2 consistently receiving the lowest compliance scores since the inception of BCBS 239 (BCBS, 2015).

Amidst growing concerns over the BCBS 239 data architecture principle, ABC Bank hired a new chief data officer (CDO) to lead an enterprise data architecture initiative. Before his organization had even been fully formed, the CDO was already getting pressure to present his vision and plan to the members of ABC Bank's BCBS 239 committee, which was made up of corporate functions and line-of-business leaders. The data architecture team created four artifacts to articulate the CDO's data architecture vision, key deliverables, and expected outcomes.

1. **The vision and goals of data architecture**—the objective of this product was to provide context for the various architecture owners and illustrate the interdependencies among their activities, deliverables, and goals. The focus of this module was to facilitate agreement among architecture owners to collaborate and establish the operating model to do so.

2. **The business information taxonomy**—a classification of information objects, data definitions, and rules for proper use, the primary goal of which was to establish a framework for data ownership and accountability.

   To illustrate the taxonomy, a conceptual data model was flattened and the cardinality rules were removed to simplify the view. Each major branch of the taxonomy was considered a data domain and was assigned to an owner. Finally, a cross-domain data governance forum was chartered to maintain the taxonomy.

3. **The information value chain**—an end-to-end architecture of data capabilities from the point of entry to the final point of information provisioning intended to align data owners to a common reference architecture, thereby enabling consistent information processing and use across the enterprise.

   Key components were named in business-friendly terms rather than technical ones—for example, "standardized data sourcing" versus "ingestion," "adaptable integration" versus "ETL," and "efficient storage" versus "data warehouse." The information was presented in the form of a "flashback," as described previously.

4. **A draft strategic road map**—the road map had three stages that were purposefully framed to appeal to business common sense: foundational capability build, data and report migration, and legacy decommissioning.

Each data domain owner would have different prioritized timelines to account for factors such as urgency of the use case, the natural complexity, and the extent of known data quality issues in the domain. Interdependencies among data domains and alternative critical paths were highlighted. Each critical path alternative had a list of business benefits, rough sizing, risks, assumptions, and other factors considered as prioritization rationale. The draft road map served as a starting point for data domain owners to discuss and reach final consensus.

When the CDO presented these artifacts, they were well received by business stakeholders. The artifacts helped the business and technical teams to understand the goals and deliverables of data architecture in relation to the primary business objectives.

A data architect should focus on developing his or her articulation skills and consultative business acumen to interact more effectively with business stakeholders.

Conclusion and Final Thoughts

In summary, a data architect should focus on developing his or her articulation skills and consultative business acumen to interact more effectively with business stakeholders. After the architecture team held a "lessons learned" workshop, they identified the following insights and suggested follow-up steps.

- Business communication and soft skills training should be continuously provided at all levels, not limited just to leaders.

- When facing confusion or opposition from business partners, the architect's typical reaction is to overcome conflict by demonstrating his or her technical expertise and knowledge. Be wary of becoming "the smartest person in the room." Projecting expertise with humility can be far more engaging and effective.

- The first engagements are always the hardest, especially if business stakeholders have had very little previous exposure to architecture. Be patient and open to teaching and coaching.

- Learn as much as you can about each stakeholder's baggage, unexpressed agendas, and any underlying political and power dynamics in advance of key meetings. Interview past major project leaders, study the

history of legacy platforms, and identify the "tribal" silos, along with their biases and mental models.

- Aside from producing architectural deliverables, the architecture team should also consider its role in fostering a stronger partnership between business and technical stakeholders by leading an inclusive and effective architecture governance process.

Most importantly, be prepared to face setbacks and longer than expected timelines. Accept the fact that there will never be a perfect architecture—one that everyone accepts and that never changes. Adopt the mindset that the journey itself is just as important as the artifacts the team is producing. To borrow from General Dwight Eisenhower's idea about strategic planning, "The architecture is useless, but architecting is absolutely essential." ■

References

Basel Committee on Banking Supervision [2015]. "Progress in adopting the Principles for effective risk data aggregation and risk reporting," http://www.bis.org/bcbs/publ/d348.htm

Fugere, Brian, Chelsea Hardaway, and Jon Warshawsky [2005]. *Why Business People Speak Like Idiots: A Bullfighter's Guide*, Free Press.

Miller, George. A. [1956]. "The magical number seven, plus or minus two: Some limits on our capacity for processing information," *Psychological Review*, 63, 81–97.

Ross, Jeanne, Peter Weill, and David C. Robertson [2006]. *Enterprise Architecture as Strategy*, Harvard Business School Press.

Westerman, George and Peter Weill [2005]. "What Makes an Effective CIO? The Perspective of Non-IT Executives," *MIT CISR Research Briefing*, Vol. V No. 2C.

Fighting Fraud with Advanced Analytics

Troy Hiltbrand is chief digital officer at Kyäni and an associate editor of the *Business Intelligence Journal*.
thiltbrand@kyanicorp.com

Troy Hiltbrand

Abstract

Business fraud is on the rise and businesses need to have confidence that the person on the other end of the transaction is legitimate and that the transaction is valid. With a large volume of transactions happening in real time, businesses have to quickly identify which are legitimate and which are fraudulent. Advanced analytics provides the platform and method to accomplish this.

Advanced analytics provides a mechanism to effectively automate the discovery of transactions that don't quite "look right." Uncovering such fraud can save companies millions of dollars in unrecoverable sales.

Introduction

From increased globalization and stricter regulation to a consumer community hungry for personalized digital interaction, doing business today is a challenge. As the business environment grows ever more complex, so does the increase in digital fraud. Not only are consumers becoming more digitally literate and interacting with businesses in entirely new ways, so are criminals looking to identify weaknesses in corporate defenses. In the digital arena, this nameless and faceless threat can make companies feel helpless, but this is not the case. Advanced analytics can provide businesses with the weapons they need to combat this growing problem.

With identity theft and credit card harvesting becoming more common, there is a corresponding increase in the use of these stolen identities and payment methods against businesses. Criminals look for ways to monetize the information they have acquired—opportunities where it can be turned into cold, hard cash—and businesses are often left liable for the costs.

There are two main costs that businesses have to cover. First, the true owner of the breached financial

account has to be refunded the money that was illegally compromised. Financial institutions are good about acting as advocates for their customers in this regard. Second, however, the lost inventory and freight costs of the transaction also have to be covered. Often, both of these costs land squarely in the court of the business that accepted the fraudulent transaction.

Businesses need to have confidence that the person on the other end of the transaction is legitimate and that the transaction is valid. With a large volume of transactions happening in real time, businesses have to quickly identify which are legitimate and which are fraudulent. Advanced analytics provides the platform and method to accomplish this.

These methods and techniques can be highly effective when applied to fraud detection, but they also can be applied more generally to other business problems where the unknown has to be predicted with a high degree of probability.

Understand the Business

Deploying advanced analytics for fraud management requires that the business fully understands the business situation and deploys the set of analytics that is optimized to address their specific challenges.

The first step in this process is to understand what the perpetrators of fraud are doing. This will be highly dependent on the business and its policies and procedures, including returns, exchanges, and payment options.

Those committing fraud have an agenda. Most often, this has to do with using information to generate financial rewards for themselves at the expense of others. Another common cause of fraud against a business is a perpetrator's personal conflict with the business. This could be based on the business's stance towards a social or political issue, or based on a soured relationship with the perpetrator.

Once you understand the end game of those committing fraud, you can start to look at historical transactions and identify where fraud has happened. If it has happened before, it will probably happen again. The patterns of historical transactions form a basis for anticipating future fraud. Criminals don't like to change their modus operandi unless they have to, so the past is a pretty good indicator of the future.

Unfortunately, with financial transactions, fraud often manifests itself only weeks or months after the transaction. Therefore, a set of relevant data must include many months of transactions.

Unfortunately, with financial transactions, fraud often manifests itself only weeks or months after the transaction. This is when the person whose identity was compromised comes forward to dispute the transaction with his or her financial institution, which usually ends in a chargeback to the business supplying the good or service. Therefore, pulling together a set of relevant data must include many months of transactions so that the patterns of historical fraud can be adequately defined.

Define a Training Data Set

To learn from the past and predict the future, you need data to develop and train an analytics model. Once you have a business understanding of how a criminal might be using transactions against the company and examples of where it has happened, you are ready to start processing the data.

As you analyze fraudulent transactions, it is essential to develop a profile for each one. This will include information related to the specific transaction and patterns of how a particular transaction relates to others.

During this step, your company will often discover that certain data that would be highly beneficial to include

in the profile is not readily available. For example, with credit card transactions, perpetrators often test out the card by making a few small transactions with it at different businesses before making a large transaction. Raw information about where the card has been used prior to the transaction being investigated exists, but is often outside the purview of what your business controls.

In this case, a surrogate attribute might be required in the transaction profile that can represent suspicious activity that exists outside of the business. Many financial institutions have a process for prescreening transactions. As part of the prescreening process, these financial processors will return data in the form of a flag or score that may be used as this surrogate attribute in the transaction profile.

Depending on the completeness of the data environment and the organizational barriers that exist (whether technical or political), pulling together an accurate profile of fraudulent transactions can be extremely difficult, often requiring time and collaboration across organizational divisions. Some of the data that is most relevant to identifying fraud is sensitive in nature and must be carefully governed to ensure that its use doesn't expose the company to further liability.

Once you finish profiling these transactions, you will have a large list of potential attributes associated with them. At this point in the model development process, you do not have to define the rules associated with determining if a transaction fraudulent or not; you are just arranging the data so that it can be used to develop a statistical model later.

This step might generate hundreds of attributes associated with each transaction. Some of them will have no bearing on defining the model. Others will only be applicable when used in context with another attribute or set of attributes. Still others will be highly correlated to whether the transaction is fraudulent. During the data preparation stage, the goal is to identify a good set of relevant attributes, centralize these attributes, and cleanse the data so that it is reliable for modeling.

Given the costs associated with covering both sides of a fraudulent transaction, even a small amount of fraud can be very expensive for the business. However, examples of true fraud are usually dwarfed in a sea of valid transactions, so it can be challenging to get a good sample size of fraudulent transactions with which to develop your model. At times, this will require that the training set be disproportionately loaded with fraudulent transaction records as compared to a full data set. This will allow the model to be sensitive enough to be able to identify fraudulent transactions in the future.

As you dissect the problem statement and the business environment, your list of factors may include:

- Transaction amount
- Time of day and day of the week when the transaction was made
- Historical quantity and quality of transactions from the same or similar customers
- The true location of the transaction's origin, obtained from attributes embedded in the network packets that created it, not necessarily the information entered by the ordering party
- Frequency of attempted transactions

These will make up the transaction profile.

Within the data set, there is one critical attribute—the indicator of whether the transaction is fraudulent or not. As this is not always available in the transactional system, this will often have to be engineered using a combination of other data sets. Typical sources of data used in constructing the fraud indicator include financial records related to chargebacks, returns and metadata related to the reason for the return, and customer notes. With data such as customer notes, the content is often unstructured and requires preprocessing so that it can be used in the model development process. The resulting fraud attribute needs to have only two potential values: TRUE or FALSE.

In advanced analytic modeling, this fraud indicator is known as the *target attribute*. A target attribute is one that is known during the modeling process, but unknown during the process of prediction. With fraud, the goal is to develop a model that will ascertain this fraud attribute from a set of known attributes.

To enable real-time detection, the profile attributes need to be limited to only those pieces of information that are known at the time of the transaction.

One word of caution: as you identify attributes associated with fraud, ensure each attribute is one that would have existed at the time of the transaction.

Take an example where returns are part of the fraud pattern. Returns only happen after the transaction. If these are part of the model, real-time fraud detection will be impossible because the model will be optimized to take a longer-term position and will only be able to identify fraudulent transactions once the full picture—including returns—is known. To enable real-time detection, the profile attributes need to be limited to only those pieces of information that are known at the time of the transaction.

Building the Model

The term *advanced analytics* is often misrepresented as a single method or technology. In reality, it represents a category of approaches and not a specific technology. Advanced analytics includes multiple algorithms and algorithmic approaches that use mathematics and statistics to extract unknown information from a known set of data.

When using advanced analytics to detect fraud, the goal is naturally to determine if a transaction is fraudulent. This could be in real time as the transaction is happening or in near time prior to the completion of the transaction and the delivery of goods or services to the perpetrator. When fraud is headed off before the transaction is complete, the costs of lost inventory can be avoided. Additionally, declining the transaction can often eliminate the future chargeback on the account and its associated fees and penalties.

Fraud detection most often falls into a category of analytics known as *supervised learning*. Supervised learning techniques create a model by iteratively cycling over the data, optimizing the model's performance by adjusting its parameters.

With fraud, the model's target is to identify likely fraudulent transactions without throwing too many false positives. True positives caught in a timely fashion can save the company money, but false positives can deter legitimate business activity and negatively affect relationships with valid customers.

As a result of this supervised learning process, you create a process by which a known set of inputs is transformed into a prediction of the transaction's fraud attribute.

In the field of supervised learning, the analytics community has identified multiple techniques for accomplishing this model development and optimization. Depending on the technique, the resulting models range from being easy to visualize and explain to highly complex, where the business has to treat the model as a "black box" and accept understanding the inputs and the output, but not the transformational model.

Methods such as decision trees are easy to lay out graphically, which makes it easy to walk people through the process from raw inputs to fraud decision. The decision tree uses a set of divisions in the data to lead to a final answer. The decision tree can be simple, only dividing the data into a couple of sets, or it can be multiple levels deep—using different attributes to subdivide the data into smaller and smaller groups, each representing a defined output. As the target variable is either TRUE or FALSE, the decision tree algorithm will have multiple routes to each potential answer.

Other methods, such as neural networks and support vector machines, are more complex and so more difficult to track back from a result to the raw data inputs.

Ensemble methods use multiple techniques jointly to transform raw inputs into a fraud decision. To fully understand the result, you must understand both the ensemble scoring method as a whole and each of the methods used to generate input for it.

Some analytics techniques only function with certain types of data. For instance, methods using vector mathematics and linear algebra to develop the model will often require that all of the attributes be numeric. Such methods would require that attributes that are not numeric either be converted from categorical information into numeric values or be left out of the model development process.

As these techniques and their resulting models are based on known statistical and mathematical concepts, even these complex algorithms can be built from scratch with every part of the "black box" understood. The challenge with this approach is that it can be a costly investment for a company whose goal is to simply develop an effective model that will allow it to predict whether transactions are fraudulent in a timely fashion. This is where you have to understand the relationship between the cost of developing a model and the value it provides the business by heading off fraudulent activity.

The fastest way to develop these fraud models is not by coding the model generation from scratch but through the utilization of a platform which has prebuilt tools for performing this function. Leaders in this space include SAS, RapidMiner, IBM, Microsoft, and Oracle. Another popular tool for developing fraud models is R, an open source platform developed and maintained by a community of analytics practitioners. There are companies, such as Revolution Analytics, that provide commercial support for the R language.

R and Python, another open source programming language, have together become very popular among the advanced analytics community for developing models. It is not always as easy to program models in R and Python as it is with the commercial platforms, but the entry cost for the technology is much more attractive.

The fastest way to develop these fraud models is not by coding the model from scratch but through the utilization of a platform which has prebuilt tools for performing this function.

As you select a tool for developing fraud models, there are a couple of key considerations that need to be part of the tool selection process.

- Different technology platforms used in the model building process will expose differing levels of insight into what the model is doing.

  Some will generate code that can be implemented in multiple computer languages and systems. This code can be broken apart and analyzed to understand what is happening. Others provide only key attributes of the model and the rest remains entirely inside the tool. You will need to determine how much transparency you will require into the model being run.

- It is critical to identify a tool your organization either has or can develop skills with to build the model.

  Effectively acquiring and preparing data, developing a model and deploying it in production are all job functions of what is now known as a data scientist. This role is highly sought after by organizations due to the potential impact of advanced analytic models in optimizing business. This also creates a relative scarcity in the market for individuals who possess these skills. Due to this scarcity of skills in the market, it is important to factor your existing and targeted knowledge of both the tool and process into tool selection.

Testing for Accuracy

The magic in advanced analytics comes not through simply developing a model, but iteratively testing that model and refining it. During this process, alterations to the data used to train the model and optimization of the parameters used to configure it drive the model to be continually better. Success is achieved when the model is able to accurately identify true fraud, while minimizing false positives.

When measuring the model's effectiveness, it is important to test using data that was not part of the data used in the development phase. During the development process, there is a tendency to over fit the training data. As a result, the model can produce spectacular results when run against the same data that was used to train the model, but will fail miserably when assessing real world data.

Deploying the Model into Production

Once the business process has been analyzed, a profile created for existing fraud cases, and a model developed, tested, and refined, the next step is to roll out the model into a production environment and allow it to process real transactions.

Oftentimes, practitioners believe that once a model is tested and deployed, it will run forever and continue to return the same level of quality results. The problem with fraud is that as it is uncovered and prevented, perpetrators evolve their tactics, and existing models will be ineffective against these new methods of fraud.

There is no predefined amount of time that a fraud model will run effectively because it depends on how quickly the market evolves as the model exposes and stops fraudulent behavior from happening. In cases where fraud is highly lucrative, models might have to be rebuilt on short time frames such as weeks or months. In less lucrative cases of fraud, the business might be able to go for quarters or years before having to rebuild their models.

The best indicator of a model's effectiveness is to track how many cases of fraud continue to pass through the established filters. As the fraud starts to increase to unacceptable levels, it is an indicator that the models need to be recreated to better address the fraudulent behavior in the market.

In this complex business environment, companies are faced with more challenges as they strive to succeed. Among these challenges, fraud has become more prevalent and thus a more expensive part of doing business. Being able to adequately address fraud can help a business succeed in these perilous times. If done correctly, advanced analytics provides the platform for effectively detecting fraudulent transactions and stopping them before they cost the business its future viability. ■

Drone Safety Calls for Complex Event Processing, Operational Intelligence

Linda L. Briggs writes about technology in corporate, education, and government markets. She is based in San Diego. lbriggs@lindabriggs.com

Linda L. Briggs

Abstract

Like it or not, drones are slowly being integrated into the national airspace. Although they raise privacy concerns, they also offer unparalleled opportunities for data collection in public spaces. Controlling them requires sophisticated real-time data capabilities.

In just a few years, the growing popularity of the unmanned aircraft system (UAS), or drone, has created a need for complex data handling abilities to safely coordinate and control the devices from the ground.

As unmanned vehicles (especially larger ones planned for commercial purposes), drones must track and process vast reams of data about their surroundings, including location, weather, wind, altitude, battery life, and of course additional objects in the air such as commercial planes and other manned and unmanned air traffic.

As the Federal Aviation Administration moves cautiously to open the airways more fully to drones—in 2012, Congress passed the FAA Modernization and Reform Act, which provides for the integration of civil unmanned aircraft systems into the national airspace system—it opens vast opportunities. However, there are still limitations on the regulation side—the devices must be kept within the line of sight of the operator, for example, and cannot exceed 55 pounds.

Operational Intelligence Guides Drone Use

As drone traffic grows, so do safety concerns. Ensuring that drones can be safely integrated into the national

airspace requires not only regulations, but technical solutions that can incorporate data from multiple sources to present the operator with a complete picture.

Given that, plenty of companies see rich opportunities in drones.

One of them is Simulyze, which began in 2000 with a focus on providing an operational intelligence (OI) platform to handle real-time analytics and visualization for the Department of Defense and intelligence community. As the use of drones grows, CEO Kevin Gallagher now sees a huge opportunity for his platform.

"Operational intelligence has proven to be key to commercial drone adoption and is positively impacting drone safety," Gallagher says. "The technology has already been tested and proven in real-world, noncivilian applications—and those uses are showing that OI can hold the key to more widespread use of drones for commercial purposes."

Operational intelligence, big data, and data analytics, including real-time alerts and complex event processing, Gallagher says, are critical for moving beyond visual line of sight. "That's where the industry will get the most bang for the buck. ... We think our operational intelligence technology is ready to go and can help accelerate things."

Simulyze Platform Combines Disparate Data Streams

The Simulyze operational intelligence platform, called Mission Insight, processes a range of data collected both by a drone in flight and from other data sources. Mission Insight is built on Simulyze's proprietary OI platform and processes and analyzes large streams of data from disparate sources to provide UAS operators with a common operating picture in a customized graphical interface.

One crucial data set, for example, is telemetry from the vehicle itself, as the drone constantly reports its aerial location back to the ground station or hand control that is designed to process that data. "You can then put the UAS on a map," Gallagher says. "You can see where it is in context to everything else. Where do you want to fly? Are there obstacles you need to be aware of? What's the weather, especially with a short flight time? Will there be a tailwind or headwind, because that can impact whether the drone has enough power to return."

Data might also include information from the FAA on air traffic, local aircraft feeds, weather data from a local weather station, maritime data if the flight is over water or near a ferry terminal, and general aviation radio traffic. During the flight, the drone can report where it is, combine that with the real-time weather reports and information about other aircraft in the area, both manned and unmanned, and alert the operator to other vehicles or off-limits areas such as commercial airfields.

The challenge of working with the variety of data that comes in during a drone flight is "our bread and butter," Gallagher said, and the platform has been designed to collect large amounts of data quickly and handle big data problems. Data streaming into the Simulyze platform varies in format, quality, timing, and how it is collected.

"We have a lot of analytics processing going on to try to normalize it, synchronize it, correlate it, and fuse it together," he says. "This is traditionally thought of as a big data analytics problem and it really is. It's just a little different from the more traditional business intelligence side of things." The collected data is both processed in real time and displayed on a dashboard for the operator, and it's saved for learning purposes later or in case of an accident.

Test Flights for Disaster Relief

In June, Simulyze used its Mission Insight application in a test flight on the New Jersey coastline in which medical supplies were delivered from ship to shore as a first-time experiment in using drones for that purpose.

An augmented dashboard displayed real-time information including location of the drone, velocity, and the artificial horizon center, along with extra displays showing details such as the current draw from the battery and the individual rotor speeds of the device's six-rotor propeller. Gallagher says that his software also sent data to an FAA

technical center nearby and provided geospatial augmentation to the ground stations "so that operators could see where they were in context to everything else."

The flight was part of a "Drones In Disaster" event convened by a nonprofit disaster preparedness group called Field Innovation Team. Agencies from the United Nations, the American Red Cross, and leading academic and drone technology companies participated in the hopes of educating the private sector about the potential of drones as disaster relief tools—and to conduct aeronautical research about integrating drones into the national airspace system.

The drone was provided by Flirtey, an independent drone delivery start-up that has flown the first FAA-sanctioned drone deliveries. In the most recent case, Flirtey sent a drone on a preprogrammed half-mile flight to an empty house in Hawthorne, Nevada, dropping a box containing food, water, and a first-aid kit with no human intervention.

Users at Many Levels of the Flight

Users of the Mission Insight software can come from many tiers of a flight operation, Gallagher said, ranging from the drone operator to those managing the fleet operations centers to the air boss responsible for flight safety and staying within FAA regulations. The software can also be used as a workflow tool to submit flight plans and transmit aircraft position data.

Beyond Amazon's statement last year that it would eventually use drones for package delivery, other industry applications include utility inspections, real estate surveys, aerial imaging, precision agriculture, and firefighting. One example of a useful drone application in real estate could be flying a drone over a building or land for sale, capturing pictures for potential clients that would otherwise be unavailable or very costly.

Checking a remote utility line or pipeline for problems is another example, Gallagher says—a drone can be deployed relatively inexpensively to fly along an entire pipeline. "There are so many different applications right now," he says. "It's really accelerating. As more regulations [are put in place] and we're able to do more, we'll see a lot more applications."

Drones and the Future

For now, operators must keep drones under 55 pounds and within visual line of sight, but Gallagher sees that changing eventually.

He points out two areas that need to advance in order to open up the use of drones. The first is technology, including operational intelligence, "which we have pretty well in hand." Other technologies include detect-and-avoid systems and vehicle-to-vehicle communication; he concedes that there is work to be done in those areas.

"The longer path is the regulatory one," Gallagher says. "It's not just regulations, but making sure you can fly safely individually within the national airspace, including with other drones and with manned aircraft."

The real promise of drones, Gallagher says, will come as the FAA slowly issues its next tier of regulations—eventually opening drone traffic beyond the operator's line of sight. "If I'm doing a utility inspection of a power line or a pipeline, there are advantages already to using a drone," Gallagher said. "To really be able to take advantage, however, I need to be able to fly beyond line of sight."

Within that example, Gallagher said, lies the challenge of managing drone traffic. "That's where big data truly comes into play. With everything within visual line of sight, [the data] is certainly a big help, and it supports safe flight and integration, but when we go beyond line of sight, I am really relying on my data sources and my data. It has to be there and it has to be good." ■

Outthink Cognitive Hype: Creating a Business-Driven Cognitive Strategy

Steve Williams is president of DecisionPath Consulting and specializes in BI strategy and program planning.
steve.williams@decisionpath.com

By Steve Williams

Abstract

The advent of the big data era in decision support heralded the ability to store massive amounts of unstructured digital content cheaply. Although the term *big data* predominated, what was mostly being talked about was *digital content*—news articles, social media posts, images, audio clips, research reports, and so on. Such digital content can be stored in computer systems, but it is not data in the traditional business sense. Rather than the facts, qualifiers, and derived values that have been the primary inputs for previous generations of business intelligence, analytics, and decision support applications, it is *unstructured* data.

If we think of the big data era as delivering the technical tools for storing, retrieving, and processing massive amounts of unstructured data, the cognitive era is when we began to consume that unstructured data and convert it into data that can be used for decision support and other organizational purposes. As with most technology-driven innovations, however, there is a fair amount of hype to sort through before organizations make the leap into the world of cognitive computing.

This article seeks to bring a business perspective to the task of formulating a successful strategy for next-generation decision support systems (NextGen DSS). Such a strategy balances the needs to capitalize on cognitive technologies, use traditional BI and analytics technologies as appropriate, and avoid the business risks of investing too heavily or too soon in cognitive computing technologies not yet fully mature or widely adopted.

What Is Cognitive Computing?

Cognitive computing is a broad sales and marketing term. As it is being used by vendors and analysts, it includes two main types of processing: traditional analytical methods that are

applied to structured data and analytical methods that have historically been called *artificial intelligence* (AI). In academic and other research settings, AI is an umbrella term that encompasses techniques such as natural language processing, speech synthesis, expert systems, and machine learning.

We use the term *cognitive computing* here because of its prevalence among vendors and analysts. However, when evaluating cognitive computing offerings, it is important to determine whether the offering is traditional BI and analytics based on the use of structured data or more advanced analytics and decision support based on AI methods.

At the applied level, what is new and most promising about cognitive computing is the ability to leverage massive computer processing power, cheap storage for large amounts of unstructured and semistructured data, and advances in computer science techniques. For example:

- Speech recognition and language processing technologies are enabling advances in a computer's ability to engage with people using normal speech. Potential business applications include customer service, technical support, and consumer market research.

- AI techniques enable subject matter experts to teach computers by feeding in large numbers of questions and answers that collectively capture a relevant domain of knowledge. Once taught, computers can serve as highly capable assistants to people such as technicians diagnosing a mechanical problem or customer service representatives helping consumers select the best product or service for their situation.

- Computer image recognition technologies and AI techniques allow experts to teach a computer to analyze, characterize, and index an image as well as make inferences about its content and characteristics.

- Document analysis technologies allow documents to be analyzed, characterized, and indexed for potential future use. For example, completed business forms might be analyzed to identify the number of errors and to develop a probability distribution for which parts of the form contain what percentage of the overall errors.

Put into historical context, the cognitive computing movement seeks to commercially exploit computing capabilities that academic and industry researchers have explored for decades. From a BI, analytics, and decision support perspective, the key question is how to leverage cognitive computing techniques in a way that delivers ROI.

The Strategic Context for Cognitive Computing

From a strategic perspective, vendor proselytizing of "cognitive business" is a recent development. Although there are a number of prominent companies that offer cognitive computing capabilities and services, IBM (the bellwether in the IT industry) offers a good example of the current state of cognitive business offerings.

In 2014, facing persistent revenue declines in its traditional core hardware, software, and consulting businesses, IBM declared a strategic imperative around big data and analytics and articulated a strategy that included cognitive solutions. As part of its transformation plan, the company formally changed its segment reporting structure for filings with the Securities and Exchange Commission (SEC). It formed a Cognitive Solutions and Industry Services segment effective for its fiscal year 2016.

In investor briefings, Cognitive Solutions reported revenue of $17.8 billion in 2015. That reported revenue was largely thanks to $17.1 billion in revenue delivered by the Global Business Services unit, now part of Cognitive Solutions. The new segment also rolls in revenue from IBM's Software segment, which includes its Solutions Software and Transaction Software products.

As a key marketing strategy for Cognitive Solutions, IBM is promoting its Watson platform for software-as-a-service analytics, cloud-based platform-as-a-service, and industry-focused cognitive solutions yet to be developed. That said, it is difficult to determine the actual revenue attributable to its cognitive computing offerings. In other words, how much is old revenue and how much is new revenue from cognitive computing offerings?

That IBM, a leader in the cognitive movement, is still in the early stages of developing a strategy supports the argument that the use of cognitive computing techniques

has not yet crossed from early adopters to mainstream acceptance. This is important to understand when formulating a business-driven decision support strategy that incorporates the use of cognitive computing techniques. It indicates both technology risk and business risk, and augurs that a bottom-up exploratory strategy in advance of any major financial commitments is the wisest course.

The Enterprise Context for Cognitive Computing

Each company's situation is different in its particulars, but a generalized view is that a cognitive strategy needs to take into account the current state of decision support at a host organization. With this in mind, we can use a framework such as the one shown in Figure 1 to guide the formulation of a next-generation, cognitive-enabled decision support strategy.

Using the numbered circles to guide our discussion, we can use this framework to align a cognitive strategy with an existing traditional decision support environment.

Numbers one through three depict a generalized value chain where traditional structured data is processed and made available to business users for decision support purposes, typically in the form of BI, analytics, and decision support models and simulations.

1 Most companies already leverage traditional structured business data from their financial, operational, management information, Internet, and mobile systems by extracting data from those systems and moving it downstream for decision support.

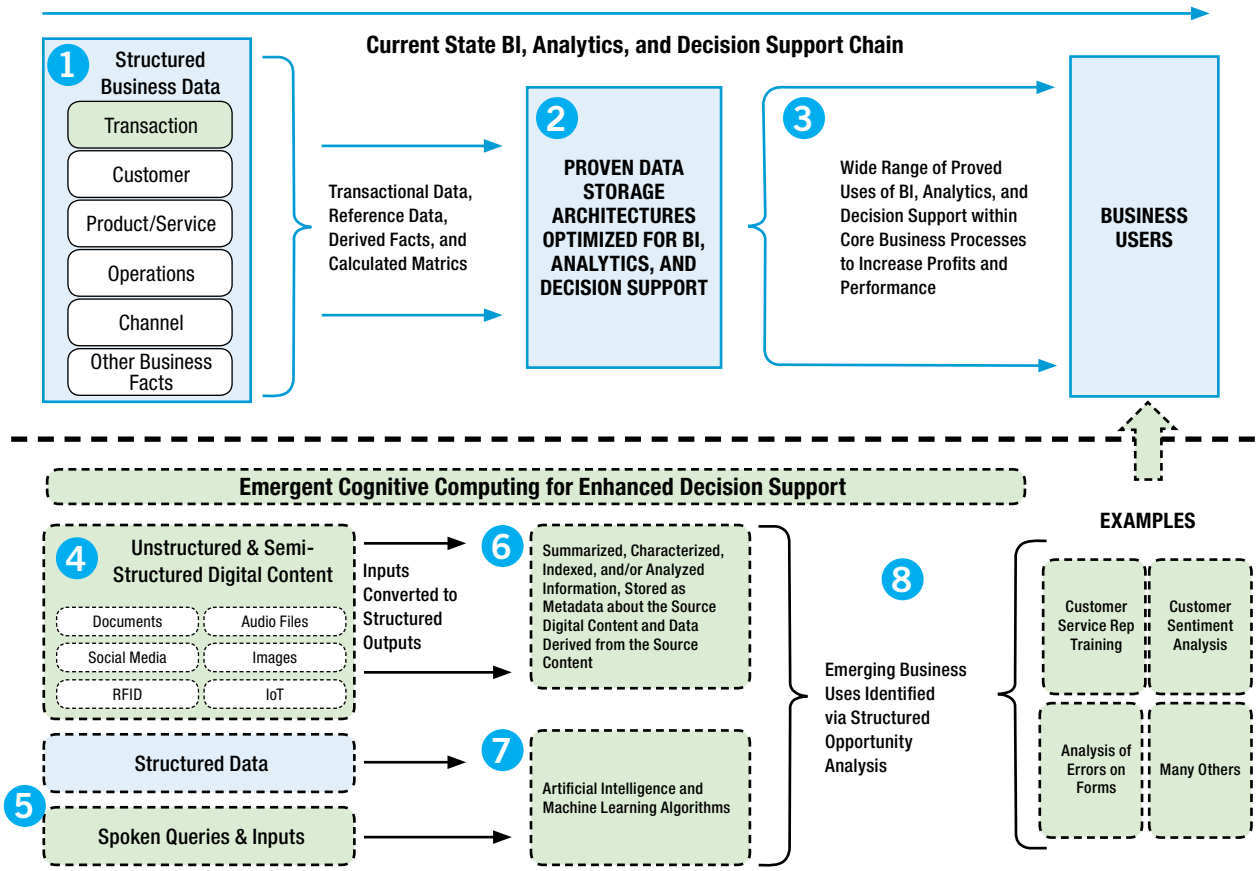


Figure 1: Framework for a business-driven decision support strategy in 2010.

2 Transactional data, reference data, derived data, and metrics are stored in various data stores using proven architectures, such as operational data stores, data warehouses, and data marts using relational or newer types of databases (e.g., columnar or in-memory).

3 Stored data is available to business users for BI, analytics, and decision support—either as structured, standardized applications using typical BI and analytics tools or within sandboxes that permit ad hoc data exploration.

The use of structured business data for decision support is a mature practice that uses proven methods and technologies and widely available, somewhat commoditized skills. There are commonly known best practices and a variety of proven use cases aligned to most business functions within an organization.

Emerging Cognitive Computing for Decision Support

Continuing with number four, we can see within the framework how cognitive computing enhances traditional BI and analytics capabilities and creates a new parallel value chain.

4 Organizations begin to use cognitive computing techniques to leverage the vast amounts of unstructured digital content more efficiently and effectively than is possible using human analysts alone.

As a hypothetical example, a human reader might be able to review, index, catalog, and enter data from 100 paper business forms a day. The resulting data set could then be analyzed to determine some quality or performance metric, such as an error rate. Using cognitive computing techniques for the same purpose, an organization could process thousands of paper business forms a day and produce a much richer set of quality and performance metrics.

Other applications could include using speech-to-text programs to process call center recordings for analysis or using language analysis programs to analyze the tone of consumer comments left at an organization's website.

5 Cognitive computing is said to also encompass the use of structured data—both traditional structured data and newer forms, such as some types of sensor data generated by the Internet of Things (IoT). However, even if some types of structured data are new, the methods of processing and analyzing such data are not—though there can be challenges associated with the sheer volume of data points.

For example, electric utilities used to read meters once a month or so. With smart meters connected to the Internet (an example of IoT in action), there can be millions of meter readings per day, so more compute power and storage is needed to analyze such data for decision support purposes.

Cognitive computing also encompasses a computer's ability to interact with humans using spoken language. Siri on Apple iPhones or Cortana on Windows devices are great examples. This ability has great potential for applications such as customer service and technical support.

6 Cognitive computing methods that take unstructured digital content (such as those in number four) convert such inputs into structured data that can be stored, modified, and analyzed for decision support uses.

For example, cognitive computing techniques can parse a document and store facts about it such as the name of the author, the presence of one or more concepts, the types of entities contained within it, the types of emotions conveyed, keywords, and the language in which the article is written. Having a database with the metadata and other facts derived from millions of documents would enable a computer to answer both traditional queries and to apply AI techniques to synthesize and summarize information from a large universe of articles.

In a business setting, consumer marketers might wish to analyze social media commentary about their products using cognitive computing capabilities to

learn how their products are perceived and measure advertising effectiveness.

7 Central to cognitive computing are techniques that fall under such general headings as rule-based systems, expert systems, AI, and machine learning.

Essentially, the newer cognitive computing capabilities we've been discussing fall under these general headings. One example of AI is the so-called rule-based system, one which humans have taught to process problems and questions and provide expert diagnoses and recommendations.

Suppose a commercial pump manufacturer has a customer support function staffed by mechanical engineers. If the company were to develop an expert system that could help diagnose pump problems, it could potentially reduce the time the mechanical engineers spend on calls by speeding diagnosis. It might also allow the redeployment of some mechanical engineers for other business purposes, such as product development or presales support.

More broadly, AI and machine learning capabilities can take inputs from numbers four, five, and six and process them via complex algorithms and heuristics to support and enhance organizational decision making.

8 Cognitive computing techniques and capabilities described have the potential for enhancing decision support by:

- Analyzing a richer set of alternatives
- Speeding the decision-making process
- Analyzing a broader set of relevant unstructured data
- Analyzing relevant data that would not otherwise be accessible to humans
- Embedding the use of unstructured data for decisions within operating processes

Factors to Consider

Whether your company's use of structured data is advanced or not, the current enterprise context establishes a starting point for formulating a strategy for next-generation decision support systems (NextGen DSS). Some factors to consider include:

- The strategic importance of traditional BI, analytics, and decision support for your company or industry
- The stage of development of your company's current decision-support value chain, as well as any projects in the pipeline for enhancing the value chain
- How cognitive computing techniques might be applied for decision support within the your company's core business processes
- How much cognitive computing projects would augment current applications, replace projects currently in the pipeline, or deliver new capabilities
- How best to organize and manage investments in cognitive computing
- Where to draw the line between operational uses of cognitive computing and decision support uses (e.g., deeming the processing of paper invoices for accounts payable management outside the scope of a NextGen DSS strategy)
- The degree to which your company culture values fact-based decision making and the use of decision support techniques
- What architecture to deploy for cognitive computing and how to align it with the current technical architecture
- Whether or not to use a business-driven or discovery-based approach to unstructured data via cognitive computing techniques

By formulating a NextGen DSS strategy that addresses the above factors and related considerations, your

company will be better able to get behind the hype around cognitive computing. A key part of formulating such a strategy is understanding the specific opportunities your company has for leveraging cognitive-enabled decision support.

Opportunity Analysis for Cognitive Computing

There are various business-driven methods for identifying opportunities for traditional decision support to improve revenue growth, productivity, and profitability. These same methods can be augmented and extended to identify opportunities for cognitive computing. Specifically, we can systematically evaluate whether and how unstructured data might be used within an organization's core business processes.

We emphasize unstructured data here because it offers an opportunity to leverage cognitive computing techniques.

The specific example analyses shown in Figure 2 are for a retail company and a manufacturing company, but the same method can be applied to any organization or sector of the economy.

To illustrate the thinking, let's look at the marketing function of CPG Retail. Within the marketing function, the example shows four typical business processes: loyalty, category management, merchandising, and returns.

To create this analysis, we would have met with the business owners of these processes and discussed with them whether they know of or can envision opportunities to leverage specific types of unstructured data (shown in the column headers). As you can see from the chart, based on these discussions, we have determined that Web logs might be useful to one or more of the marketing business functions, whereas video clips are not.

More broadly, this same line of analysis can be applied across all functions and core processes of an organization. For example, Figure 2 also shows that CPG Manufacturing has seven opportunities to leverage unstructured data, as indicated by the stars in the appropriate cells of the table.

These seven opportunities can be thought of as a "cognitive portfolio"—the basis for developing specific business

INDUSTRY & POTENTIAL USAGE OF UNSTRUCTURED DATA FOR BI, ANALYTICS, AND DECISION SUPPORT			TYPES OF UNSTRUCTURED DATA					
LEGEND								
★ = good potential fit	= low potential fit	? = possible fit	web-logs	images	audio files	video files	text	social
CPG-RETAIL	BUSINESS PROCESSES							
Marketing	Loyalty, Category Mgmt, Merchandising, Returns		★	?	?		★	★
Sales	Circulars, Co-op Ads, Promos, Pricing		★	?			★	★
Supply Chain	Purchasing, DSD, DC Ops/Inventory, Distribution							
Store Operations	Revenue, Shrink, Inventory, Labor, Maintenance							
Department Mgmt	Demand Management, Ordering, Labor							
Financial Mgmt	Performance Management, Profit Management							
Human Resources	Skill Base, Training, Compensation		?			★		
CPG-MANUFACTURING	BUSINESS PROCESSES							
Marketing	Brand Mgmt, Product Mgmt, Category Mgmt		★	?	?		★	★
Sales	Trade Mgmt, Broker Mgmt, Forecasting		★	?			★	★
Supply Chain	Purchasing, Inbound/Outbound Logistics, DC Ops/Inventory							
Operations	Manufacturing, Safety, Quality, Yield, Maintenance							
Financial Mgmt	Performance Management, Profit Management							
Human Resources	Skill Base, Training, Compensation		?			★		

Figure 2: Example opportunity analysis.

Source: Business Intelligence Strategy and Big Data Analytics: A General Management Perspective.

cases for each opportunity. That process is one of drilling down within a given opportunity to specify how the data would enable applications that would improve process efficiency and/or efficacy and thereby improve revenue growth and/or reduce costs.

The portfolio of cognitive opportunities is a key input to a NextGen DSS strategy. It provides a comprehensive view of what an organization intends to accomplish by investing in cognitive computing techniques and how it intends to achieve a return on that investment.

Adapting for the future and avoiding the risks of disruptive innovation is a basic management challenge for any organization.

It is also important to identify ways to use AI and machine learning algorithms to analyze a broader set of relevant unstructured data. Such algorithms can also analyze data that would not otherwise be accessible to humans, provide richer sets of alternatives, speed the decision-making process when needed, and automate complex recurring analyses. One such opportunity is to apply AI to managerial and analyst work. In effect, there will be many opportunities to do what ERP systems did for business operations: standardize, automate, and interconnect purported best practices for given decision processes.

For example, inventory managers and analysts in manufacturing, distribution, and retail companies have to optimize inventory levels and locations in order to meet stated service levels. This requires continual monitoring and dynamic adjustments using both structured and unstructured data along with BI, analytics, and decision support applications of varying degrees of sophistication.

Within a NextGen DSS strategy, a company could use a combination of traditional BI and analytics, business process management software, an inventory expert system, and a demand-forecasting expert system to substantially enhance the efficiency and effectiveness of its inventory management function. This use of cognitive computing, when applied to a specific industry or business function, illustrates the kind of cognitive solutions that IT vendors will soon be bringing to market.

Becoming a Cognitive Organization

Adapting for the future and avoiding the risks of disruptive innovation is a basic management challenge for any organization. Cognitive computing has already begun to disrupt the use of information for analysis and decision support, and developing a strategy that incorporates it is key to making sure your company isn't a victim. Such a strategy would blend a cognitive strategy with a strategy for traditional BI to move forward without overinvesting or taking on undue risk.

At a high level, a NextGen DSS strategy would take into account the fact that the economic performance of the organization is already being captured by its enterprise business systems and stored as structured data. These applications have largely worked well and therefore probably do not need to be replaced.

Companies should also exercise considerable care in developing a cognitive strategy by using a business-driven opportunity analysis to identify and prioritize opportunities *before* investigating specific cognitive solutions.

Vendors are investing enormous resources to promote cognitive computing in an effort to shore up declining revenues in traditional segments of the IT industry. By understanding the nature of cognitive computing and the ways it applies—or doesn't apply—to their situation, you can reduce the potential risk of outsourcing your ability to think through decisions about certain core processes.

An effective strategy will also be focused on *augmenting* the existing decision-support value chain, not replacing or replicating it. The best uses of cognitive computing are to extract meaning from the massive amount of unstructured big data and to support exploratory ad hoc analysis of any kind of data using schema-on-read approaches.

The history of computer science and IT is one of continual innovation. Often the technical innovation is far ahead of business adoption. Accordingly, a balanced perspective is advisable—one that takes into account the major workflows associated with the execution of enterprise or business unit BI programs. The NextGen DSS strategy is thus a guide for developing a cognitive organization, a process that is further developed through these six workflows.

1. Strategy, Organization, and Management

The purpose of this workflow is to determine, specify, and document:

- How the organization intends to leverage traditional and cognitive DSS methods and technologies to improve profits
- Its current state of decision support maturity and readiness
- How it intends to insert newer technologies into its existing technical architecture
- How it will organize and manage the effort to mitigate risks and ensure success

One can think of this portion of the strategy as establishing a vision for decision support and aligning the organization to realize that vision.

2. Iterative Development of DSS Applications

This workflow defines the technical activities required to develop and deploy production-ready DSS applications. There are many proven methods for executing traditional decision support projects and using agile methods specific to the unique needs of such projects is recommended.

Nevertheless, the methods and tools for developing and deploying cognitive-enabled DSS applications are still emerging (use of the Hadoop ecosystem and various platform-as-a-service offerings for cognitive computing are common). One can think of this portion of the strategy as defining how the technical work will be managed and executed.

3. Business Process Improvement

This workflow encompasses the critical activities for integrating DSS applications into core business processes. Having earlier determined how the organization intends to use traditional and cognitive DSS methods to improve profits or efficiency, there must be suitable follow-through to ensure that resulting applications work and are used as intended.

The strategy would typically prescribe a pragmatic and systematic approach to injecting DSS applications into targeted business processes, often adopting Six Sigma or similarly proven process improvement methods. The possibilities range from embedding analytics into operational processes to using business process management software in conjunction with DSS applications, in order to automate and structure enterprise decision making in business areas such as performance management and process improvement.

4. Technical Infrastructure and Operations

This workflow addresses how emerging cognitive computing technologies will be aligned with the organization's information strategy and integrated with the computing infrastructure.

It addresses considerations such as what will be done on premises and what will be done in the cloud, how vendors and platforms will be selected, how new technical processing will be managed, and other aspects of modern data center operations. The strategy also systematically addresses the relationship between existing DSS technologies and the newer big data and cognitive computing technologies.

5. Data Governance

This workflow focuses particularly on how to manage unstructured data for production applications. It takes into account the distinction between exploratory R&D-like activities of data scientists working in sandbox environments and the more rigorous data governance requirements of production applications in core business processes.

The strategy also addresses the relationship between ongoing data governance for structured data and approaches for governing unstructured data, much of

which is generated outside the organization. Since the whole point of DSS applications is to improve enterprise performance, it is important that the strategy takes a business-driven approach to data governance versus approaching data governance as an end in itself.

6. Change Management

This workflow focuses on gauging the magnitude of business and cultural change involved in implementing cognitive-enabled DSS and then managing that change.

Different organizations have different cultures with respect to decision support systems in general. Some have embraced traditional DSS and therefore may be more willing to embrace a cognitive-enabled version. Others have been late adopters of traditional DSS and might not be ready for a big leap into cognitive-enabled DSS. Changing an entire organizational culture of fact-based DSS is a major task, whether the facts in question come from traditional structured data from unstructured data or are processed via AI techniques and output as decision recommendations.

On the other hand, a small R&D-based approach to cognitive computing for DSS might not require a lot of change. The strategy would assess the degree of culture and process change needed using proven change management methods, such as stakeholder analysis or business process analysis. Then it would prescribe an approach accordingly.

Given the state of cognitive computing, the evolution of organizations into cognitive organizations is likely to take at least a decade. Having worked with large successful clients in a wide range of industries, I would say that most are still working to better-leverage structured data for DSS purposes. Many will have solid opportunities for using cognitive-enabled DSS applications, but others may be well served to adopt a smaller-scale provisional approach.

The People Impact of Cognitive-Enabled DSS

Since the Industrial Revolution, machines have increased productivity by replacing human labor. The people impacted were often frontline blue- and gray-collar workers. Some could adapt and others could not. Should cognitive-enabled DSS applications come to deliver the capabilities and pervasiveness being forecast, the people replaced this time might be managers and business analysts.

For example, the ability to analyze unstructured data might offer market research departments new ways to conduct focus groups and influence customer behavior. Alternately, expert systems and machine learning techniques might offer operations groups the ability to more dynamically forecast demand and adjust inventory targets or production plans.

It is much too early to tell if cognitive-enabled DSS will live up to the hype. Even if the technical potential is there, it may not be fully realized due to human barriers to adoption. We tend not to trust things we don't understand or that threaten our jobs. Full adoption of cognitive-enabled DSS will have to overcome both of those hurdles to fulfill its true promise. ■

References

Rometty, G. [2016]. "IBM Investor Briefing 2016." http://www.ibm.com/investor/events/investor0216.html

Watson, H., and M. McGivern [2016]. "Getting Started with Business-Driven Data Governance," *Business Intelligence Journal,* Vol. 21, No. 1, pp. 4–7.

Williams, Steve [2016]. *Business Intelligence Strategy and Big Data Analytics: A General Management Perspective*, Morgan Kaufmann.

BI Experts' Perspective

Avoiding Storms in Your Move to the Cloud

Ravi Chandran is the chief technology officer of XtremeData, Inc.
ravi.chandran@xtremedata.com

Norman C. Nicholl is principal consultant specializing in enterprise and cloud architectures.
norm.nicholl@teradata.com

Tracy Ring is specialist leader in the technology alliances group for Deloitte Consulting LLP.
tring@deloitte.com

By Ravi Chandran, Norman C. Nicholl, and Tracy Ring

Stacey Bostian is the BI director at Superior Lighting, a leading manufacturer and seller of lighting products for home, office, and outdoors. Stacey has been with Superior Lighting for 10 years and was responsible for creating the 10-terabyte data warehouse, and selecting and implementing a leading BI tool to generate reports, dashboards, and mobile applications.

Stacey's team has dabbled with BI in the cloud, using her BI vendor's cloud-based platform to develop several applications. Once successfully developed, the applications were moved on premises, where everything else currently resides. Based on these successful SaaS experiences and the belief that the cloud is the future, Stacey is interested in exploring what, how, and when she should start moving things to the cloud. There are currently discussions within Superior Lighting to move other data and applications to the cloud, including its ERP system.

She does have questions that perhaps you can help answer:

- **What kinds of conversations should she have with senior management and IT about a possible move to the cloud?**
- **What are the major pros and cons of staying on premises versus moving to the cloud?**
- **How should the fact that other data and applications might be moving to the cloud influence her thinking?**
- **What should she do with her on-premises warehouse? Keep it where it is? Gradually move its data to the cloud? Take a hybrid approach with some on-premises data and the rest in the cloud? What's the best way to think about this issue?**
- **There will be questions about security. Although Stacey feels the major cloud providers have security as good as—or better than—Superior Lighting's, what is the best way to respond to security concerns?**
- **What conversations should Stacey have with her current BI vendors about a possible move to the cloud?**
- **What are the pitfalls that Stacey should know about?**

RAVI CHANDRAN

Stacey is clearly at a stage where the cloud seems an attractive option, so she is cautiously considering next steps—a stage common to many these days.

From the perspective of one who offers an analytics platform on multiple clouds, the first issue I recommend she guard against is being **locked in to a single cloud provider**. The software you use for your BI and analytics stack should give you maximum flexibility, which includes on-premises deployment as well as multiple public clouds.

As Stacey is probably well aware, there are strong competitive pressures between the major cloud providers (AWS has had more than 50 price reductions to date), and we can expect further differentiated offerings in terms of both features and price. Stacey should position herself to fully leverage these new offerings by **planning ahead for migration** from one cloud to another.

It appears Stacey has had a positive experience with the cloud so far. She has found that the cloud enables easy and rapid application development. These features—agility, low-risk, low-cost, try before you buy, and pay-for-use—are the driving factors behind cloud adoption. They are also probably the motivation for other groups within Superior Lighting (such as ERP) to move towards the cloud.

It is much easier to plan for operational systems—for example, to build out capacity in your data center for the next three years. This would be nearly impossible for analytics systems.

However, unlike operational groups, **an analytics group has much more demanding requirements**. Growth in data volume and data sources for analytics (social media, location, online ads) dwarfs that for operations (transactional data). Therefore, it is much easier to plan for operational systems—for example, to build out capacity in your data center for the next three years. This would be nearly impossible for on-premises analytics systems; they need the elasticity of the cloud.

On the flip side, **on-premises systems offer a familiar landscape of tools** (no learning curve) and may provide reduced operational costs in the long term, albeit at the cost of up-front capital expenditures.

Increasingly, though, the same on-premises tools are also being offered in the cloud, either by the same tool vendors or by the cloud providers themselves. This is certainly the case for identity and access authentication tools, such as Microsoft's Azure Active Directory and AWS' Identity and Access Management (IAM).

Similarly, there are **numerous solutions for data security available**. Cloud providers support encryption of data at rest and associated services for key management (AWS Key Management Service, Azure Key Vault). They also support the concept of isolated networks (a "virtual private cloud") and isolated machines not shareable by other cloud users. Additionally, there are numerous commercial and open source solutions available for ensuring secure network access to the cloud from a desktop. In short, access control and data security should not be stumbling blocks to Superior Lighting's cloud migration plans.

Since Stacey's analytics group has to deal with large and growing volumes of data, her top three priorities should be **storage, storage, and storage!** To most people, cloud evokes the image of on-demand computing via "virtual machines." This is partially correct, but it ignores the significant storage options that the cloud also provides.

Clouds offer several tiers of storage, each with different features and price points. Broadly, they can be classified in order of increasing cost and flexibility of access as "Archive, Object, Block" (these options are explained in more detail in the 2015 article, "DW-On-Demand: The Data Warehouse Redefined in the Cloud," *Business Intelligence Journal*, Volume 20, No 1, pp. 8–13). From

the perspective of a BI user, you should **view analytics platforms as systems with decoupled compute and storage**. Storage in the cloud offers many more features than storage on premises at dramatically lower costs.

For instance, replication across geographically separated zones is available for relatively low marginal cost. This means your data is protected and available for disaster recovery. Decoupling compute from storage enables recovery and independent elastic scaling of both storage and compute.

Another significant benefit derives from the **relative cost difference** between cloud storage and cloud compute. Storage cost is very low ($25–$50 per TB per month) in comparison to compute cost ($1–$5 per machine per hour).

With decoupled storage and compute, data can be safely maintained even when compute resources are shutdown. This avoids the per-machine-per-hour costs during idle times (nights, weekends, and holidays) and can result in cost savings of more than 60 percent.

The elasticity and ease of use that the cloud provides could also cause some headaches, though. **Without proper governance** in place, there is a danger that users in Stacey's group could rapidly propagate cloud deployments and run up significant usage bills without her knowing—until the invoice comes, of course! Stacey should implement internal processes to control and monitor cloud usage, ideally before her first real cloud project is undertaken.

NORMAN C. NICHOLL

To help Stacey take her company's business forward into the cloud and leverage her team's experience delivering on-premises BI solutions, I'm going to walk through some key cloud discussion topics.

To prepare for the discussion with senior management, Stacey needs to understand the **overall corporate cloud strategy** and which cloud vendors Superior Lighting may want to use for hosting.

Getting educated and connecting with someone who has experience working with the identified vendors will go a long way toward understanding (at a high level) how the cloud will fit with the current BI implementation and will help identify any capability gaps.

With that knowledge about the overall direction and fit of the current solution with the cloud, Stacey can have a more meaningful dialogue with senior management around cloud corporate strategy—including timing and priorities for systems and applications migrating to the cloud.

Getting educated will go a long way toward understanding how the cloud will fit with the current BI implementation.

This is also **an opportunity to educate management** on how moving the BI infrastructure to the cloud can improve the company. This includes explaining new capabilities the cloud brings for operational integration, mobile applications, increased agility, and so on. In this way, she can paint a vision for next generation BI capabilities at Superior Lighting.

Her willingness to move quickly with some portion of cloud deployment—even if it's only development and testing—without putting the business at risk will help accelerate overall corporate adoption.

The question of on-premises versus cloud remains an interesting one, though. Continuing operations on premises allows you to maintain the status quo with **reduced short-term risk** while maintaining existing development and deployment life cycle and budgeting policies. This ensures resources are focused on immediate business and technical needs.

However, here's why cloud still matters:

- On-premises operations require continued capital expenditure budgeting and paying for peak demand; in contrast, the cloud only requires you to pay for what you use, which usually comes from the operational expenditures budget.

- Enabling new capabilities on premises requires hardware and software investment, which can take time to acquire and implement; in the cloud, you can try new capabilities cheaply by instantly spinning infrastructure up and down to prove business value.

- Capacity caps due to limitations with on-premises infrastructure (such as power, cooling, network or hardware capacity, and licensing) can impede business growth; the cloud brings agility to better support the ebbs and flows of a business.

So why does moving to the cloud matter? **Data in motion takes time**, introduces analytics latency, and costs money. Therefore, having the BI infrastructure and processes close to the data provides many advantages.

Once in the cloud, the business can take advantage of more seamless integration and enhanced capabilities. In addition, sharing governance, standards, and procedures with other business units going to the cloud will help reduce the learning curve. The analytics team must be part of this process or they risk quickly becoming a "legacy" team.

> Move the current warehouse technology stack if—and *only* if—it is optimized for the cloud and can take advantage of the real benefits of moving.

What should Stacey do with her on-premises warehouse? It should eventually be migrated to the cloud—assuming the underlying technology supports it and there are no other mitigating factors. Although warehouse options in the cloud are abundant, many are still maturing.

I would suggest moving the current warehouse technology stack if—and *only* if—it is **optimized for the cloud** and can take advantage of the real benefits of moving to the cloud (e.g., flexibility, agility, scalability, cost). Alternately, she can look at new platforming options during the migration, if time and funding can support it.

Stacey should **consider starting with a hybrid approach**. With the warehouse on premises and BI in the cloud, she'll have time to analyze the broader impact on capabilities, governance policies, sourcing, ETL, and downstream users. Choosing a hybrid implementation to start could be a more agile, risk-averse approach.

Stacey's inevitable security discussions must be based on reality, not "blue sky" wishes, so engaging the corporate information security or governance team early is critical to keeping the conversation fact based and accurate. The information security team can:

- Review and validate current on-premises data and security practices for applications, infrastructure, network traffic, encryption, and so on

- Capture any new near-term security requirements, including simplifying existing processes

- Map those requirements to cloud provider capabilities and support to ensure the security provided by the cloud vendor will be equivalent or better than the current on-premises solution

What conversations should Stacey have with her current BI vendors? To begin with **pricing is different in the cloud** and understanding licensing options ahead of time will help. Questions such as "Can I bring my own license?" and "Does the pricing model support scaling up or down on an hourly basis?" can make a big difference to the financial model.

Improvements that a vendor has made to their products or services to

specifically take advantage of cloud capabilities (microservices, dynamic scaling, provisioning, automatic routing, global availability, and so on) should be another important consideration.

Transitioning to the cloud may have pitfalls, of course. It is critical for Stacey to understand the technologies and capabilities of the cloud, even if Superior Lighting's BI doesn't ultimately move to the cloud right away, because all BI will likely need to interoperate with the cloud at some point. She would also benefit from getting on board early with the company's cloud strategy so she can make a meaningful contribution and ensure her point of view and requirements are being met.

Understanding how pricing in the cloud works for her requirements is another critical component. In some cases, pricing for compute power, disk space, and software can be deceivingly low—until you calculate those costs over time for a 24-hour-a-day, 7-day-a-week fully functional production environment. Stacey may need to look at alternative cloud solutions to help keep the cost down.

Finally, **don't get too enamored with the new technologies**. There are many and they're maturing quickly, so be sure they give you the agility, features, and scalability you need when you need it, at the right price.

Don't get too enamored with the new technologies. There are many and they're maturing quickly, so be sure they give you what you need.

TRACY RING

When it comes to management discussions, Stacey should first and foremost be working with her CIO or chief technology leader. Many CIOs and CTOs have adopted a cloud-first strategy and aligning with their overall vision can only strengthen her case.

She can take this opportunity to discuss her experiences and success using the cloud. She can also explain that, while she was waiting for the cloud to mature, she was testing its capabilities and is quite pleased with the results. Based on these **successful SaaS experiences** and the belief that the cloud is the future, Stacey can build a strong case for why a move to the cloud should be a strategic consideration.

When presenting this to the larger team, she may want to help the executive team by showing **how often the cloud is used in their everyday lives**. They are most likely already using cloud services to solve routine problems without even knowing it.

Sharing personal use cases where her peers are entrusting the cloud—such as family photos, personal email accounts, group calendars, and other cloud solutions—will help Stacey build connections with the team and encourage a move to the cloud.

There are several benefits of moving to the cloud:

- Lower total cost of ownership
- New capabilities and better business value
- Alignment with future state architectures and modernization readiness
- The ability to attract top talent with an eye for innovation, differentiating IT as a strategic capability
- Focusing company resources on development and high-value activities versus hardware and maintenance

However, a move to the cloud undeniably **changes the face of corporate IT**. In some cases, it may prove challenging to individuals who feel threatened by the shifting workplace paradigm. Even so, as with previous technological shifts, Stacey should see this as a positive change and an

opportunity to communicate to the rest of Superior Lighting that the move to cloud solutions actually opens up new opportunities for the IT staff.

Stacey should also consider the consequences of being fixed to one provider. The perception of being "locked in" is a huge barrier. To mitigate this risk, Stacey should **conduct sufficient due diligence** and work together with her procurement organization or a consulting firm to help her through some of these negotiations.

Knowing that other data and applications might be moving to the cloud, Stacey will continue to see vendors using a cloud-first strategy. In most cases, once a CIO understands the value proposition of the cloud and sees the potential decrease in overall risk, he or she will open up more opportunities for IT to focus on development and other tasks instead of traditional ROI.

Concentrating her strategy on where IT vendors are spending more will ensure that Stacey's architecture stands the test of time. She shouldn't shy away from **asking about a vendor's R&D budget**—it's a leading indicator of where an organization is investing and, of course, one of the best ways to understand the direction of future architecture developments.

What should she do with her on-premises warehouse? As Stacey will soon realize, **most companies are not quite ready to bet everything on the cloud**. They want certain features of the cloud while maintaining the comfort and control of having their data on premises. This concept of a hybrid cloud is gaining more traction among IT professionals who want the benefits of the cloud without having to give up all their locally controlled data.

As Stacey will soon realize, most companies are not quite ready to bet everything on the cloud.

One way she may be able to **gently encourage change** is to include all the enhancements to the data, dashboards, and visualizations in the new platform, while keeping the old platform intact for an 18- to 24-month transition period. The potential benefit for moving to the cloud would then be to take advantage of new metrics, data sources, and tools that can better enable the business.

New platforms coupled with change management and strong communication may ease the pain of moving to the cloud. Over time, the abilities of the cloud will likely grow and evolve as new vendors establish themselves in the cloud-based infrastructure.

Security is the most talked about concern; in fact, it is currently **one of the highest barriers to cloud adoption**. Therefore, Stacey must be prepared to respond to any security concerns that her company may have regarding the move. Nonetheless, Stacey's gut feeling on security is typically in line with what we see from many clients.

In addition to understanding security concerns about the cloud, Stacey should also **begin having conversations with the executive who is managing security** of their environment (typically the CISO or controller). This will help her develop a better understanding of the DLP tools and other data security intelligence tools that are already in place internally.

Stacey may also be able to leverage what they are already using in building the new infrastructure. An important question Stacey should ask herself in this process is how she can partner with the CISO to **drive a cloud-first strategy**. This may be an opportunity for her to work with her CISO to best categorize and monitor sensitive data and understand data proliferation. It may also be an opportunity to increase their capabilities around security in the cloud rather than constraining their security capabilities on premise.

Stacey should understand their product road map and their approach to data modernization. As her architecture evolves in the cloud, there may be an opportunity for additional players. Thus it is very

important that Stacey **look at the entire platform** and not just one or two aspects of it.

Stacey should develop a clear view of new and existing vendors, and how they may come into play as new technology emerges. This is the perfect time for Stacey to **revisit her BI strategy**, which takes into consideration anticipated technology changes over the next 18 to 24 months, and will be the best way to chart a cloud future.

Although we have discussed the many great opportunities evolving around the cloud, Stacey **should not have reservations if certain pieces of her architecture aren't good candidates for the cloud** because of geography or the need for a subsecond response. There are still applications that may need to sit on premises and that is okay. Even in a cloud-first strategy, on-premises applications can still be a part of the architecture.

Stacey should also prepare herself for the likelihood that many people will still want to stay on premises. Most importantly, Stacey should know that the focus on security cannot be underestimated. As mentioned earlier, it is the biggest limiting factor we see in adopting the cloud.

By nature, the cloud is a shared resource. It is a model for enabling convenient, on-demand network access to a shared pool of configurable computer resources that can be rapidly provisioned with minimal management, effort, or service provider interaction. As Stacey is negotiating performance, she needs to remember that even though her data is secured in the cloud, it is still a shared resource. Stacey should work with a procurement or consulting firm to ensure that all performance statistics are clearly stated and that she uses proven, qualified cloud providers with reliable service and support. ■

Instructions for Authors

The *Business Intelligence Journal* is a quarterly journal that focuses on all aspects of business intelligence, data warehousing, and analytics. It serves the needs of researchers and practitioners in this important field by publishing surveys of current practices, opinion pieces, conceptual frameworks, case studies that describe innovative practices or provide important insights, tutorials, technology discussions, and annotated bibliographies.

The *Journal* publishes educational articles that do not market, advertise, or promote one particular product or company.

Visit tdwi.org/journalsubmissions for the *Business Intelligence Journal*'s complete submissions guidelines, including writing requirements and editorial topics.

Submissions

For more information and complete submissions guidelines, please visit tdwi.org/journalsubmissions.

Materials should be submitted to:
Peter Considine, Managing Editor
Email: journal@tdwi.org

Upcoming Submissions Deadlines

Volume 22, Number 2
Submission deadline: January 27, 2017
Distribution: May 2017

Volume 22, Number 3
Submission deadline: May 26, 2017
Distribution: September 2017

Using Lean Methods to Advance the Business Intelligence and Analytics Organization

Timothy Sullivan is the director of business intelligence reporting at Cleveland Clinic.
sullivt@ccf.org

Eric Hixson is the data scientist for Cleveland Clinic Business Intelligence.
hixsone@ccf.org

Andrew Proctor is the senior director of business intelligence at Cleveland Clinic.
proctoa@ccf.org

Christopher Kucharik is a process improvement specialist of the Continuous Improvement department at Cleveland Clinic.
kucharc@ccf.org

Timothy Crone is the medical director of business intelligence at Cleveland Clinic.
cronet@ccf.org

By Timothy Sullivan; Eric Hixson, Ph.D.; Andrew Proctor; Christopher Kucharik; and Timothy Crone, MD

Introduction

To handle significant changes in regulatory requirements, payment structures, consumer expectations, and scientific discovery—as well as making big data business as usual— healthcare organizations are turning more and more to business intelligence (BI) teams. Structured as hybrids of business and technical resources, they are often the principal source of intuitive, accurate, and actionable information to drive decision making at all levels of the organization.

However, effective BI requires effective data management and reporting. Given the myriad internal and external pressures, it is not surprising that many organizations struggle with these tasks. Typical BI enterprise stakeholders include IT, finance, business units, and (in healthcare) clinical service lines. Without a comprehensive strategy and strong governance, these stakeholders can often drain limited resources by duplicating requests, effort, and capital expenditures.

Therefore, it is critical for organizations to have a unified vision and interdepartmental communication regarding data management and reporting priorities. Institutions that implement and adhere to such policies are better equipped to navigate the inevitable changes in healthcare delivery and compensation.

"[W]e have been able to think and act as a unit."

—Dr. George Crile,
Cleveland Clinic cofounder, 1918

BI According to the Cleveland Clinic Philosophy

Dr. Crile's philosophy guided the founding and clinical success of the Cleveland Clinic and has been equally influential in the implementation and execution of our BI strategy.

BI is an enterprise resource under the shared governance of finance, IT, and medical operations—a broad reporting portfolio that includes patient access, productivity, clinical and hospital operations, quality, safety, and patient experience. This three-pronged organizational structure effectively delivers prioritized results tightly aligned with strategic and operating initiatives across functional areas, supporting the Cleveland Clinic's growing data-driven culture and increasing appetite for analysis and reporting.

Over time, the BI department has seen significant growth in demand for services. In 2000, the department consisted of only a handful of analysts. By 2012, it had grown into a diverse group of managers, analysts, database and web developers, trainers, and project managers. The department manages an integrated data warehouse; designs and develops enterprise dashboards, data marts for self-service data exploration, and Web-based data entry and reporting tools; and supports numerous organizational initiatives with customized analytics solutions.

The Challenge

The organizational environment at the Cleveland Clinic would be familiar to most readers. Project requests differ in size, scope, and complexity and many arrive with limited notice and different customer expectations about when they will be completed. Staff members frequently work on multiple projects simultaneously, often in different capacities. Priorities are often fluid, resulting in a juggling act for both staff and governance processes.

As might also be familiar to readers, the department's success created a new set of challenges and questions. The sheer volume of requests required a different approach to effectively track and prioritize; internal coordination and customer communication were stretched and often broken. Identifying available resources for new projects became complicated.

BI's organizational role is primarily supportive and its outputs are usually not revenue producing. Consequently, ROI is most often demonstrated indirectly through organizational achievement using BI deliverables. However, we were directed by executive leadership to demonstrate direct organizational value through our work products and by lowering production costs.

As a result, we recognized the need to turn our methods on ourselves and develop the "BI of BI" to address critical questions about becoming more effective and efficient—what measures of success make the most sense, how do we quantify our organizational value, and how do we achieve these goals while addressing staff engagement and morale?

Lean Thinking

Even the most effective systems and processes have opportunities for improvement. These efforts, though difficult at times, can provide quantifiable benefit to organizations and teams that implement them. For any business "to be successful in the long term, they must be engaged in a relentless quest to make things better" (The Shingo Institute, 2016).

There are a number of effective continuous improvement methodologies. We embraced lean manufacturing. Lean principles have been applied to data warehouse architecture and software design (Lean Enterprise Institute, 2016; Wambler, 2013). We applied them to improve the BI organization itself.

The term *lean* was coined in the late 1980s by a research team (led by Jim Womack) to describe Toyota's manufacturing renaissance. Simply stated, it describes the

ability to do and create more with less. (Lean Enterprise Institute, 2016; Womack, Jones and Roos, 1990; Womack and Jones, 1996)

Lean's foundation is the drive to achieve a perfect target system. A "perfect system" may always remain aspirational, but continually striving for perfection pushes the people and processes to become better and more efficient. Lean involves thinking systematically about processes and embracing a scientific mindset for creating value through continuous improvement. Targeting processes or systems that align with strategic organizational goals is a critical success factor for implementing lean.

Our goal for lean BI was a sustainable program that could rapidly develop and implement high-quality, customized decision-making tools for the organization. Success would be measured in terms of departmental efficiency, waste reduction, customer experience, and organizational value.

We also recognized the opportunity to demonstrate the relevance of a methodology emphasizing standardization and efficiency in an environment where the deliverables are predominantly customized solutions.

Creating Lean Culture

The new culture we set out to create had to permeate department activities while continuing to operate within the larger organization, as well as navigate many external constraints beyond our control. We fostered initial buy-in through educating staff about the need for change, committing leadership to the change process, and providing the support structures to encourage and celebrate staff involvement.

The perception of lean can be overwhelming at first and is a major shift in mindset. Securing staff buy-in required continual communication about the reasons and goals for change. The responses ranged from fully engaged to cynical. This initial reaction was not surprising as it is a common challenge in continuous improvement initiatives (Langley et al, 2009).

In any professional environment, individuals develop technical skills and knowledge, roles and responsibilities, and the ability to navigate their corporate environment. Improvement initiatives can be perceived as unnecessary and a threat to that equilibrium. However, reluctance to buy in is often in reality risk-avoidance and uncertainty about the value of the new processes.

Leadership carried the team through this initial phase. Formal and informal leaders embraced the need for change, believed in the concepts, and incorporated lean principles and terminology in their routines. They found the initial small wins to keep the team motivated until the practices became intuitive. The best leaders realize partnerships are a necessity and collaboration with staff is essential (Lean Enterprise Institute, 2016).

Communication was also critical. As W. Edwards Deming stated, "create constancy of purpose toward improvement of product and service, with the aim to become competitive and to stay in business…." (Hunter, 2015). Leaders' communication demonstrated they were also on board, and that lean was a department-wide effort not restricted to a select few, and not a novelty being tried on the periphery.

Creating an environment that supported employees feeling free and safe to suggest improvement opportunities and different solutions was equally important. Frontline staff members know how processes are configured and often have well informed ideas for improving them. Additionally, buy-in to interventions increased when the staff being asked to change designed the change to be made.

Some initial projects felt like "one step forward, two steps back" because teams were trying to solve real problems, but using a new way of thinking about problems and new tool sets they were still learning. Leaders helped this process by insulating the team from some of the usual time pressures and the need to find an optimal solution the first time.

Finally, leaders supported progress by celebrating it regardless of size. Celebrations recognized individuals or whole teams, were formal or spontaneous, and were department-wide or confined to smaller groups. Celebrations provided a sense of accomplishment and achievement; but they were also a means of educating

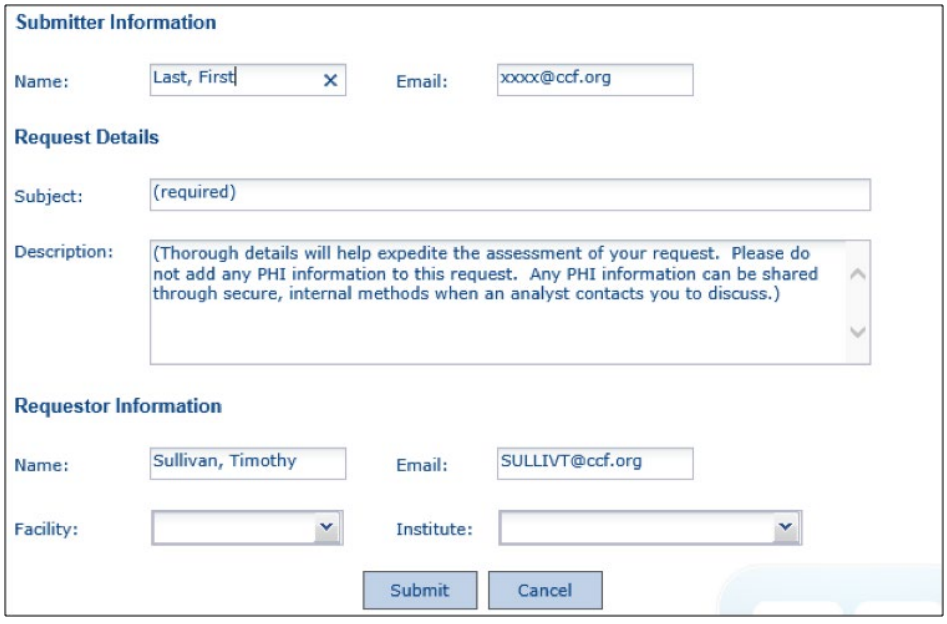


Figure 1: Online Request for Service form

the department, communicating about newly designed processes, and a powerful means of broadening enthusiasm about the lean program.

Organizational Improvements

Cleveland Clinic's lean implementation needed to be completed without reduction in the current workload and so was done incrementally. It included formal staff training, an engagement survey, developing new scales to describe department work, and applying specific lean tools and techniques to projects focused on our improvement goals.

A lean training curriculum was developed and systematically rolled out to the entire department through instructor-led classes, peer-to-peer mentoring, and standing department meetings.

Department feedback was gathered from two surveys. An organization-wide engagement survey captured general perceptions of workplace stress, manager relations, likelihood to recommend the hospital as a place to work, and compensation. A second questionnaire was administered within the BI department to gain more focused insights about those domains and garner ideas for improvement.

The survey identified five areas of employee concern:

1. Assurance that they are working on the right things and that those things are important
2. Lack of insight into customer use of—and feedback about—BI tools
3. Reassurance that improvement opportunities be focused on what is practical and achievable
4. Proper balance and focus of staff workload
5. Improving communication and transparency within the department

Our first discovery occurred very soon after the program launched. Where we thought we had a standardized intake process for new projects, we soon realized it was actually disjointed and in need of a consistent method to define and describe work requests.

Previously, we implemented an online form to facilitate submitting BI requests in a standardized format. However, we found there were actually multiple alternative

paths by which BI staff members were receiving requests and being assigned work—hallway conversations, direct phone calls, emails, and even texts to employees across BI. Staff felt accountable for these added projects, but because they bypassed the intake form, they were not counted in department totals and managers had no line of sight into them.

The online form was revised to simplify its content and improve customer experience, integrated into the reporting portal, and configured so anyone could submit a request (Figure 1).

A large, organizational education effort was made to communicate the streamlined process and the value of a standardized intake process to BI customers. BI staff members were educated in how to redirect requests made directly to them and assist customers with their submissions.

The revised intake form also interfaced with a project management system that automatically populated the request information and subsequent follow-up detail, as well as added timestamps to document request receipt, project start, comments added, and project completion.

Staff logged hours spent on work tasks, entered comments, and provided updates to communicate project status. These efforts resulted in a complete accounting of work requests.

As a result, managers were better informed about staff work queues and more empowered to help balance workloads, customers had increased confidence their requests were not lost or forgotten, and project communication improved.

Setting Priorities among Projects

Work requests are heterogeneous by nature—they vary in complexity, fluctuate in level of customization, and require different staff capabilities, skill sets, and subject matter expertise. To better prioritize and quantify workloads, we developed standardized scales to characterize work requests and allow consistent and accurate comparisons.

Each work request was assigned a weighted project score (*P* score) using a point system composed of four subscales: Request Type, Complexity, Organizational Value, and Sponsor Engagement (see Figure 2).

***P* = (Request Type × Complexity) + Organizational Value + Sponsor Engagement**

Domain	Levels	Value	Description
Request Type	Project	1.0	Estimated to require more than 40 hours of work and often multiple skill sets
	Routine process	1.0	A request, or group of requests, to manually refresh existing reports or applications that repeats monthly or quarterly
	Analyst request	0.5	Estimated to require less than 40 hours of work and often a single resource
	Maintenance/Support	1–10	Routine maintenance, patches, or bug fixes to existing production products
Complexity	Simple to Complex	1–10	Rated according to the professional judgement of the lead analyst, manager, or director on the project, considers multiple factors including degree of customization, number of estimated hours to complete, project phase, new development or enhancement of an existing tool, project risk, preestablished or calculated deadline, number of components or data sources, and anticipated innovation
Organizational Value	Low to High	1–10	Rated according to the professional judgement of the lead analyst, manager, or director on the project based on the sponsor and comparison to legacy projects
Sponsor Engagement	Low to High	1–10	Rated according to the professional judgement of the lead analyst, manager, or director on the project

Figure 2: Criteria for calculating a *P* score.

***R* = Sum (Request Type × Complexity × Contribution Level)**

Domain	Levels	Value	Description
Request Type	Project	1.0	Estimated to require more than 40 hours of work and often multiple skill sets
	Routine process	1.0	A request, or group of requests, to manually refresh existing reports or applications that repeats monthly or quarterly
	Analyst request	0.5	Estimated to require less than 40 hours of work and often a single resource
	Maintenance/Support	0.1	Routine maintenance, patches, or bug fixes to existing production products
Complexity	Simple to Complex	1–10	Rated according to the professional judgement of the lead analyst, manager, or director on the project, considers multiple factors including degree of customization, number of estimated hours to complete, project phase, new development or enhancement of an existing tool, project risk, preestablished or calculated deadline, number of components or data sources, and anticipated innovation
Contribution Level (Complexity ultiple for each assigned resource)	Major	100%	A major contributor—usually the project lead for a specific role; assigned main tasks and usually logs the most hours
	Minor	66%	A minor contributor—active on the project but less than the lead
	Consultant	33%	Most-senior contributor or subject matter expert—assists the lead or other contributors with training or guidance

Figure 2: Criteria for calculating an *R* score.

Each project was also assigned a weighted resource score (*R* score) using a point system composed of Complexity and Request Type, with the addition of a measure for Contribution Level to quantify the work effort of each resource needed for a request (see Figure 3). The aggregated resource score was the sum of the calculated scores for each resource allocated to the project.

The combination of these two scores provided multiple, ongoing benefits. The diverse portfolio of work requests could now be quantified and compared in meaningful ways, as could the workload for each individual staff member. This allowed managers to view workload and determine availability for future new work requests.

By knowing the importance, complexity, estimated effort, duration, and resource requirements of new requests and those actively being worked on, decisions on priority were greatly simplified. The *P* and *R* scores, as well as the domain subscales, provided clear guidance for prioritization decisions. BI staff time is a scarce resource and so staff should be focused on complex projects with high organizational value. Work requests with low *P* scores—particularly those with low organizational value—would be prioritized lower, redirected to existing BI self-service products, or declined.

Lastly, department leadership was better positioned to effectively and objectively communicate what the BI department was working on, the queue of work in progress and projected work, the reasons some projects were selected over others, and how their efforts added value to the organization.

As BI professionals know, scoring and classification facilitate reporting. With the introduction of the standardized project scoring, the BI team was now able to apply the same visual management tools it produced for the organization to monitoring itself, as well as to apply the same level of transparency in reporting results.

Three Examples of Putting Lean Principles into Practice

The lean improvement program initially concentrated on three projects: production migrations, internal testing, and user acceptance testing. These projects were large in scope, carried high impact, and had significant improvement opportunities—in other words, their *P* and *R* scores were more than sufficient.

Each was organized using the lean A3 process for problem definition, analysis, corrective actions, action planning, and follow-up (Figure 4).

Production Migrations

Production migrations involve moving objects from environment to environment (e.g., from development to testing and testing to production). The length and timing of a migration are critical because they affect server downtime and customer use of the products.

Best practices were documented and organized into optimized, repeatable work. Migration timing was standardized to occur on days of the week where server activity was low. Checklists were developed and implemented for all migrations. Issues were logged, addressed, communicated, and recorded so they would not occur in future migrations. These efforts produced a 30 percent reduction in server downtime.

Internal Testing

Internal testing is the review, testing, and sign-off of a new product by the project team prior to releasing it to the customer for user acceptance. As part of this review, the project team verifies final data validity, tests all product functionality, and confirms that all project deliverables in scope were completed.

Process maps were used to clarify the standard workflow steps, with responsible parties, specified time allotments, and sign-off criteria. These efforts are still under assessment, but preliminary results indicate the number of

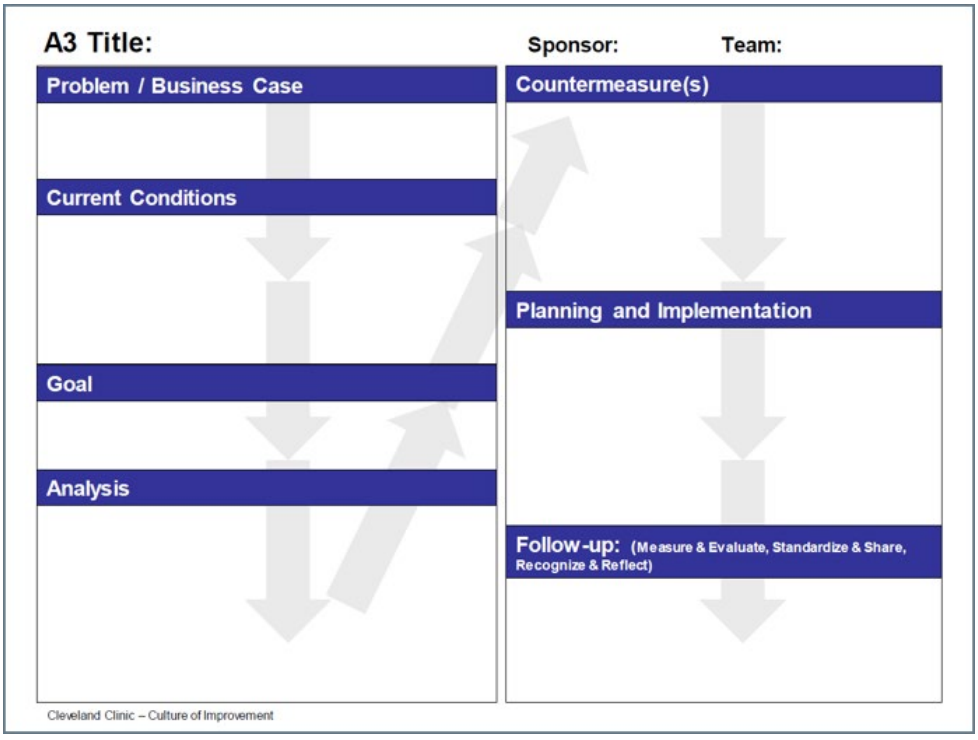


Figure 4: Lean A3 process.

product issues or errors reaching the production environment have been markedly reduced.

User Acceptance Testing

User acceptance testing (UAT) is the final customer testing prior to production migration. Project sponsors use the product and confirm it meets all their requirements for functionality and content.

The improvement team identified an opportunity to improve communication and expectations for UAT by initiating the planning at the beginning of the project rather than at the end. In this way, by the time products were ready for UAT, customer testers were already identified, had clear instructions regarding testing and communicating results or issues, and had an agreed-upon timeline for completion. Both customers and project team members report greater satisfaction with the new planning process.

Conclusion

Development of the lean program in the BI department at the Cleveland Clinic is still ongoing, but our initial experience has met our goal of establishing a sustainable program of improvement. Internal efficiency and effectiveness have improved with better customer experience and engagement. In addition, there was no significant change in the organizational environment, reduction in the volume or pace of work requests, or material change in BI's portfolio diversity throughout this process.

Our epiphany, however, was that lean did not require eliminating every form of variation by standardizing every process and product. What it did require was disciplined thinking focused on identifying the proper targets for standardization and efficiency to eliminate unnecessary variation. Lean requires significant effort, but the benefits have been worth that effort. ■

References

Deming, W. Edwards [2000]. *Out of the Crisis*, MIT Press, as quoted in Hunter [2015]. "Create Constancy of Purpose," https://blog.deming.org/2015/09/create-constancy-of-purpose/

Langley, Gerald J., Ronald D. Moen, Kevin M. Nolan, Thomas W. Nolan, Clifford L. Norman, and Lloyd P. Provost [2009]. *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance, Second Edition*, Jossey-Bass.

Lean Enterprise Institute [2016]. "What Is Lean?" http://www.lean.org/WhatsLean/

The Shingo Institute [2016]. The Shingo Model for Operational Excellence, Utah State University, p. 9.

Wambler, Scott [2013]. "Agile Data Home Page," http://www.agiledata.org

Womack, James P., Daniel T. Jones, and Daniel Roos [1990]. *The Machine That Changed the World*, Harper Collins.

Womack, James P. and Daniel T. Jones [1996]. *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*, Simon & Schuster.

Winners: TDWI Best Practices Awards 2016

TDWI's Best Practices Awards recognize organizations for developing and implementing world-class business intelligence and data warehousing solutions. Here are summaries of the winning solutions for 2016.

For more information, visit tdwi.org/bpawards.

Analytics and Data Science

Toyota Financial Services

One of the toughest questions any company can ask itself is, "Do we really know what our customers want?" With 13 million customer interactions per month, customer information dispersed across multiple applications (billing, payment, services, and marketing) as well as numerous unstructured interactions (logs, call center notes and audio, and surveys), getting to the answer can seem daunting. To begin with, we need to identify our customer experience pain points; assess the current and future customer needs; and determine which customer channels provide the best returns.

By applying machine learning and Toyota's Kaizen philosophy of improvement, TFS performed what we call a digital *genba* walk (an observation of work) by combing through millions of customer interactions per month, parsing out insights, and diagnosing root causes of problems. Both supervised and unsupervised machine learning methods were employed on over 12 TB of customer data, which enabled TFS to better understand the customer voice and map the customer journey.

The *genba* walk has allowed for effective, efficient, and automated identification of process improvements. TFS is now able to explain channel usage, identify pain points, predict future behavior, and prescribe recommended actions, resulting in significant financial savings. Machine learning has also allowed for an automated performance measurement system for planning new improvements, setting targets, checking results, and providing course corrections.

BI and Analytics on a Limited Budget

Oregon State University

Supported with existing resources and staff, the Cooperative Open Reporting Environment (CORE) system is a

21st century business intelligence solution for Oregon State University. The six participants formed the Business Intelligence Center (BIC), which successfully re-architected the data warehouse and built the infrastructure to support next-generation university BI practices.

Throughout CORE development and design, university subject matter experts provided short-term assistance to the BIC team. The engagement of data owners to create agile data governance has created a position- and role-based security access model that makes data readily available to employees—with a user-centric design for meaningful reports in formats that meet informational needs.

The evolution of CORE continues and has shifted the long-standing centralized control of university data and information. CORE will become the standard—the only source for accurate information. During this era of limited resources for public universities, additional funding for CORE is not anticipated. This small band of "misfits" will continue to engage users, broadly communicate the CORE story, and develop a reporting and analytics system that informs tuition and fees, budgets, and student metrics for years to come.

BI, Search, and Data Discovery

Microsoft

Microsoft's Finance Business Intelligence and Finance Cross Service Platforms IT teams partnered to create and deliver on a vision to provide an innovative enterprise-level Measure Management Service (M2) through a three-year initiative. Currently 18 months into that journey, Microsoft's Sales and Marketing Group (the first adopter of the M2 service) is seeing dramatic gains in data quality, timeliness, and business measure publishing support.

Other Microsoft groups are preparing to implement M2 to take advantage of the service's ability to easily define, maintain, and change their measure definitions and calculations; mesh with governance and data steward processes; and provide and manage source data from various systems. The groups that implement Microsoft's M2 will find a service that provides the timely, high fidelity, highly governed business performance measures widely used by Microsoft decision makers to measure company performance-to-target, leading indicator trends, and forecasts.

Microsoft's M2 Service initiative continues to deliver on the vision set forth 18 months ago—and by delivering a complete set of innovative services that help business leaders and decision makers understand and monitor business performance, and guide business decisions to improve that performance, M2 intends to make a big mark at Microsoft.

Big Data

Verizon Wireless

The rise of big data created unprecedented challenges in providing Verizon employees with data-driven decision-making capabilities.

The goal was to manage internal and external data with powerful and secure storage and processing solutions at a relatively low cost—while providing more employees, in both the IT and business organizations, with access to data.

We extended our analytics ecosystem to integrate big data discovery with existing analytics capabilities to give employees access to data and insights via easy-to-use tools that facilitated widespread adoption, data discovery, and analysis at scale. The result: integrated big data helps drive our Verizon Lean Six Sigma (VLSS) program, reduce churn, improve customer satisfaction, and lower business costs—leading to billions in measurable benefits.

Some specific benefits include:

- Verizon Wireless leads the industry in customer loyalty with the lowest retail postpaid churn rate among all carriers at 0.96 percent.
- Increased self-service and millions fewer calls in 2015.
- Add a Line (AAL)—big data platform analyzes behavior, propensity models predict most likely device for each customer and CRM launches multichannel AAL campaigns to target customers with best device offer. The result: a significant increase in AAL gross adds in 2015 representing millions in monthly revenue.

Data Management Strategies

VMware, Inc.

Solution Sponsor: SAP

Our program began with the realization that VMware needed to manage data as a true, reliable corporate asset. We created a vision for master data management (MDM) to invoke a hub-and-spoke, publish-and-subscribe architecture with embedded information governance and data quality measures. This was initially implemented on Oracle platforms, but we are in the process of migrating our ERP and MDM capabilities to an SAP multidomain platform. We have completed the first phase with a new vendor master using SAP Master Data Governance and continue to drive value in the legacy environment, given a 2–3 year migration plan to SAP.

The success of our data management program is due to the active support of the information governance business and IT community, which agreed to include individual, annual performance goals for information governance—the framework that drives vision into value. Moving from reactive data cleanup to proactive data management and governance requires extensive change management around business practices and processes. The technical part, however, is easy compared to the people side of change management. Constant education, along with understanding the value achieved, is key to maintaining the necessary participation.

We have discussed our program with many peer organizations that have not achieved a comparable level of success in their data management and information governance programs. We believe that VMware's success makes us pioneers in managing data as a true, reliable corporate asset.

Emerging Technologies and Methods

Arizona State University Analytics & Data Services (co-winner)

Solution Sponsor: Snowflake Computing

Arizona State University's Self-Service Analytics Initiative, launched at the beginning of 2013 and completed in the fall of 2014, has brought a revolutionary change to the way that data is accessed, manipulated, and published around the institution.

The core goal of the project was driven by an increasing demand for university data across all of ASU's departments, administrative units, and leaders. With this goal in mind, the University Analytics & Data Services Team created their vision: to provide an easy-to-use, Web-based reporting platform that facilitates efficient access to data.

To meet this challenge, the initiative developed a two-tier solution leveraging the Microsoft BI stack of tools. For power users and the analytics team, the initiative selected tabular and multidimensional modeling tools (SSAS, PowerPivot) to curate, organize, and publish data sets. For casual users and other university constituents, the team developed a custom-themed SharePoint environment designed to host analytics assets as various as Excel Web Services, SSRS, Tableau, and Splunk.

With this combination, Arizona State has created an innovative self-service environment that meets the needs of its constituency and leaves room for the addition of new tools and methods as soon as they become available.

Emerging Technologies and Methods

IAC Publishing Labs, operator of Ask.com (co-winner)

Solution Sponsor: Snowflake Computing

By completely transforming its relationship to the data upon which it depends, IAC Publishing Labs' BI team went from business barrier to business value creator. The legacy data environment simply could not keep pace with IAC Publishing Labs' rapid growth.

Cloud data warehouses represent the leading edge of analytics and BI innovation. By choosing to replace its legacy environment with the cloud-based Snowflake Elastic Data Warehouse, IAC Publishing Labs immediately jumped to the forefront of modern best practices.

Achievements include:

- Establishing one source of truth in a centralized data warehouse that serves both data scientists and business analysts and provides value across all the company's businesses

- Consolidating technologies and eliminating outmoded systems, which has given the data greater responsiveness and usefulness and lowered IT costs
- Providing enhanced BI service levels because the data environment is more consistent and responsive
- Enhancing visibility into the data analytics for business executives and marketing teams as well as data scientists and BI experts
- Changing the BI team from a cost center to a value center and making it a respected, often-consulted resource within the company
- Significantly decreasing expenses (by 78 percent) for the data warehouse environment

Real- and Right-Time Operational Intelligence

Aircel Limited, India

In today's Indian prepaid mobile market, understanding a subscriber's interests, needs, preferences, and consumption patterns in real time and providing personalized products and services across all customer touchpoints is a unique strategic advantage. With this in mind, Aircel initiated a project to capture customer transactions in near real time and to leverage that data for event-based actions, as well as to update customer profiles for use by various touchpoints, campaign systems, and the end-user community.

By virtue of having customer consumption and network usage data available in real time, Aircel is committed to recommending services and products best suited to the needs of every customer. A few examples of business value derived with this solution include:

- A campaign management system that leverages the Unified User Profile (UUP), increasing monthly revenue by more than $7 million by promoting best-fit offers to selected customer segments
- The ability for retailers to fetch best-fit recharge options for prepaid customers in real time (on USSD channel), generating more than $18 million in tertiary revenue
- Realizing synergy among network teams and sales and marketing by having a single view of network and business KPIs in near real time

Data Warehouse Innovation

- No Winner Selected

Data Visualization and Storytelling

- No Winner Selected

BI StatShots

Predictive Analytics—An Update from the Field

By Martin Pacino, Industry Research Analyst, TDWI

Past TDWI research indicates that organizations are reaching critical mass in their use of predictive analytics. The technology is now mainstream and its value is well understood. What's more, those organizations not using predictive analytics are interested in using it.

Our San Diego 2016 survey followed up on similar questions about predictive analytics use asked in 2013 to understand the current challenges users are facing and to see how the landscape has evolved. One hundred and seventy-four people responded to this survey and therefore this should be considered simply a "quick pulse" survey. However, results do map to what we've seen in other research.

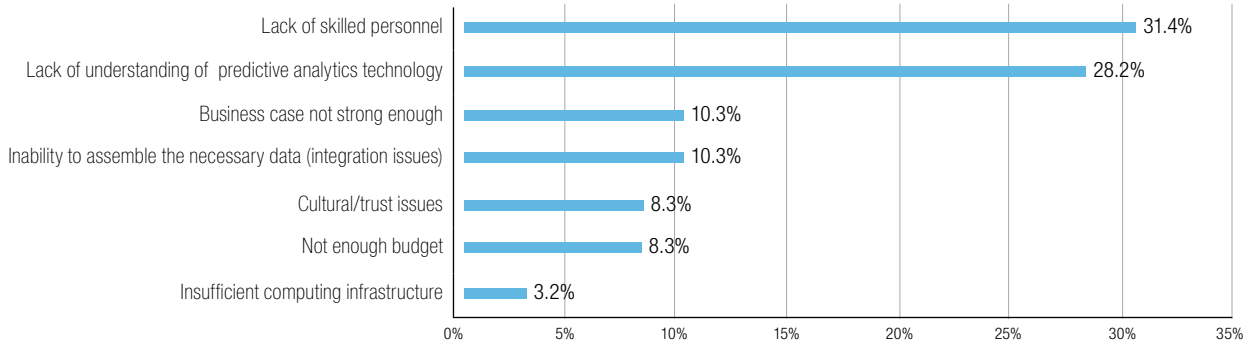
- **Despite high levels of interest, use of predictive analytics appears to have plateaued.** Thirty-nine percent report that their company currently uses predictive analytics, with a similar proportion intending to use it in the future (42 percent). This is roughly equivalent to the levels seen in July of 2013, when a combined 81 percent reported either currently using predictive analytics or the intention to.

- **Skilled personnel and usable software are pressing issues.** Thirty-one percent said that a lack of skilled personnel is a barrier to implementing predictive analytics. A similar proportion (28 percent) reported that a lack of understanding of predictive analytics software is their biggest barrier. However, as only 10 percent thought there is no business case to use predictive analytics, further education efforts and demonstrations of viable software options would be valuable.

- **Data scientists are the primary builders of predictive analytics models,** according to two thirds of respondents (64 percent). Interestingly, this reflects a near total inversion from 2013, when 64 percent stated that business analysts were developing these models. This may be owing to the growing presence of data scientists in companies, the increasing complexity in predictive analytics practices, or perhaps simply the fact that media popularization of the title is leading people to misapply it to their organization's model builders.

This combination of factors presents an opportunity for predictive analytics vendors and educators to meet would-be customers halfway. With the market near the saturation point, shrinking the gap in knowledge and resources will be critical to the ongoing use of predictive analytics—goals furthered by increased education. The results also further suggest that the recent industry turn towards simplified tools such as prepackaged tools and DIY predictive analytics packages is a reasonable response to market demands, but will need to work in conjunction with further education. ■

What is—or do you expect to be—your biggest challenge with getting your organization to use predictive analytics?



tdwi ONLINE LEARNING

The Industry's Best BI and Analytics Education

AROUND THE CLOCK, AROUND THE WORLD, ON YOUR SCHEDULE

For more than 20 years, TDWI has been dedicated to the success of BI and analytics professionals through our world conferences, Onsite courses, and our prized CBIP certification. In fact, 95 of the *Fortune* Top 100* turn to us.

Now with our new Online Learning program, there are no limits. If you can get online, you can get the education-based career boost that our members rely on. Whether you're an advanced analytics professional or a newcomer looking to get familiar with the concepts, BI and analytics courses from TDWI Online Learning will give you the skills to tackle your toughest workplace challenges today—and be ready for the challenges of tomorrow.

*based on attendee data and the top 100 companies from the 2015 Fortune 500

To learn more, visit us at online-learning.tdwi.org

TDWI Partners

These solution providers have joined TDWI as special Partners and share TDWI's strong commitment to quality and content in education and knowledge transfer for business intelligence, data warehousing, and analytics.

tdwi PARTNER

Platinum Partners

IBM

Partners

cloudera

DELL Software

Hewlett Packard Enterprise VERTICA

Information Builders

LIAISON

MAPR

MarkLogic

SAP

sas

snowflake

tableau

talend

TERADATA

WhereScape