Chapter **1**

Context: The Case and Place for Smart Data Strategy

Experience is the name everyone gives to their mistakes.
— Oscar Wilde

1.1 VALUE OF DATA TO THE ENTERPRISE

The context in which executives address performance is truly selective in that one can choose to consider performance of a function or department, of a product or asset, or even of an individual. When we talk about performance optimization it is in the enterprise context that is discussed at some length. Only the CEO and executive staff sit in a position of neutral view, the all encompassing enterprise view of performance. Performance metrics from all the pieces or elements of the organization aggregate to comprise a composite view.

Performance management expert Frank Ostroff, speaking to the Association for Enterprise Information (AFEI) on the subject of "Designing and Implementing Horizontal, Collaborative Organizations to Improve Performance," emphasized the criticality of having executives spend at least 20% of their time addressing managing change or he said "it isn't worth it."

Well, we are not talking to people who somehow make change happen through consultation and cajoling from the outside; we are talking to *you*, the executive, about what *you* can do personally to make change happen. We are not talking about a special program; we're talking about executives asking the right questions and demanding the right data to answer them. That is the strategy, plain and simple.

Smart Data: Enterprise Performance Optimization Strategy, by James A. George and James A. Rodger

The easier and faster it is for you to answer questions and make decisions, the greater your capacity to optimize enterprise performance. That is a critical element of the smart data focused strategy.

However, to optimally manage performance, the executive must be able to identify deviations and improvement opportunities by interpreting indications and drilling down to root causes. That is particularly the case for mature requirements and demands. To accomplish this you need a framework, methods, and algorithms upon which you can depend with confidence, and which we advocate be defined through the smart data strategy.

Responsibility to lead and to see ahead, to accurately anticipate the future, and to formulate vision rests with senior leadership. In both instances, managing the operation and envisioning the future require actionable data and analytical support. While ultimate decisions rest at the top, end user customers and employees are major contributors. Moving data from them in an actionable state to that of the executive is a deliberate engineering challenge. MIXX.com CEO Chris McGill says that "sifters" are needed to help cull the data, referring to application of his social media service to help accomplish this.

On April 19, 2009, the *Washington Post* headline [1] is "Obama Picks Technology and Performance Officers."

Aneesh Chopra is the nation's first Chief Technology Officer (CTO). Jeffrey Zients is the Chief Performance Officer (CPO). The Chief Information Officer is Vivek Kundra. Significant is the fact that there are three positions—Technology, Performance and Information—and that they are expected to collaborate, in our words to help optimize enterprise performance. They do this not as operations executives managing complex bureaucracies, but by providing leadership and expert direction and input to the president and to the cabinet secretaries.

Can we assume that the CPO will be a performance analyst, examining government operations to seek out improvement opportunities? Can we assume that the CTO will be evaluating and improving the technology underpinning that spans the spectrum of research beginning in a discovery mode to identify strategic technologies that will not only optimize government enterprise but will energize the economy? Can we assume that and the CIO will focus on hardware and software systems that enable all aspects of government to optimize performance, by leveraging the best information technologies? Those *are* our assumptions.

According to the Department of Labor is *Occupational Outlook Handbook 2008–2009*, "chief technology officers (CTOs), for example, evaluate the newest and most innovative technologies and determine how these can help their organizations. The chief technology officer often reports to the organization's chief information officer, manages and plans technical standards, and tends to the daily information technology issues of the firm. (Chief information officers are covered in a separate *Handbook* statement on top executives.) Because of the rapid pace of technological change, chief technology officers must constantly be on the lookout for developments that could benefit their organizations. Once a useful tool has been identified, the CTO must determine an implementation strategy and sell that strategy to management."

The last sentence is a problem because it implies all technology is a tool, and that is too narrow a definition and context. Broadly, technology is the application of science. The National Institutes of Health (NIH) defines it as "a body of knowledge used to create tools, develop skills, and extract or collect materials; the application of science (the combination of the scientific method and material) to meet an objective or solve a problem" [2].

The Department of Labor (DoL) defines the duties of chief information officers as follows. "CIOs are responsible for the overall technological direction of their organizations. They are increasingly involved in the strategic business plan of a firm as part of the executive team. To perform effectively, they also need knowledge of administrative procedures, such as budgeting, hiring, and supervision. These managers propose budgets for projects and programs and make decisions on staff training and equipment purchases. They hire and assign computer specialists, information technology workers, and support personnel to carry out specific parts of the projects. They supervise the work of these employees, review their output, and establish administrative procedures and policies. Chief information officers also provide organizations with the vision to master information technology as a competitive tool."

General descriptions such as these are helpful in understanding normal expectations, although these are extraordinary times and keen attention must be given to how these positions fit into the enterprise schema. In fact, the enterprise must be engineered to perform in pursuit of aims that are defined by the chief executive officer (CEO) and supported by the executive team. The CEO tweaks and guides staff leadership positions to accomplish priority requirements.

First and foremost, the CEO needs and should want greatly improved data and data analytical support in order to optimize enterprise performance.

Comparing our brief assumptions with the DoL descriptions, we believe that the CTO position as staffed by the government and by many commercial organizations should not be so tightly coupled with the CIO because much technology is not information technology per se. New technology can shake the paradigm completely, such as synthetic fuels that can make all current engine technology obsolete overnight.

The chief performance officer (CPO) position is a green field notion that a dedicated professional needs to be able to objectively assess departments' performance and to make recommendations about their metrics. We embrace this idea as it emphasizes objective performance analysis that is most certainly data centric.

Commercial enterprises too are staffing these positions with parallel expectations in the commercial context. Perhaps the most significant assumption, one that is shared, is that superior products and services are the expected outcomes. Nothing less than superiority wins in a global economic environment and that means optimizing the interplay between government and commerce.

While regulation is receiving a resurgence of attention, complementary and balanced attention must be focused on the government providing incentives for innovation in the pursuit of superior commercial performance such that the capital

engine is reinstated and made healthy again. Therefore the aim must be to achieve higher performance and higher standards than ever before.

These pursuits are made real with associated transparency and accountability with a data focus and with a smart data strategy such as we recommend.

For the *Washington Post*, referring to the CPO position, Max Stier wrote [3]: "When it comes to government performance, one of the best ways to improve it will be to improve the way we measure it." He added: "Right now, our federal government relies on lagging indicators to let us know how our government is doing. In other words, we need leading indicators, not lagging ones."

The CPO is a new position and after a false start at filling the post, the challenge is to grasp what President Obama wants to do with it. We have some ideas as discussed herein. Needed is a management approach to optimizing enterprise performance that is shared from top to bottom, and that is known and understood by everyone—a *transparent* management approach.

It is a nontrivial activity to change the way our enterprises work, from being backward looking to forward looking.

While we are surely data focused, we want to clarify that it takes more acumen than numbers to run an enterprise. A story by Al Ries says it all in the title of an article by him: "Metric Madness: The Answer to Mathematical Failure Seems to Be More Math, If You Run a Company by Numbers Alone, You'll Run It into the Ground" [4]. The article describes the need for people with marketing and creative skills, to be able to generate new products and services that people need and want. Smart data is needed as much by the creative professionals as it is by accountants.

The contemporary performance environment includes an internal view and an external one. The internal view is something that we will discuss in this and later chapters, which describe technology that you need to have on board, or access to, to support smart data. Having access to technology accommodates the notion of "cloud computing," that is, where your enterprise subscribes to computing and communication services on a pay-as-you-go basis versus having in-house infrastructure. Cloud computing is advancing and emerging rapidly and will surely enable smart data strategy as an extension to or as a replacement for how you manage computer and communications technology enablement.

On April 23, 2009, Reuters carried an announcement that "International Business Machines Corp. plans to launch cloud computing services this year, taking on companies such as Amazon.com Inc., Microsoft Corp., and Google Inc." [5].

You also hear terminology describing various levels of Web computing: Web 1.0–4.0. These generations of Web development are defined as follows:

- Web 1.0—First generation and we are all there: interlinking and hypertext documents that are web accessible via browsers, web pages containing text, graphics, pictures, and so on.
- Web 2.0—dmiessler.com distinguishes Web 2.0 as being a difference in how applications are implemented versus being a change in technology. Differentiators include web-based communities, hosted services (i.e., social networking sites), wikis, and collaboration.

- Web 3.0—This is Tim Berners-Lee's dream come true: the semantic web featuring web content expressed in natural language and in a form that can be read by software agents. Herein is one major characteristic of smart data.

Can you have smart data without semantic interoperability or without the World Wide Web Consortium (W3C) completing a standard to tag data to make it smart? Well, our answer is that you can have smarter data by attending to the three domains of our definition. You can make considerable progress until the tagging schema pushes us to the pinnacle of this round of achievement.

- Web 4.0—Marcus Cake (www.marcuscake.com) provides a graphic (Figure 1.1) depicting the transformation that concludes with Web 4.0 being called *transformational*, including the notion of global transparency, with community sovereignty over channels of information.

Our view is that we already have a toe into Web 4.0 as social media pushed us there as did the global economic calamity. So, for our executives, you need to be aware of your changing environment and know that a part of this is relevant to achieving our

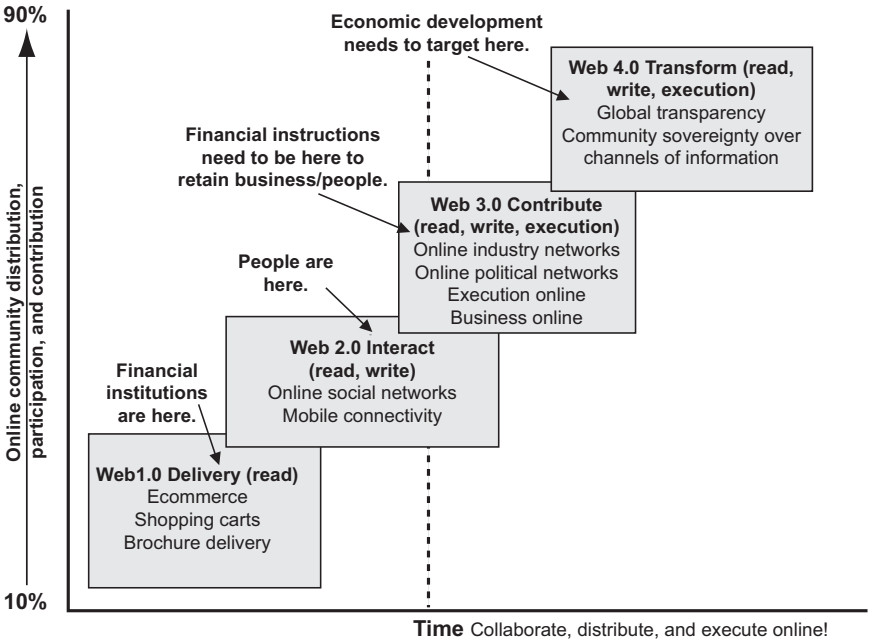

Figure 1.1 Economic development: Web 1.0 (distribute), Web 2.0 (participate), Web 3.0 (contribute), and Web 4.0 (transform). (used with permission from Marcus Cake, http://www.marcuscake.com/key-concepts/internet-evolution [6].)

highest expectations for smart data. However, there is much work to be done by you right now before being concerned about all of the technical details.

Social media are technologies and practices leveraging the Internet that enable people to share opinions, insights, experiences, and perspectives, including user-created media such as a blog, wiki, or hosting site. We propose another definition for executives: social media accelerate communications among communities of practice and interest.

How does this relate to smart data? Smart data moves on the Web and it moves among your enterprise community, sometimes leveraging applications of social media based on your strategy and active participation. The process and the medium become a generating source of data as well as a user of data.

Distributed computing has evolved from the days of remote method invocation. Today, cloud computing, the semantic web, SOA, and SOE have transformed what was once simply referred to as a three-tier client–server architecture. In our discussion of smart data, we have portrayed it in the context of the enterprise view, interoperability of technology, and a view of the data that addresses appropriate methods and algorithms.

Companies such as Google and Amazon have embraced this full-service cloud and definitely have moved the ball up the hill. However, today's executives who are operating in the shadow of a global crash must keep their eyes on the target. They are aware of their enterprise needs and are surrounded by technical support and a lot of data that may be used to optimize their systems. The answer is not to buy more legacy systems or to recapitulate the phylogeny of data from RDF and OWL. The answer is to produce business rules that direct the executive toward the right answers and outcomes.

President Obama has decided to discuss privacy issues at www.data.gov and he has empowered his CIO and federal CIOs to make government data transparent.

In a similar vein, Health and Human Services (HHS) has always operated with a socialization approach to policy development and implementation that can greatly benefit from applying social media such as www.mixx.com to accomplish this mission. Mixx.com is one example of a company that has a public face to a user community while offering customization services and branded facing through government and commercial websites. For organizations that are communications intensive in developing and advancing policy, this can have significant advantage by (1) reducing conference and meeting time while (2) increasing participation and (3) speeding the communications process.

Much of the transfer of this information will take place on what is currently termed Web 2.0. The term Web 2.0 describes the changing trends in the use of World Wide Web technology and web design that aim to enhance creativity, communications, secure information sharing, collaboration, and functionality of the Web.

The concept of cloud computing is based in distributed computing concepts such as grid computing and incorporates components of the grid such as infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) as well as Web 2.0 and other recent technology trends that have the common theme of

reliance on the Internet for satisfying the computing needs of the users. Some examples of Web 2.0 include social-network or video-sharing sites, blogs, peer to peer file sharing, and WebTV.

WebTV can be software-based, in which you access the online TV by downloading a specific program or software. WebTV can also be webpage-based, in which you can access the online TV by entering a specific website and run ActiveX. Software-based examples include Live TV and PPLive. Webpage-based TV includes www.hulu.com and www.cnn.com/live. There are advantages and disadvantages to WebTV.

Pros

- Most are free
- No location restriction
- International TV
- No hassle schedule
- Commercial-free
- More choices
- Watch movies at home

Cons

- Requires a high speed connection
- Sometimes low audio or video quality
- Risk of programs bundled with virus and malware
- Lost productivity at workplace
- Piracy
- Not a very large selection of things to watch
- Copyright and illegal links

So where is the payoff for President Obama and his CIO in making government data transparent and how does this fit our concept of smart data? Executives such as President Obama must realize that today's Internet users are deserting the mass media. Whereas previous executives used fireside chats and television, today's leaders must realize that the smart data resides on the Web. There is an old query that asks if a tree falls and no one is around to see it fall, did it really fall? Today's leaders ask if they have communicated.

Also at an AFEI luncheon, we listened to president Obama's former campaign strategist, David Plouffe, talk about the Obama campaign strategy. He made a number of points worth quoting.

Referring to the Obama campaign, Plouffe said: "We were not a democracy." By that he meant that once the executive had chosen the strategy and message, it was time for everyone to execute in synchronization and to be consistent and stable, such that grass roots could absorb, understand, and reinforce. He said: "We had to model the electorate."

Plouffe went on to say that "when we had problems it was because we did not step back to test our assumptions. We had to ask, what's the strategy and do our assumptions align properly? Decisions are based on strategy."

Optimizing enterprise performance requires knowing the strategy; floating the strategy to subordinate organizations such that they are contributing properly; and measuring contributions against the strategy.

How frequently do you change strategy? Our answer is: as often as the data indicates that you should.

Bob Bragdon, Chief Security Officer (CSO.com), commented: "Assigning monetary value to data is a concept that few businesses seem to be embracing, although, even from a cursory examination, it would seem to make a great deal of sense to do. We treat everything else in our organizations as assets with measurable financial value: inventory, people, property, etc. But when was the last time you sat down and calculated a dollar value for your customer database? My guess is that it has never been done" [7].

Our view is that the situation is worse than Bragdon describes. Executives do not associate assets with data and data with costs and value. Accounting systems do not work that way, unless you move to activity-based cost accounting, for example. In fact, the effort to introduce these practices needs resurgence and reinforcement. Accounting rules and methods constrain the ability of companies to plan ahead and require addition and change to support capacity to forecast and predict.

In 2003, one of the authors (George) was at a conference at which Michael C. Daconta talked about designing the smart data enterprise [8]. Daconta was talking about how semantic computing would impact enterprise IT. We really didn't remember this until Dr. Brand Neiemann, at the CIO's office at the Environmental Protection Agency, brought it to our attention. So, with this prompting we went back in time and found the reference. If Daconta was talking on this subject in 2003, why is it not in wide practice today? In fact, George was speaking about semantic interoperability at the same conference and is astonished at the lack of progress.

Daconta went on to become a metadata guru for the Department of Homeland Security, where we are certain he made significant contributions. He laid the foundation in thought that should have advanced much further today. Because IT messages do not often resonate with the executives and because executive leadership was in constant flux, progress is less than expected.

Daconta went on to write a book, The *Semantic Web* [9], published by Wiley, that surely made an impact. Yet, how can we accelerate the progress that the concept deserves? The answer is to engage executives in the subject in ways that matter to them most, as it should.

If we made a mistake, it is from not aligning the subject with the executive audience that most needs the solution.

A demonstrable example of how technology can take off when aligned with executives' needs and interests is social media. President Obama put the technology to work and now it has sprung an industry. Of course, we may not imply that President Obama is responsible for social media, though we can say that he and David Plouffe demonstrated how to use it, as it put the community organizer on steroids, so to speak.

One could assume that data is valuable or we would not be all consumed by the process of creating, storing, publishing, and otherwise transacting it in every aspect of our daily lives. However, how much do we respect data? For many executives and their enterprises, respect does not manifest in a strategy. Data does not appear as the first words in the annual report, though it very well should.

The following excerpt is from the 2006 Annual Report from AIG. "Dear Fellow Shareholders: 2006 was an excellent year, starting with the resolution of our significant regulatory challenges and ending with record financial results" [10].

To be fair the historical data was there, it just didn't say anything about the future. The place to begin is to make certain that we have a common agreement about critical terms. What do we mean by "data"? What do we mean by "smart data"? We can address this need straightforwardly, though the context would be missing. Which should come first? We'll try it two ways.

We don't like trite sayings, one of which is "smart data is that which knows what to do with itself." Yet there is merit for this and that is where we are headed. Data is engineered and with application of today's technologies it can be attributed such that it operates in a system enabling it to appear to qualified users in an actionable state and in a timely, accurate, and complete manner. Internet inventor Tim Berners-Lee proposes accomplishing this by tagging the data, although the W3C has not yet completed the standards for doing this as part of the semantic web initiative, where he has renewed commitment: "W3C refers to the World Wide Web Consortium created in October 1994 to enable the WWW to realize its full potential by developing common protocols to promote its evolution and ensure its interoperability. Having developed more than 20 technical specifications for the Web's infrastructure, including XML, DOM, and SOX, the W3C is now laying the foundations to meet the needs of the next generation of the Web as computers, telecommunications, and multimedia technologies converge" [11]. We advocate that executives commit to accelerating this work.

From Michael Daconta's posting, we summarize what he describes as the historical evolution of data: "Historically, data began as a second-class citizen locked away in proprietary applications.... Data was seen as secondary to the programs processing it. This incorrect attitude gave rise to the expression garbage in, garbage out or GIGO. GIGO basically reveals the flaw in the original argument by establishing the dependency between processing and data. In other words, useful software is wholly dependent on good data" [11].

Computing professionals introduced the notion of data quality and that data are assets.

Daconta reported that object-oriented facilities developed that made data more important, although internal to applications, so that vendors could keep data proprietary to their applications for competitive reasons. Daconta commented that data proliferated in volume and number of formats in the era of personal computers. "The most popular formats were the office document formats like Microsoft Word. These data formats and databases had proprietary schemas understood only by the applications processing the data" [11].

"The growth of the Internet in the 1980s and the World Wide Web in the 1990s began the shift away from proprietary formats." Daconta wrote that in 2003 [11] and we observe that many organizations remain trapped by continuing investments in old paradigm software.

He observed that the evolution of data has sped up in recent years, with major shifts occurring more rapidly. This speed is an effect of a more mature understanding of how to model data. "We are now moving into a new phase of the data evolution—the age of semantic models—in which the standardization of the World Wide Web Consortium (W3C)'s Ontology Web Language (OWL) will be the catalyst" [11].

It is important to understand that data evolution will not end with this phase; there are more fine-grained shifts and more follow-on phases to come. Daconta's treatise pretty well brings us up to today's fine-grained refinements.

Executives should be aware that their organizations need to support and participate in W3C standards initiatives.

We will present a case example in which we demonstrate model-based data exchange employing an open international standard, ISO10303 Product Life Cycle Support (PLCS), for instance.

At the time of this writing, the federal CIO Council is grappling with how to implement "data.gov," which is intended to make a broad array of government data available to everyone. From this, we expect improved policies regarding individual privacy protection and overall data security. The subjects related to smart data are a moving target.

House Subcommittee on Telecommunications and the Internet

Privacy, secrecy, and security are separate and related topics that we cannot ignore as part of smart data strategy, although the subjects and disciplines are a specialty that calls for independent treatment.

The follow excerpt is from the testimony of Larry Clinton, President, Internet Security Alliance, to the House Subcommittee on Telecommunications and the Internet, May 1, 2009:

> At her confirmation hearings two months ago, Secretary of State Hillary Clinton said that the single biggest threat to our country was the proliferation of weapons of mass destruction, and she identified four categories of these weapons: nuclear, biological, chemical and cyber.
>
> The former Director of National Intelligence Advisor to President Bush, Mike McConnell, has argued that "the ability to threaten the U.S. money supply through a cyber attack is [the] equivalent of today's nuclear weapon."
>
> Just 10 days ago, Melissa E. Hathaway, Acting Senior Director for Cyberspace for the National Security and Homeland Security Councils, previewed the report on cyber security she has provided to President Obama by saying: "The Internet is neither secure enough nor resilient enough for what we use it for today and will need into the future.

> This poses one of the most serious economic and national security challenges of the 21st century."
>
> The cyber security threat is much more serious than the well publicized massive losses of personal data. There are now recorded instances of our government and industry's electronic infrastructure being compromised by criminals, nation states, and even potential terrorists. The result of such attacks could be disastrous ranging from the possible shutting down of our electrical grid to having our own military weapons turned against us.
>
> On a purely economic basis, if a single large American bank were successfully attacked, it would have an order of magnitude representing a greater financial impact on the global economy than September 11. But the threat is not just speculative. Today, cyber security injuries are already substantial: some estimates now place the economic loss from known cyber thefts at more than $300 million per day. [12]

All of that is sobering to say the least. When accomplished to the level that we envision, smart data and associated tagging schemas combine with credentialing and privileging procedures that should greatly increase security and reduce vulnerability, although work must continue.

Definition of Smart Data as Envisioned

Generally, smart data is the product of engineering, such that it presents itself to the right users at the right place and time in support of planning, problem solving, decision making, sense making, and predicting. It comes with methodology and algorithms that support intended uses. It is interoperable and supports qualified users with diverse needs in an ubiquitous environment. It is the product of advanced data engineering technologies, including modeling and metadata management, and smart application of open standards. It accounts for credentialing and privileging as a dimension of security.

This definition will continue to expand as you read further. To be certain, smart data is our invention, as we want to make smart data and smart data strategies more explicit to people that depend upon it.

The U.S. government, on one hand, says that it wants to invest in technology that is readily available; on the other hand, it wants to make investments that are strategic. Investing in smart data capability is a strategic investment that includes (1) maximizing use of available technology while (2) pushing forward initiatives like the semantic web that will yield substantial enablement of smarter data.

General Discussion About Data from the Information Technologist's View

As an executive, you may hear the following perspective from your IT professionals. Our discussion here will help you understand what they are saying.

When using the term "data," it can be confusing because nearly everyone has an idea about what it means to them. IT guru Daniel S. Appleton instructed that "information = data facts + meaning." Data comprise entities that have attributes. George moved this definition along and said that knowledge = information in context. Wisdom = applied knowledge = knowledge + experience.

That sounded pretty good until we thought about it some more. In a layperson's terms, data describes a person, place, object, event, or concept in the user environment about which the organization wishes to maintain information. Surely, IT professionals will describe data as comprising entities and attributes. Open Source defines entity as "a self-contained piece of data that can be referenced as a unit; you can refer to an entity by a symbolic name in the Document Type Definition (DTD) or the document. An entity can be a string of characters, a symbol character (unavailable on a standard keyboard), a separate text file, or a separate graphic file" [13].

Attributes provide more information about data elements. In computing, attributes are entities that define properties of objects, elements, or files. Attributes usually consist of a name (or key) and value.

Is such a definition compelling to a CEO or cabinet secretary? It is probably not captivating because it sounds so abstract, however technically essential.

Some say that "facts" are things that can be shown to be true, to exist, or to have happened. Therefore data can be a representation of facts such that humans and machines can process it. The data facts require additional attributes that provide meaning and the attributes are themselves data. Data are not always factual. Data describing data is called metadata.

While writing this section we read a newspaper headline in the *Washington Post*: "Math Error to Cost Maryland $31 million." Apparently the state overpaid 17 schools for which it really didn't have the funds. The mistake was said to be a computational error, made by someone using a calculator in the Department of Taxation and Assessment. This is an example of dumb data, and surely not the product of smart data strategy.

The source erroneous: 17 school systems accepted money and intended to use it, even though the amount was more than they had budgeted.

Data are the equivalent of the atomic elements of information. Data is the plural of datum. Data are used to describe virtually anything including their representation in any medium. Information technologists need to differentiate data and programs such that programs are instructions for manipulating data: they say data on its own is without meaning.

Discussion About Data with an Enterprise CEO and Department Secretary

Euclid addressed the subject of data, associated with his work called *Elements*. It is a complex work and King Ptolemy requested a more elementary text on geometry. Euclid replied, "There is no royal road to geometry." We might give a similar reply to an enquiring executive about data, except to say that a data strategy map will make it easier to chart the course to higher performance.

How might the data strategy map appear? It most surely would include a comprehensive enterprise data model depicting the entities and attributes of all the data needed to manage and optimize performance in the enterprise. It most surely would include process models describing how work gets done that are attributed with data inputs and outputs, controls, and mechanisms. It might even depict data flow, and who uses it.

As you might imagine, a data strategy map would be complex, stored in a computer, supported by software, and served up for executives, users, and technologists to analyze and to act upon.

Data is the atomic matter behind government and commercial enterprise performance. It is the essential grist for the mill, although this metaphor stops far short of complex meaning.

Some people ask "Is not your book just as valuable to commercial enterprise as it is to government enterprise?" The short answer is yes. Others ask, "What is the relevance of government to commercial enterprise?" A longer reply is needed for this question. Although current history certainly illustrates the close relationship, albeit too close for some.

A prevailing notion for the past 12 years is that best commercial practices and technologies are best for government. While there is merit to this idea, application is not always direct and universal. We observe that government is in a perpetual catch-up situation, lagging commercial enterprise by 4 to 20+ years. There are structural, bureaucratic, political, economic, and cultural reasons for the lag, and we are certain that commentators like CNN's Lou Dobbs will find no good excuses.

Our approach aims at cutting through excuses. It also introduces another set of views about the relationship between commercial enterprise and government.

Global asymmetrical threats to national security demand that the U.S. government performs with superiority in the comprehensive realm of enterprise management that is enabled by superior technology. Since the government is in the lead of enterprise management, the institution cannot afford to wait for commercial enterprise or to wait to catch up—it must lead in the role of ultimate customer representing citizens. Note that this would be a strategic policy shift to demand that government not default to commercial industry in pursuit of technology leadership.

Commercial enterprises that support the government as prime contractors are often not servicing commercial customers. They are locked into servicing government and therefore are not in the position to be the sources of "best commercial practices." (There is a much larger question related to the military–industrial complex, although that is out of scope.)

Business algorithms for commercial enterprise are often the antithesis of business algorithms for government enterprise. Therefore one-to-one applicability must be challenged.

The common thread among government and private enterprise is data.

There is a pecking order in this book. Government executives, called department secretaries, are the equivalent of CEOs. Commercial executives in this book are considered in the context that their organizations are participants in the government-led enterprise as prime contractors, subcontractors, and suppliers. They are stewards

of their shareholders and employees, and their being served is contingent on serving customers in a superior manner.

Citizens are the recipients of services provided by government enterprise and they are the end-user customers. As stewards of citizens, government executives are customers who are served by participants in the supply chain. All participants in this scenario have symbiotic relationships. Success is measured by the following aggregation:

- Best for citizens
- Best for government
- Best for government trading partners
- Best for allies

A nation's capacity for "best" is constrained by capital, infrastructure, resources, institutional design, and intellect, among other things. A factor critical to overall success is the ability of the elements of enterprise to collaborate in achieving ultimate outcomes.

Collaboration is a characteristic describing how enabling mechanisms interact to produce outcomes from processes that are shared under complex rules by members of the supply chain and by customers themselves. Data are provided as inputs and created as outputs throughout the process, and agreements determine to whom the assets belong.

There are only a few executives who are performing on behalf of citizens. They are supported by a large number of organizations, operations management, and bureaucracy. They are dependent on a large number of commercial contractors and suppliers. The relationships are governed by laws, regulations, business rules, and contracts.

The term "enterprise" is employed to describe the aggregation of organizations—public and private—that work together to satisfy missions, goals, and objectives on behalf of citizens aka constituents aka communities. Enterprises have certain building blocks:

- Control architecture (more contemporarily called leadership and integration architecture)
- Inputs (capital, material, data)
- Processes that are prescriptions for how work gets done (activities that define how work gets done to produce required and desired outputs)
- Outputs (outcomes, products, services, assets, data, and results)
- Enabling mechanisms that may be considered the technical architecture (people, organizations, and technologies)

Note again that data are both an input and an output. Data can be an asset and can also be noise, just as outcomes can be positive or negative (i.e., value or costs, revenue or expenses). This is the construct by which Dennis Wisnosky, Chief Technology

Officer at DoD, led the Defense Enterprise Modeling effort. The artifacts from this effort will become assets and input for future administrations.

To optimize enterprise performance, CEOs and secretaries must be able to track data through processes and to realize data as assets, or to otherwise assess it as being costly noise. The degree to which executives and their operations management are competent at accounting for data assets can be realized as competitive advantage or operational superiority.

From a citizen's perspective, government should be performing optimally, maximizing service while exploiting scarce resources to best use and advantage. From a shareholder's perspective, commercial organizations are expected to provide the best value to customers while returning optimal profitability that is competitively best. Both scenarios are dependent on best uses of data and associated accountability.

The Information Clearing House (ICH) definition of enterprise is "a system of business endeavor within a particular business environment. Enterprise architecture (EA) is a design for the arrangement and interoperation of business components (e.g., policies, operations, infrastructure, and information) that together make up the enterprise's means of operation" [14].

In the U.S. government there have been and continue to be "enterprise architecture" initiatives intended to engineer performance as expressed in models. Making enterprise performance explicit is a good first step. It is also a continuing necessity. It is the modern day translation of laws and regulations into automated operations and automated contracting environment.

Unfortunately, government leaders are rarely around long enough to see the effort completed. Worse still is that leadership often does not know how to use the models because the IT people have not made them operationally useful.

The reality is that government leaders are elected to manage and make a difference *today*, and not so much in the years after they are gone from office. It is a rare occurrence when architectural assets trickle down to operational management and bureaucracies with any positive effect as they are trapped in the day-to-day necessities, by the legacy as it were.

Something gets done anyway. For instance, wars get started and waged. Departments' consolidation gets initiated on massive scale. Executives get in trouble for short-circuiting the system in ways that appear to be abusive of the system.

Government auditors declare as foul the discovered waste, fraud, and abuse, while they too are on a short term and subject to resource manipulation that undermines their effectiveness. This is not an acceptable picture and at the bequest of American constituents, new leadership is pressured to find better solutions to which this book contributes by addressing data, the atomic matter behind government enterprise performance.

Automated real-time auditing enabled by smart data is an area worth exploring and improving in this regard. Such would provide management with actionable controls as it would simultaneously provide independent flagging of potential problems.

Further complicating government enterprise performance is the fact that government leadership performs in short terms while addressing problems and needs of great scale and magnitude. As a nation, we have trouble reconciling this circumstance.

The way to stabilize the situation and to assure continuity in focus and progress is by providing a legislative framework that is supported by operational management policies. While government laws and regulations may flow to participants in the government supply chain, private enterprise may address its own continuity issues via accounting and auditing practices.

Assets

All assets are defined by data; however, not all data are assets.

Definitions of "asset" include all real or intellectual property owned by the enterprise that has a positive financial value; resources and/or property in the possession of an individual or business entity; everything a corporation (or government entity) owns or that is due to it—cash, investments, money due it, materials, and inventories, which are called current assets; buildings and machinery, which are known as fixed assets; and patents and goodwill, called intangible assets.

One cannot just substitute government for corporation and apply the definition because there are definite nuances in how government accounts for assets versus corporations. However, for our purposes the first definition is probably best.

Government may make investment in assets at various stages in the process of transforming from raw materials to end-use products. Government may contract with private enterprise to care for the assets, and while in custody by private enterprise where legal title may reside subject to contingencies. The assets may be destined for ultimate possession and use by government customers. Government may do this for the purpose of managing investments and limiting risk and liability, and to preserve flexibility in reassigning assets to alternative uses and purposes.

We observed this specifically in the case of material stockpile management, where investments are made in staging metal products in certain quantities at various stages in the production process, for instance. This is done to have surge capacity.

The Health and Human Services Department maintains surge capacity in medical suppliers. The Department of Homeland Security maintains surge capacity in logistics capability for use in nationally declared emergencies. All such departments must track materials for these purposes as assets for which they must provide visibility to planners and prospective users, as well as to property managers.

To keep track of assets, one needs to know the contractual business rules as well as the associated processes and how information feeds processes, and how information is transformed and transferred as a result.

Information transformation describes how data is changed. Information transference describes the status of the asset as it moves through various processes and by various custodians. These are separate, though associated, ideas. Complicating this further is the fact that some assets are perishable and have limited life cycles.

The authors recall an instance where the DoD procured a new jet fighter under a performance-based logistics contract. In the process of developing the autonomic logistics component of the contract, disparity between what the government thought it was buying and what the contractor thought it owned was revealed in discussions about protecting weapon system data. Such revelations are common as

tracking layers of data and elements is dynamic and subject to interpretation and legal determination.

Sometimes multiple government departments and agencies contract with the same supply chain for materials needed for surge capacity, for instance. The chances are that all departments will call on the surge inventory at the same time. Will the suppliers be able to respond to these multiple requests? We suspect not, because from our direct experience we believe the process for accounting for materials in such events is flawed as a result of the absence of smart data and smart data strategy.

Data Assets

All assets are ultimately expressed, represented, and defined as data. Therefore it is imperative to be able to track and manage the disposition of data assets. It is equally important to keep the system clear of extraneous data that is simply noise in the system.

All data that is in a system consumes resources to be processed, stored, and maintained. Delinquent data, that which is not maintained, can dilute the system's scarce resources. That is why Sid Adelman, Larissa Moss, and Majid Abai, authors of another book, *Data Strategy*, are consumed by the idea to rid the system of redundant data, for instance [15].

While eliminating unnecessary redundancy may be a worthwhile pursuit, what are the other options and priorities that represent strategic focus? It seems to us that a more direct strategy is necessary, such as knowing your data, tracking and protecting your data assets, and optimizing data transacting with your enterprise trading partners.

What Is the Work of Management?

From an executive perspective, managing the enterprise is all about the data.

Smart data will improve the enterprise via enterprise level data resource management with improved operational management reflected by more accurate and timely information, situational awareness, and decision making at all levels.

Enterprise management, executives, and operational management own the data. Information technology and information system specialists provide enabling support.

Using the three-schema architectural view of data, there is (1) an external view and that is as seen by the end user and their applications and in our case that is the executive; (2) a conceptual view and that is from the system designer's view and independent of storage mechanisms; and (3) internal schema that describes how data is organized, stored, and manipulated independent of the application. In our discussion about smart data, we traverse these different views and schemas but always with the executive view and the enterprise view assuming paramount importance.

For executives to receive the level of support they need from information technologists servicing a smart data strategy, they must know what they require, and this involves expecting data to be engineered with certain characteristics that make it smart, as well as expecting the infrastructure and services to be engineered to

support smart data strategy. Because you are pioneering a developing new idea, albeit with existing and emerging technology, it is an iterative process requiring close collaboration between executive customers and IT solution providers. You will be producing a capability that is best for your enterprise and that will produce superior results.

It is important to say this because we believe that, too often, the data ownership and responsibility is passed to IT or slips away from management control and visibility, getting lost in a sea of information technology. We want to make explicit how to channel data, and how to optimize enterprise performance with complete visibility.

For instance, the DoD embarked on a strategy to employ enterprise resource planning software as a panacea for escaping a myriad of legacy systems supporting defense logistics. The notion is that by having everyone on the same software, all organizations will be able to integrate operations. Trouble began when organizations were asked to throw out their business rules and to adopt those embedded in the software.

First, the ERP software, for instance, was originally developed for private enterprise that is motivated by a completely different set of metrics than government enterprise, notably profit motive versus constituent service utility motive. Second, the solution of choice is often from a foreign vendor, making U.S. defense systems and operations dependent on foreign software.

These large issues loom among many others and the circumstance has moved responsibility for enterprise data away from management and into the hands of supporting information technologists. How did this happen?

Some key government executives and former military flag officers ended up working for technology vendors as a part of the revolving door practice that compromises objectivity.

Alright, we know that American defense is dependent on foreign trading partners, customers, and allies. It is not necessarily bad to engage foreign suppliers in U.S. government programs. However, one must ask what capabilities and assets should a nation own or be able to deliver domestically in order to remain secure? These are executive questions that need constant attention as they affect policy and practice and, most important, data.

Accept the Premise

Concluding the discussion about the value of data to the enterprise, accept the premise that all assets are ultimately expressed, described, and accounted for as data. To optimize enterprise performance, the process is to take capital and material and, through enterprise processes, produce outcomes that are higher yield products and results, whereby the processes operate under constraints and the work is performed by people and technologies performing in concert. The people are aligned with organizations, and the relationships among organizations are contractually bound.

From the viewpoint of the enterprise head, the enterprise creates certain data that is critical to its performance. The enterprise receives data from external sources and transforms this data for internal purposes that translate into higher yield outputs. The

outputs are sent on to other users in the community—some of which are end users while others will employ the data as input into their own processes.

The purchase, lease, and use of data falls under constraints and agreements whereby there is consideration—monetary value exchanged for use either explicitly or implicitly.

As an enterprise executive, you must know your data assets, as that is what you must protect, and it is the collateral for which you receive revenue or benefit. You must know the data on which your enterprise performance is dependent. You must know about the sources of data. You must understand that you may influence the sources of data such that the data you receive is optimally engineered for your use.

Likewise, you must know where your data outputs are headed. You must be concerned about how well your outputs satisfy the needs of the user community.

Of parallel importance is to know where and how data are applied to support

- Sense making and predicting
- Planning
- Problem solving
- Decision making

A prerequisite is to know enterprise processes and how data feed processes, and how processes produce data. A part of describing or modeling processes is defining and accounting for business rules as constraints. Conversely, executives need to identify what data is needed to optimize enterprise performance. As Gertrude Stein might have put it, "What is the answer? What is the question?"*

Understanding data is dependent on semantics (meaning) and syntax (format) that manifest in lexicons (vocabulary), data dictionaries (reference books), and ontologies (organization of knowledge).

In a complex enterprise, associated with data are functional profiles. That is, your data has certain attributes that affect associated costs of access, storage, processing, publishing, and usage. The data attribute profiles have direct bearing on your own internal costs as well as having impact on the aggregate enterprise performance where you are a part of something larger.

There remains a clear field for accountants to think about how to ensure that aggregate benefits are distributed among the community, and how those that introduce excess costs pay for their deficiencies as the expense is born by the source and by users.

Exiting the disaster of 2008 and the Bush era, U.S. business must reestablish credibility, trust, and integrity, as must the government. Stimulus from government spending is intended to buy time for commercial enterprise to regain footing, although such cannot begin until the banking, finance, and lending system is reengineered to a suitable international standard that is under development simultaneously with the fix.

* In one account by Toklas, when Stein was being wheeled into the operating room for surgery on her stomach, she asked Toklas, "What is the answer?" When Toklas did not answer, Stein said, "In that case, what is the question?"

The management approach proposed in *Smart Data* can surely increase the probability of success while reducing associated risks as it promotes a scientific and quantitative process to enterprise performance management.

1.2 ENTERPRISE PERFORMANCE VERSUS ENTERPRISE INTEGRATION

CEOs and department secretaries talk about enterprise performance. When speaking with IT professionals, they talk enterprise integration. What's the difference? Answering this question takes considerable explanation, although in so doing, we can introduce some perspectives that will ultimately permit our readers to appreciate the focus on smart data to which we are headed.

Enterprise Integration

There are two principal definitions on the Web, for instance. One says that enterprise integration (EI) "is the alignment of strategies, business processes, and information systems, technologies, and data across organizational boundaries." We discussed this with electronic magazine publisher guru Bob Thomas, who had a magazine called *Business Integration Journal*. He changed the name to *Align Journal*, emphasizing that alignment of IT strategy and enterprise strategy is essential to optimizing performance. We subscribe to this idea.

Another definition of enterprise integration "refers to internal coordination processes among different core activities." The first definition expresses a more external and global view, while the second definition is more inwardly focused. The trouble with inwardly focused enterprise integration is that the result is often stovepiping.

What is wrong with stovepiping? It adds extra costs and often results in islands of automation that require extra effort, ranging from manual intervention to extra programming and maintenance to share and process data.

A popular notion the past few years is transformation. Applied to the enterprise, it means qualitative change. This was the primary topic in the 2008 political campaign and is a central theme of the new presidency.

According to some, transformation applied to defense is the term used for new methods in warfare integrating communication and technology. We think that it is much more and note the absence of the key term, data, from this definition.

We advised the DoD Business Transformation Office that transformation requires an enterprise view. With that we developed what we called the service-oriented enterprise (SOE) management paradigm. We draw from this experience to explain our views about enterprise integration. SOE is a management paradigm and strategy directed at improving performance from enterprise integration and from engineering information for interoperability. The Electronic Logistics Information Trading Exchange (ELITE) was a specific program managed by the Office of Secretary of Defense, Unique Identifier (UID) Program and in concert with the Logistics Enterprise Service Program Office at the Defense Logistics Agency (DLA).

The SOE paradigm was advanced as a best-practice approach to achieving results from enterprise integration placing enterprise context at the forefront and endpoint of all enterprise integration initiatives. SOE is presented as a replacement for traditional systems integration approaches that are too often narrowly focused, and from which results are functionally siloed, rarely producing the magnitude of improvement planned and costing far more than expected.

The SOE initiative addresses the causes and the remedies and identifies what actions and elements can improve optimizing enterprise performance with computer automation.

SOE was developed in concert with a subset of U.S. Department of Defense (DoD) customers who manage the most complex enterprise in the world. Because optimizing performance in defense enterprise is dependent on supply chains and complex relationships between government customers, contractors, subcontractors, and suppliers, this may be the source of best practices in addressing associated problems and needs. Needed is leadership that is aware of this: perhaps retaining Secretary of Defense Gates will increase the possibility that a strategy may take hold.

Transforming the U.S. government, including the DoD, is a long-term challenge for which economic pressure demands improvement from information technology investments and higher results from performance improvement initiatives. Continuing to keep U.S. commercial enterprise competitive in an increasingly global market also demands superior application of advancing technology, and the SOE initiative demonstrates what leaders of enterprises must do to achieve and sustain leadership positions through this means.

At the outset, it is imperative to have executive sponsorship. It is also imperative to staff the initiative with those who have domain expertise, as this is essential in garnering support from those who are responsible for the area of change and improvement. Improving operational performance with advanced technologies will result in changes in work design and resource deployment over the life cycle of the initiative.

A part of the SOE argument is that upfront investments in engineering information for interoperability will have significant downstream payback. From past experience, we know that organizations must be prepared for change such that they anticipate improvements from doing things differently.

In this section we outline topics that are prerequisites for team participants in helping adopt an enterprise viewpoint for developing solutions that contribute to optimizing performance.

A Comprehensive Review of the Service-Oriented Enterprise

SOE is the wide angle view of the enterprise and presents a context in which to address smart data. A film maker might open a movie with a wide angle view of the landscape before zooming into the cowboy on a horse, for example. SOE is our landscape.

Is data the horse or the cowboy in this metaphor? Neither, as data is more like the air we breathe in wide open spaces to keep living.

The goal of enterprise integration is to optimize performance through seamless automation of business and technical operations. Accounting for this, consider efficiencies realized from better work design and operational performance required to support the enabling information technology. Also, consider the impact of resulting automation realized as benefits among the community of participants: users, customers, members of the supply chain, and other stakeholders.

Accounting for benefits is addressed by answering what they are, who realize them, how they are measured, and how the enterprise rewards their attainment and discourages deficiencies.

Up to now there have been different strategies:

1. *Standardization-oriented strategies* emphasize getting members of the enterprise to adopt standards and implementation conventions and rigorously comply with them.
2. *Enterprise applications integration(EAI) strategies* depend on middleware and the development of interfaces that translate and convert information from one form to another for automated processing.
3. *Enterprise resources planning(ERP) strategies* depend on wholesale adoption of proprietary enterprise software to effect complete change to a common platform.
4. *Enterprise information integration and interoperability (EII) strategies* leap beyond standardization, EAI, and ERP and create a new generation of capability based on preparing information for integration in a manner that is open, nonintrusive, and loosely coupled.

It is in the latter category that we are developing and promoting smart data, although we extend the idea further with consideration of semantic web and data tagging techniques.

Introducing the SOE framework or paradigm reminds me of how Steve Martin introduced the notion in a play called *Picasso at the Lapin Agile* (the agile rabbit) [16]. EINSTEIN asks a question about a Matisse painting.

Sagot: I'll tell you what makes it great [taking the painting from the frame].

Gaston: The frame?

Sagot: The boundaries. The edge. Otherwise, anything goes. You want to see a soccer game where the players can run up into the stands with the ball and order a beer? No. They've got to stay within the boundaries to make it interesting.

Einstein: That frame is about the size of my book.

Sagot: Well, I hope that you chose your words carefully. Ideas are like children: you have to watch over them, or they might go wrong.

Preparing the canvas involves frameworks, architectures, reference models, and concepts of operations. The process is telescopic, microscopic, and otherwise kaleidoscopic.

Service-Oriented Enterprise (SOE) Topics Overview

- SOE Main Idea and Key Terms
- SOE Demand: Problems and Opportunities
- SOE Description: Framework
- SOE Engineering Disciplines: Practices
- SOE Business Case: Customized to the Situation
- SOE Enabling Technology: Methods and Tools
- SOE Implementation Plan: Strategy and Process

The information presented here is a review of a briefing that is intended to prepare participants embarking on enterprise performance improvement employing the SOE approach that is a framework in which smart data strategy may flourish.

SOE Main Idea and Key Terms Understanding the demand for change and improvement establishes context. "Enterprise" is the context for application.

Today's demand for change and improvement begins with the requirement for government to right the U.S. banking and financial system so that capital is available to drive the commercial enterprise engine. In all cases, resources are scarce and there will be increasing demand on government for higher performance from a much smaller footprint. Therefore automated support systems must be smarter and reliant on smart data.

Defining terms is essential to improving performance in an enterprise. Understanding the meaning of terms in context with their use is called *semantics*. Semantics is crucial to applying today's most advanced technologies that produce seamless automation among organizations, corporations, and individuals sharing information and completing functional transactions among them.

You may have heard the terms semantic web, semantic mediation, and semantic interoperability. These terms are used to describe the newly emerged and emerging state of modern computing.

They describe the circumstance whereby entities, organizations and people, can communicate automatically (self-governing) from their unique viewpoints while retaining their own lexicons of words, and are able to be understood accurately and completely by others with different viewpoints and terms. This describes our current state of pursuit, although information must be engineered to accomplish the result.

By contrast, heretofore, information technologists were more preoccupied with *syntax* the rules for formatting or structuring of information for automated processing. Semantics and syntax are subjects essential to understanding new strategies for producing higher enterprise performance through better use of information technology.

Wrestling with syntax and semantics, aerospace and defense prime contractors worked for over 10 years to harmonize the basic terms used to execute sales to the government that centered to a large extent on business rules and the terminology for

conducting business using electronic data interchange (EDI) standards and implementation conventions.

At the start of the process, EDI was already 10 years mature but had penetrated only a small percentage of the supply chain. Ten years later, harmonized EDI was accomplished among a shrinking number of remaining aerospace prime contractors and about 20% of the suppliers. EDI was too hard, too brittle, and too inflexible to afford full potential from electronic business automation.

EDI is an example of a standard that is by design inflexible. It results in users having to design brittle interfaces to support changes and improvements, which require a high maintenance effort and expense. Selecting this standard was not smart and therefore not reflective of smart data strategy. The goal was elusive, and the introduction of XML and Web-based service oriented architectures (SOAs) destabilized the effort. XML has more desirable characteristics than EDI and affords greater flexibility, though no less rigor among user participants to maintain.

On the one hand, great investment was committed to "standardizing" with the use of EDI standards. On the other hand, introduction of new Web-friendly standards opened the window for greater flexibility.

Industry faced a dilemma: (1) continue to wrestle with a rigid standard or (2) expect potential for improvements from a new standard. A third alternative was to consider the application of new data exchange technologies possible in SOA that would increase flexibility in applying standards that are inherently rigid.

Using advanced methods and technologies, enterprises can greatly reduce the effort required to maintain standards-based interfaces by engineering information for interoperability upfront.

The ultimate goal is to provide benefits from automation to all enterprise participants, increasing information sharing and reducing operational costs.

Returning to the term "enterprise," CEOs and senior executives and managers have active roles and responsibilities for managing technology-oriented improvement strategies. While CIOs and information technology professionals are partners in the process, superior results are produced through optimal working relationships throughout the enterprise.

Generally, "enterprise" refers to the undertaking necessary to accomplish something. An undertaking needed to produce an automobile that solves consumers' transportation needs and the need for fuel economy must certainly be considered a complex enterprise. An undertaking to develop a metrology machine used to measure an automobile piston or to measure the precision of jet engine turbine blades may also require a complex enterprise.

The enterprise needed to produce national security for the United States of America is enormously more complex than any of the previous examples. Integrating 22 disparate departments and agencies into an effective homogeneous system for delivering security to the homeland is a task of similar complexity. On the premise that these initiatives began with a flawed management approach, it is unlikely that the condition today is as it should be.

The National Institutes of Health describes enterprise by the following definition in its handbook of "net" terms: "In the computer industry, an enterprise is an

organization that uses computers. In practice, the term is applied much more often to larger organizations than smaller ones." we don't find this definition particularly useful.

Another definition is more suitable; this one is from the Interoperability Clearing House (ICH): "Systems of business endeavor within a particular business environment. Enterprise architecture is a design for the arrangement and interoperation of business components (e.g., policies, operations, infrastructure, and information) that together make up the enterprise's means of operation" [14].

Here is an enterprise definition that we prefer from an information technologist's view. "An enterprise is an organization with partially overlapping objectives working together for some period of time in order to attain their objectives. The actors utilize technology, competence, information and other resources in order to transform input to products that satisfy the needs of customers" [17].

So you see that one can shop for terms and meanings that best apply to a certain viewpoint and situation. That is what we all do as individuals, some more precisely than others.

When the number of constituents for your messages is large and complex, there are different strategies for producing timely, accurate, and complete communications that may be used to consummate business transactions.

All of these definitions of enterprise are applicable to our discussion. We recall a definition of enterprise used by Daniel S. Appleton, IT guru. He said that an enterprise comprises organizations that may include commercial partners and government customers that are linked together by contingent claim contracts. He emphasized the importance of understanding the dynamics of business rules and how they operate throughout the enterprise.

As a commercial organization conducting business with the government, or as a government entity operating as a part of a larger organization, both examples are enterprises with their own internal characteristics that link with other entities to form a greater whole that is itself a larger enterprise.

The degree to which the linkages are seamless or efficiently integrated to produce the desired output is what we describe as the enterprise performance domain. If an enterprise can be described by its essential operating components with a degree of consistency in method and technique, we can more clearly understand how to optimize the linkages.

Integrating linkages is not the end game. The end game is delivering the needed services or products in a timely, accurate, and complete manner at the best cost and with the best quality. This is what is meant by "service" in the service-oriented enterprise.

Integrated is a term used throughout this discussion. For software developers it may describe a group of application programs designed to share data. For an enterprise executive, integrated may describe the linking of controls (business rules), processes, and information among different members of the enterprise (i.e., trading partners).

What is common in these two ideas is shared data or information. Data = facts meaning. Information = data + understanding, or data in context. Some consider

the term "integrated" to be associated with tightly coupled linkages that are accomplished by application of standards and development of enterprise application interfaces. By contrast, some use the term *interoperability* to describe linkages.

"Interoperability" implies the ability of a system or product to work with other systems or products without special effort on the part of the customer or user. Words such as loose coupling, adaptive, and noninvasive are used to differentiate interoperable linkages from integrated linkages. These ideas are important in understanding the attributes and characteristics of enterprise performance-improving strategies and solutions.

They help the people who must plan, budget, and pay for solutions to better understand the trade-offs from the technologies enabling results. Technologies affect how work gets done and how results are produced with associated attributes and metrics.

When there is a surplus of labor or people needing work, some leaders may be more inclined to use people versus technology to perform work. When there are shortages, outsourcing and worker immigration are alternatives. The sociogeopolitical trade-offs are real for government leaders and for American commercial enterprises alike.

Metaphors are used to help explain complex ideas. Supply chain and value chain are used to describe the relationship among trading partners that comprise an enterprise. *Chains* describe dependent relationships, yet the linkages we are developing with application of new technology may be better described as neurons and synapses. (Neurons are nerve cells that can receive and send information by way of synaptic connections.) Chains are rigid and hard to break. Neurons and synapses work as more agile and adaptable means of linking.

Those performing enterprise integration need to adopt a different view from systems integration. To illustrate this, compare systems integration to that of a castle and enterprise integration to that of the open marketplace. There is only one way into the castle and that is to take the bridge over the mote that is opened and closed by request to the gatekeeper. By contrast, the open market may be accessed by anyone entering from all sides. The difference is context. A commercial enterprise participating in the government market must cross many motes and gatekeepers, though the goal is to make the enterprise more open like the marketplace.

Systems integration is more concerned about performance inside organizational or functional boundaries that are often called *silos*. Enterprise integration must consider how software solutions perform for all of the entities that share information. As much value is given to how systems perform in the enterprise context as is given to the local context.

A more advanced topic has to do with consideration of self-organization versus entropy and striking a balance between freedom and constraints. From the viewpoint of a supplier or subcontractor that is a member of the larger enterprise, and from the viewpoint of prime contractors and ultimate customers, achieving the optimum degree of freedom while ensuring enterprise-wide interoperability is a goal.

SOE Demand This section describes why a new management paradigm and strategy is needed to improve results from enterprise integration initiatives. The concept of service-oriented enterprise (SOE) is rich with substance applicable beyond enterprise integration in that the "service-oriented" values and optimizing enterprise performance are considered first.

When we discussed SOE with the Technical Operations Vice President of the Aerospace Industries Association (AIA), he asked, "Is SOE a business or technical-oriented paradigm?" That is an interesting question because managers and technical professionals often ask this question to determine if and how it fits their viewpoint. SOE addresses both business and technical functional requirements in a comprehensive and integrated manner.

In an article published in *Business Integration Journal* [18]. Hall Bailly, and George outlined the following takeaways.

1. ***Business*** "A stable framework is needed to help enterprise management configure a collaborative enterprise from independently migrated processes and applications. The ability to describe application architecture for EI is essential to the decision process. SOE will help in understanding the immediate costs of directed investment in EI and its benefits in reducing costs for ongoing interface maintenance."
2. ***Technical*** "Without an overarching strategy and associated architecture, many EI activities continue to be ad hoc and, as such, non-predictable and non-repeatable. As a strategic approach to EI, SOE blends structural elements in harmony with three software engineering disciplines. The objective is to achieve the optimum balance between attention to the elements and application of SOE disciplines."

We wanted to present evidence demonstrating the need for a new paradigm and strategy. Software vendors such as Metamatrix offered a commercial example, but we needed an objective third-party source. From a government perspective, the General Accounting Office (GAO) reported that program and project managers need more disciplined processes to reduce risks associated with enterprise integration and systems engineering. There is too much rework needed to stabilize applications after deployment.

Pressure to get systems implemented quickly has shortchanged investment in information engineering to get them right. Our view is that more investment upfront will deliver high reward downstream. This begins with smart data strategy and adoption of supporting values and principles.

Another aspect of the demand for SOE is from continuing to support the value for information sharing. Former Defense Deputy Under Secretary of Materiel Readiness, Dave Pauling, advanced the idea that "improving operational performance is realized through improved planning, sense making, and decision making that is currently too disjointed and costly to maintain." Information interoperability deficiencies are the cause. Correcting the cause eliminates proliferation of deficiencies while maximizing the distribution of benefits.

1.3 CURRENT PROBLEMS AND DEFICIENCIES FROM POOR DATA STRATEGY

Current problems and deficiencies from poor data strategy must be addressed in the context of enterprise resource management and the SOE paradigm. Enterprise resources are constrained and demand for change and improvement often outstrips available resources. Here are some considerations that we imposed in constructing the SOE paradigm:

- Enterprise scope and context
- Acknowledging that resources are insufficient to replace the magnitude of legacy systems
- Necessity to keep pace with advancing technology
- Necessity to deliver significant improvement from constrained investment
- Make application least invasive while delivering immediate tangible benefit

Therefore we concluded that a new strategy is needed based on addressing the need for information interoperability.

All enterprise environments operate under resource constraints that manifest as shortages of capital, and shortages of scarce talent and materials, for instance. Demand for improvement and change nearly always outstrips an enterprise's capacity for making change and improvement.

Different things, often working in combination, create demand for improvement. Demand for improvement and change is different for government and commercial enterprise. While some things are the same, others are the opposite. Government enterprise is about the business of maximizing citizen services. Commercial enterprise is about the business of maximizing return to shareholders. When the government is the customer of commercial enterprise, government negotiates for the best value, while business negotiates for the best margin. Buyers and sellers need one another and in a competitive marketplace, agreeable parties will connect. Correct classifications are essential to this process.

In an environment in which resource constraints manifest themselves as capital shortages, cost containment is a major concern to the SOE. Not only are correct classifications of costs important in a smart data–SOE scenario, but understanding misclassifications merit a closer inspection as well.

Smart data can be utilized, in a combination with artificial intelligence methods, to shed light on cost classification within the SOE. For example, a misclassification cost matrix can be incorporated into nonlinear neural-evolutionary and genetic programming-based classification systems for bankruptcy prediction. This smart data–SOE approach, properly implemented, could have helped to prevent and predict the 2008 Wall Street collapse, if implemented in a smart data paradigm.

Remember that smart data provides the following benefits: enterprise performance enhancement, end-to-end process optimization, and continuous optimization. Smart data also leads to better enterprise decision making—timelier, more responsive, and

increased agility—that is more responsive to change, has lower IT cost to change, and achieves continuous optimization.

In statistics, the terms Type I error (α error, or false positive) and Type II error (β error, or a false negative) are used to describe possible errors made in a statistical decision process. When an observer makes a Type I error in evaluating a sample against its parent population, he/she is mistakenly thinking that a statistical difference exists when in truth there is no statistical difference. For example, imagine that a pregnancy test has produced a "positive" result (indicating that the woman taking the test is pregnant); if the woman is actually not pregnant, however, we say the test produced a "false positive." A Type II error, or a "false negative," is the error of failing to reject a null hypothesis when the alternative hypothesis is the true state of nature. For example, a Type II error occurs if a pregnancy test reports "negative" when the woman is, in fact, pregnant*.

Most classification systems for predicting bankruptcy have attempted to minimize the misclassifications. The minimizing misclassification approach assumes that Type I and Type II error costs of misclassification are equal. This assumption may be a fallacy. Remember that wisdom = applied knowledge = knowledge + experience.

There is evidence that these costs are not equal and incorporating these costs into the classification systems can lead to superior classification systems. The principles of evolution can be used to develop and test genetic algorithm (GA)-based neural and genetic programming (GP)-based classification approaches that incorporate the asymmetric Type I and Type II error costs. By applying these artificial intelligence methods and mechanisms to decision making, it is possible to move toward the smart data "wisdom" paradigm of knowledge + experience, in which data comprise entities that also have attributes.

Using simulated and real-life bankruptcy data, we compared the results of the proposed approaches with statistical linear discriminant analysis (LDA), back-propagation artificial neural network (ANN), and a GP-based classification approach that does not incorporate the asymmetric misclassification cost. In essence, we are providing a data strategy map in order to make it easier to chart the course to higher performance. We are moving toward "wisdom" because we are investigating both a linear and nonlinear approach and comparing the two outcomes.

The results indicate that the proposed approaches, incorporating Type I and Type II error cost asymmetries, result in lower Type I misclassifications when compared to LDA, ANN, and GP approaches that do not incorporate misclassification costs. These smart data truths should be self-evident and useful as a common thread among both government and private enterprise. The smart data "wisdom" here is to investigate both the linear and nonlinear approaches and to compare the outcomes for the one that demonstrates the most cost containment.

We must remember that smart data is the product of engineering, such that it presents itself to the right users at the right place and time in support of planning, problem solving, decision making, sense making, and predicting. In this case, there is

* Wikipedia definition as of 2008.

evidence that costs are not equal and incorporating costs into classification systems can lead to superior results.

Our smart data approach to Type I and Type II errors comes with methodology and algorithms that support intended uses. Our approach is interoperable and supports qualified users with diverse needs in an ubiquitous environment. It is the product of advanced data engineering technologies including modeling and metadata management and smart application of known standards.

Bankruptcy prediction is an active area of research in finance. Several analytical approaches have been proposed beginning in the late 1960s. Among the popular approaches for bankruptcy prediction are the use of statistical discriminant analysis, artificial neural networks, decision trees, genetic algorithms, and probabilistic approaches, such as logit and probit. Researchers use analytical techniques for prediction of bankruptcy. Using statistical discriminant analysis and data on a set of companies that went bankrupt during the period 1946–1965, Altman [19] showed that statistical discriminant analysis is a viable tool for prediction of bankruptcy.

Following Altman's study [19], researchers investigated the use of probabilistic approaches to predict bankruptcy and some used a maximum likelihood estimation of the conditional logit model with the objective of making probabilistic estimates of insolvency. Logit is a linear technique that does not require any assumptions about the prior probabilities of bankruptcy or the distribution of predictor variables.

Unlike linear discriminant analysis (LDA), logit does not specify a cutoff point delineating bankrupt firms from nonbankrupt firms. The model assigns each firm a probability of bankruptcy. The decision makers can then choose a level that they are willing to tolerate. The trade-off is between choosing a higher Type I or Type II error.

Several machine learning techniques were also used for prediction of bankruptcy. Among the machine learning techniques used for prediction of bankruptcy are artificial neural networks (ANNs) and genetic algorithms (GAs). Since these techniques do not rely on any distributional assumptions about the variables, they avoid problems associated with LDA and logit.

It has been found that although LDA provides superior results, GAs are effective tools for insolvency diagnosis, since GA results are obtained in less time and with fewer data requirements. ANNs have been used for bankruptcy prediction and several researchers have reported a better performance of ANNs against statistical LDA and logistic approaches.

The foregoing discussion, about the value of correctly classified smart data to the enterprise, accepts the premise that all assets are ultimately expressed, described, and accounted for as data. Examining this data can lead to cost savings and optimizing enterprise performance. In this case, we showed how to improve the process of minimizing costs.

We examined capital and material and, through enterprise processes, produced outcomes that are higher yield products and results, whereby the processes operate under constraints and the work is performed by people and technologies performing in concert. From the viewpoint of the enterprise head, the enterprise creates certain data that is critical to its performance.

In bankruptcy prediction, and in other classification problems, the costs of Type I and Type II errors are important considerations for a decision maker. There is evidence of an asymmetric cost structure, with an estimate of Type I error cost that is higher than Type II error cost. The cost of Type I error was estimated from the loan loss experience of banks, and the cost of Type II error was the opportunity cost of not lending to a nonbankrupt firm because it was predicted to become bankrupt.

Past studies employing LDA, ANNs, and GAs do not allow a user to incorporate asymmetric costs of misclassification. As a result, the Type I and Type II error costs are considered equal in most past studies. A probabilistic approach such as logit, however, allows a decision maker to trade off between Type I and Type II error by setting the cutoff probability. The technique, however, does not provide any guidelines for deciding on the cutoff probabilities, which makes it difficult to use. Recently, GAs have been used to learn connection weights for an ANN.

Among the advantages of using GA to train ANN connection weights are flexibility of design of the fitness function and global search. The flexibility of the design of the fitness function makes it possible to use GAs to learn connection weights of an ANN so that Type I error is minimized. Furthermore, the discriminant function learnted by using GA-based ANN is a nonlinear function and is likely to perform better than linear discriminant functions developed by LDA and GAs.

The preceding discussion provides the basic underpinnings of the bankruptcy prediction problem, that is, information = data facts + meaning. However, it is not until a comparison of the various linear and nonlinear techniques are made and the evidence that Type I and Type II errors are not equal, that actual knowledge or "information in context" has been supplied to move the decision maker toward the smart data "wisdom" paradigm shift. The knowledge + experience context is to recognize that most classification systems for predicting bankruptcy have attempted to minimize the misclassifications.

The minimizing misclassification approach assumes that Type I and Type II error costs of misclassification are equal. This long-held assumption may be a fallacy. Remember that wisdom = applied knowledge = knowledge + experience. The lesson to be learned is that there is evidence that these Type I and Type II costs are not equal and incorporating these costs into the classification systems can lead to superior classification systems. The common thread among government and private enterprise is data.

In the following case, we demonstrate what we stated on page 45 how the enterprise receives data from external sources and transforms this data for internal purposes that translate into higher yield outputs and helps to prevent accounting fraud in the organization. The outputs are sent on to other users in the community—some of which are end users while others will employ the data as input in their own processes in order to prevent fraudulent activities in their enterprises.

The purchase, lease, and use of data falls under constraints and agreements, such as generally accepted accounting standards, whereby there is consideration—monetary value exchanged for use either explicitly or implicitly. Material misuse of these assets can be extremely costly to the organization. It is the job of the enterprise executive to

know the data assets and protect them, and these assets are the collateral for which the enterprise receives revenue or benefits.

Executives must know the data on which the enterprise performance is dependent. They must know about the sources of data and understand how they may influence the sources of data such that the data received is optimally engineered for their use. Likewise, top executives must know where their data outputs are headed and they must be concerned about how well their outputs satisfy the needs of the user community.

The integration of smart data, application of artificial intelligence techniques, and adoption of a smart data—SOE paradigm can be used to counter accounting fraud fiascos, such as Enron and World Com, and are set forth in the case "Utilization of Data Mining Techniques to Detect and Predict Accounting Fraud: A Comparison of Neural Networks and Discriminant Analysis" [20]. Accounting information systems enable the process of internal control and external auditing to provide a first-line defense in detecting fraud.

There are few valid indicators at either the individual or the organizational level which are reliable indicators of fraud prevention. Recent studies have shown that it is nearly impossible to predict fraud. In fact, many of the characteristics associated with white-collar criminals are precisely the traits that organizations look for when hiring employees. This case proposes the use of information systems and smart data and utilization of the SOE model to deal with fraud through proactive information collection, data mining, and decision support activities.

Here is an example of citizens being recipients of services provided by government enterprise. All participants in detecting accounting fraud have a symbiotic relationship. Success in detecting and deterring accounting fraud is measured in an aggregation that is best for citizens, best for government, best for government trading partners, and best for allies.

Results show that while traditional methods, such as discriminant analysis, yielded 50.4% of original grouped cases correctly classified, no significant relationship was found (0.149) between attitude, morale, internal controls, increases in expenditures, and whether or not fraud was actually committed. Cronbach's alpha of reliability was 0.6626 and offered somewhat reliable results in this exploratory research. Neural networks did a much better job of predicting fraud (75.9%) than discriminant analysis (50.4%). Neural networks were able to find patterns in the training set and then correctly identify more than three-fourths of similar patterns in the testing set.

Therefore it can be concluded that neural networks outperform discriminant analysis by 25.5% in this data set. It is not until a comparison of the various linear and nonlinear techniques are made that actual knowledge or "information in context" has been supplied to move the decision maker toward the smart data "wisdom" paradigm shift. The knowledge + experience context is to recognize that the application of artificial intelligence mechanisms for predicting fraud involves not only the context of a symbiotic relationship between the stakeholders, but also a realization that, in a smart data sense, we do not necessarily live in a parametric, bell-shaped world. Nonlinear relationships must be taken into account as well if we are to find true "wisdom."

Problem and Opportunity Types

Possible Problem Types

- Deficiency in leadership and integration
- Deficiency in mission, values, policy, regulations, rules
- Deficiency in strategy
- Deficiency in framework
- Deficiency in planning
- Deficiency in sense making
- Deficiency in decision making
- Deficiency in systems
- Deficiencies of knowledge
- Deficiencies of skill
- Deficiencies of proficiency
- Deficiencies of execution
- Deficiencies in the balance of consequences
- Deficiencies in tools and equipment
- Deficiencies in methods
- Deficiencies in processes
- Deficiencies in infrastructure
- Deficiencies in organization

Possible Opportunity Types

- Better use of capital
- Better use of people
- Better use of technology
- Better use of materials
- Better competitive advantage

In the past six years, we have seen the introduction of "enterprise portfolio management" to evaluate information technology initiatives and their effectiveness. We have learned that it is difficult to collect all of the information about IT initiatives, programs, and projects to perform objective evaluation. A reason for this is executives' and CIOs' failure to manage performance improvement initiatives as they might other business decisions and operations.

Managing today's highly automated enterprises places increased demand on executives and their management teams to have higher competence in the discipline of management science that is interlocked with the disciplines of computer science and engineering. It is the nature of the enterprise that forces our rethinking about the required disciplines.

Change and change management are actually very mature disciplines; however, application is now integral or embedded among those who are planning, designing, and developing enterprise systems. While the work of employees is changing in modern enterprise, so is the work of their leaders.

SOE Framework

We present the SOE framework in different ways to help understanding. At the core are structural elements of enterprise integration: smart data, smart grid, and smart services. Required software engineering disciplines include data engineering, process engineering, and grid engineering, each of which set requirements for integration into the enterprise.

Admittedly, there are questions about why we chose these labels, and questions about relationships. If this causes you to think about it and to develop a better understanding, we will claim that as our intention. Confusion without resolution is not our intention.

The highest order activity for enterprise management is to optimize performance. Figure 1.2 illustrates the key relationships needed to optimize performance in response to the need for supply chain transformation, for instance.

Inputs are arrows entering the optimization activity and include problems and opportunities, structural elements, and information. The purpose of the activity is to produce solutions with desirable metrics made possible by producing harmonious structures and interoperable information. SOE disciplines and technologies are the enablers, shown as arrows entering the bottom of the activity. Enterprise integration guidance and control is provided by the SOE paradigm and strategy, accompanied, of

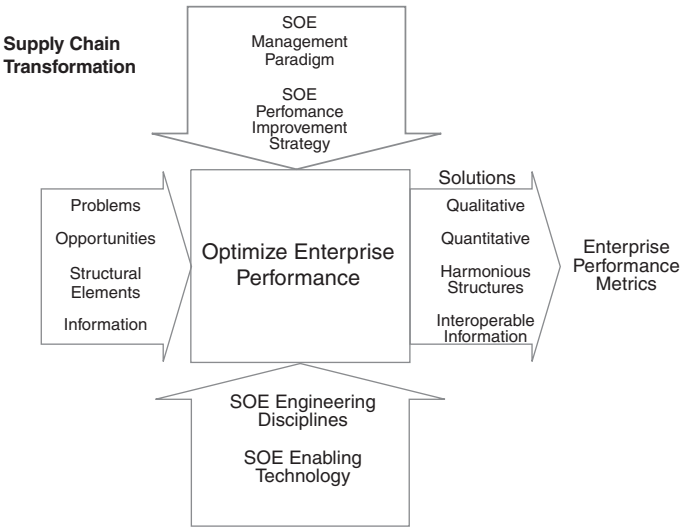

Figure 1.2 Optimizing performance with SOE.

course, by a host of other considerations that are accounted for in the details of SOE implementation.

What makes SOE different? Engineering information interoperability and embracing the value for openness and vendor independence are critical values.

What Is the Current Situation?

The current system software environment is the product of collaboration among mutually developed applications, hosted on a fixed configuration of structurally integrated platforms.

Undesirable qualities include:

- Tight coupling
- Brittle interfaces
- High maintenance
- Great difficulty in changing and upgrading

The desired state is a product of development from the enterprise view accomplished through collaboration among independently configured and structurally isolated platforms.

Acknowledged in this approach is the real-world situation that is fraught with legacy systems. Desirable qualities include more agile and less invasive solutions performing better to the benefit of the entire enterprise.

Figure 1.3 identifies engineering disciplines, advancing technologies, and structural elements in relative association. Effective enterprise integration requires an enterprise view, not silos or stovepipes. Optimizing enterprise performance requires refined engineering disciplines accompanied by advancing technology and with attention to certain structural elements (Figure 1.4).

Engineering Disciplines	Advancing Technologies	Structural Elements
Data Engineering	Enterprise Information Integration Metadata Management	Smart Data
Grid Engineering	Service-Oriented Architecture NetCentricity	Smart Grid
Process Engineering	Business Intelligence Modeling Enterprise Core Web Services	Smart Services

Figure 1.3 Starting point: effective enterprise integration requires an enterprise view.

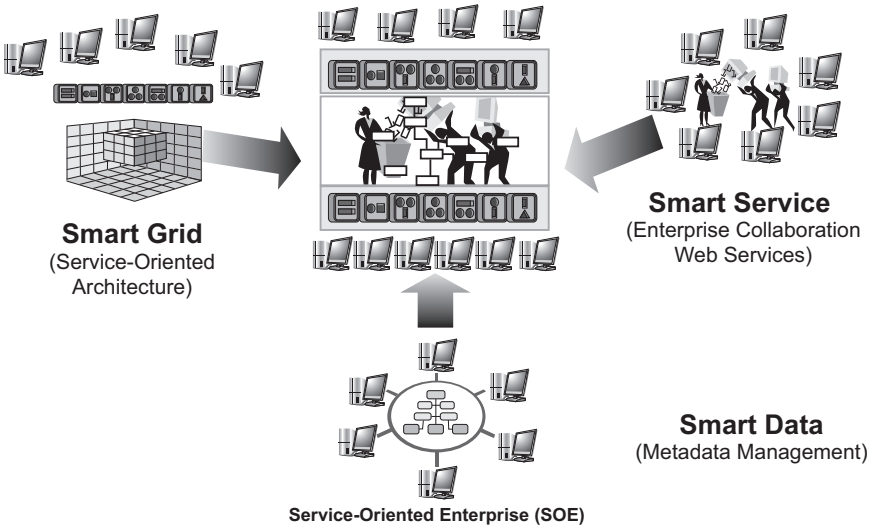

Figure 1.4 Key structural elements.

SOE includes an icon. Subsequent description will introduce the elements comprising the icon.

1.4 NEW TECHNOLOGIES

Critical to the success of enterprise-wide strategy is ownership by the CEO in private enterprise. Leaders in the top positions of any organization command, lead, and integrate with the following elements: power, time constraints, focus, attributes, and outcomes. These elements can be seen in the following two case studies. In one case, Netscape built a successful e-business—Anthill.com—was successful, and sold for a profit. The second case, Fannie Mae, is a case study in critical success factors gone wrong.

Power translates into value-setting, business rules, resource allocation, and balance of consequences—that is, making certain organizational participants clearly realize the difference between positive behavior and deficient performance. Power was abused at Fannie Mae and even though the company was successful in the short term, the corruption of power led to its downfall. Time constraints force the identification of milestones, metrics, and outcomes in a meaningful period with associated metrics such as rate and frequency.

Netscape identified these constraints successfully in Anthill.com. However, over time Fannie Mae ignored these time constraints and failed. They got away from the critical success factors (CSFs) that made them successful originally. Focus describes the domain in which change and improvements are to be realized, leading to the formation of themes. The domain for Netscape and its theme for Anthill.com stayed focused on building a successful e-business. Fannie Mae, on the other hand, lost its focus and failed.

Attributes are the detailed characteristics that describe desired process and associated performance metrics such as cost and quality, in addition to previously mentioned budget constraints (resource allocation) and time. Netscape paid attention to cost and quality metrics; Fannie Mae lost site of these important metrics.

Outcomes are the end results that include everything from physical products, assets (including information assets), and performance results. Netscape sold Anthill.com for a profit and fulfilled its outcome objectives of turning a profit. Fannie Mae's end result, on the other hand, was bankruptcy, because, in the long term, it did not continue to pay attention to its information assets and performance results.

New technologies have been affording new and different strategies for years. Today, SOE is a strategy for developing business capability and helping the organization to achieve business agility and to create an information-sharing environment that can create new applications to support changes in goal requirements, increase the speed at which services and information can be shared to benefit others, and securely connect people and systems in real time.

Case Study: Building a Large-Scale E-Business from a Small Anthill

Dr. Rodger wrote a case entitled "Building a Large-Scale E-Business from a Small Anthill: A Case Study" to gain insight into how a company can incorporate important principles learned by other successful e-businesses. This study examined four essential principles used by Netscape Corporation for building a successful e-business as they relate to Anthill.com, which was a relative newcomer to the e-business world. Data were gathered via in-depth interviews with Anthill.com executives. It is hoped that other e-businesses will recognize the importance of closely following these principles to improve the opportunity to develop into a large-scale e-business.

We systematically analyzed the practices of Anthill.com, even though there was a paucity of guiding frameworks for successful e-commerce. We felt the framework was an appropriate first step for rigorous analysis of our case.

Principle 1: Create a compelling, living vision of products, technologies, and markets that is tightly linked to action.

The greatest strength of Netscape's vision was its ability to create a tight link between senior management's high-level view of the world and the products it delivered to the marketplace. Netscape's vision did more than map a path through the confusion of the Web's early years. It also mobilized the company's troops to develop and deliver an impressive range of client and server products in a very short period of time.

From humble beginnings, Curt Matsko and Scott Alexander envisioned developing a startup Internet business into a multimillion dollar enterprise. The two friends began the Anthill.com business from a two-bedroom apartment in 1999. Their vision was to bring as many small-to-medium businesses in secondary markets across the United States together into one e-business location. These secondary markets included cities having populations between 10,000 and 250,000 people. As a result of this vision, individuals are able to

develop their business in a central Internet location that is highly visible to consumers.

Likewise, small e-business owners benefit from being able to place their business in a visible location without enormous cost or effort. With advances in Internet technology and hard work, Matsko and Alexander are beginning to realize their e-business vision. With Web access now very common and relatively inexpensive, Anthill.com is currently adding up to 1000 small business sites per week and is in the process of capturing over 10% of the entire small business market. This forward thinking is benefiting both consumers and retailers involved in e-business.

Principle 2: Hire and acquire managerial experience in addition to technical expertise.

Netscape's strategy of hiring experience was not restricted to the top ranks; it extended throughout the organization. Managers at every level tried to bring on board people who would hit the ground running. Netscape did not hire many green college graduates, fresh from studies in programming or marketing. Instead, it looked for people who had actually done these jobs.

Anthill was pleased with company growth but realized other resources were necessary for the future of the company. Matsko was sure that they would need an in-house lawyer to assist with legal issues and a possible move to bring the company public. "We didn't know what to expect, but we figured that rapid growth doesn't come without pain. I discussed with Scott the possibility of recruiting Dan Thurber for legal support. He is an outstanding and experienced attorney as well as a personal friend." Scott concurred with the choice. "I think he would make an important addition to our team. We not only need Dan, but should also approach his brother Brad. He knows the markets and is an exceptional manager" (S. Alexander, personal communication, November 5, 1999).

Since their company was made up of only the two partners, it seemed prudent to hire the management and legal experience they lacked. In addition, they knew the people they were hiring on a personal level: Matsko and Alexander thus followed this Netscape strategy. They brought in the Thurber brothers who have experiences in managing and providing legal advice for new businesses. Dan Thurber, a successful New York City attorney, was so impressed with the company's potential that he joined in 1999 while taking a 70% reduction in salary. His brother Brad quickly followed Dan's lead and joined. Brad emphasized the importance of his decision with the following comment, "the potential growth, stock options, and work environment at Anthill.com made it an easy decision for me to leave my current position as a successful stock consultant" (B. Thurber, personal communication, November 5, 1999). Substantial growth of Anthill.com has greatly increased the number of employees necessary to run the operation. Matsko developed a unique recruitment strategy that has enabled the company to add over 45 highly qualified employees. The recruitment strategy is unique at Anthill.com because it focuses on future

growth and employee ownership, rather than traditional salaries and retirement plans. Employees at Anthill.com have bought into the vision, potential, and excitement of managing an e-business. "Employees that have joined the Anthill believe in our business model, and are eager to obtain stock ownership in the organization over competitive salaries which we cannot offer" (C. Matsko, personal communication, November 5, 1999).

Principle 3: Build the internal resources for a big company, while organizing like a small one.

Most start-up companies scale their systems to meet their current needs. In fact, they usually allow their systems to lag behind their growth. One of the biggest traps for an entrepreneur is to build an organizational structure in advance of sales, profits, and stable cash flow. Far too often, wildly optimistic sales projections do not materialize, the company gets overextended, and everything comes to a crashing halt.

Attracting attention to Anthill.com is a top priority. Management is responding by a method coined by Matsko as "travel and conquer." The method involves traveling the country and providing training seminars for small businesses to help them develop a variety of commerce sites. According to Matsko, "our clients need to know us on a personal basis and understand how we can affordably help them make money with the Web" (C. Matsko, personal communication, November 5, 1999).

This approach is the mainstay of Anthill.com. It continues to bring thousands of individuals to the Anthill.com colony. Likewise, this personal interaction makes Anthill.com have the coziness of a small business as it works to expand into a large-scale e-business. Since Anthill.com competes in an industry immersed in the technological revolution, it must invest much of its profits in technology resources and people with technical skills. Top executives Curt Matsko and Scott Alexander are investing in the future. Matsko underscores this by stating that "our company realizes the importance of staying abreast of the rapidly changing technologies associated with the Internet" (C. Matsko, personal communication, November 5, 1999).

What started out as a relatively simple Web server has emerged into a complex array of active server pages, Java, Pearl, and Netscape's SSL technologies. This change drastically increased the need to be able to attract technically competent employees. Recruiting the best technical minds is becoming more difficult as the competition for good technical people is at its highest in years in the information systems industry. Anthill.com executives spend much of their decision-making time devising strategies to attract talent. Matsko's plan is to recruit individuals with experience in both managerial as well as technical areas, as well as to develop unique business relationships with individuals who do not work directly for Anthill.com.

One unique strategy utilized by Anthill.com to improve business practices is the use of independent contractors and Anthill.com affiliates. Anthill.com currently has over 1000 active contractors (individuals who have purchased

a commerce site from Anthill.com) who sell Anthill.com commerce sites to other businesses and individuals for profit. In addition, Anthill.com affiliates are compensated for each new business they bring to Anthill.com. The affiliates use an identification number when adding a business site to Anthill.com to receive compensation. The use of independent contractors and Anthill.com affiliates permits Anthill.com to develop a formable workforce similar to a large organization.

Principle 4: Build external relationships to compensate for limited internal resources.

Netscape would have been unable to keep up with the demands of Internet time without outside help. The company had a powerful vision, experienced leaders, and an organization geared toward fast growth, but ultimate success depended critically on a wide variety of external resources and relationships. These external assets compensated for Netscape's lack of scale in marketing, financing, and product development. Netscape was essentially able to exploit the Internet and other external resources to create a virtual workforce—people outside the organization who were working for free on the company's behalf.

According to Matsko, "individuals we train are equipped to promote the colony though a variety of means such as registering with search engines and maintaining promotions via traditional media" (C. Matsko, personal communication, October 15, 1999). As a result, there are thousands of individuals around the United States who work to attract attention to Anthill.com. Other methods of attracting attention to Anthill.com include mailers, online promotions, and face-to-face communications. For example, teams of Anthill.com employees meet daily with individual businesses. Brad Thurber emphasizes this facet of company growth: "Face-to-face communication helps us to personalize the Internet and our company. It is a viable and successful approach to business" (B. Thurber, personal communication, October 15, 1999). Additional external relationships have been developed with a number of Internet companies such as BedandBreakfast.com, LotteryUSA.com, Mapquest.com, CardSercicesInternational.com, sisna, FilmFrinzey.com, Astrologynet.net, Travel.com, and Barchart.com. These relationships help to enhance the quality services offered by Anthill.com without requiring additional resources.

According to Thurber, "we plan to develop additional relationships with additional Internet organizations, to help improve our site within our current budget constraints" (B.Thurber, personal communication, October 15, 1999).

Case Study: Fannie Mae

Fannie Mae is an example of how a once successful company got away from a smart data approach, with a strategy for change and improvement, and ended up as a prime contributor to the Wall Street fiasco of 2008, and how the adoption of a SOE model could make them successful. We postulate that the highest order activity shared by all enterprises independent of organization type and role is to "optimize performance." We also postulate that enterprise data is a most treasured asset, and that strategy

focused on improving enterprise data is of equally high importance and aligned with enterprise performance optimization.

Generally, smart data is the product of engineering such that it presents itself to the right users at the right place and time in support of planning, problem solving, decision making, sense making, and predicting. At one time, Fannie Mae embraced these truths and was successful. In the near term, Fannie Mae ignored these core concepts of smart data as well as the methodology and algorithms that support the intended uses of smart data.

As a result, Fannie Mae has strayed from the smart data path and has become a liability as a government bailout. Fannie Mae forgot that smart data is interoperable and supports qualified users with diverse needs in an ubiquitous environment. It refused to follow a paradigm that produced advanced data engineering technologies, including modeling and metadata management, and smart application of open standards that account for credentialing and privileging as a dimension of security.

In a nutshell, this is why Fannie Mae finds itself in financial woes today. The following case was conducted in 1999 and shows Fannie Mae adopting many smart data concepts. Compare this environment with the one that Fannie Mae finds itself floundering in, a decade later, by ignoring the smart data paradigm.

Much like "Where's Waldo," can you point out the numerous smart data concepts that Fannie Mae ignored in the following case, Managing Radical Transformation at Fannie Mae: A Holistic Paradigm? [22] In this case, from 1999, we noted that change management is a critical issue in the current fast-paced business environment. Organizations are being bombarded with global business change, innovations in communications, and rapidly evolving information systems capabilities. Since there exists a paucity of rigorous and relevant literature on change management, we chose to embark on an in-depth case study to explore how one organization manages radical change. We developed a set of theoretical propositions to guide the effort and act as a theoretical lens.

Our specific research question is: How does an organization successfully manage projects that call for dramatic change to the way one normally conducts business? The case study allowed us to test the propositions in the business arena with one organization. It is hoped that this approach will offer insights and mechanisms to help other organizations effectively deal with change.

When we embarked on this study, we wanted to explore how organizations successfully manage dramatic change. We began by consulting the literature on transformation, BPR, and sociotechnical management. The literature helped us articulate five propositions relating to successful change management. We conducted an in-depth case study of one organization (Fannie Mae) involved in radical transformation to see how it coped. We found that each of the five propositions paralleled the case study in some way.

The case study provided a rich set of data that allowed us to delve deeper than the literature. We were able to show that these propositions are in fact applied in at least one case. We were also able to show that holistic management can increase information sharing and knowledge creation. Fannie Mae approaches training and development in a holistic manner. Top management is active and committed.

Resources are allocated. People are the central focus. Finally, every training course is aligned with the business it serves and education is customized for each person to maximize potential.

Each of the propositions is supported by the case. However, analysis of the data revealed more depth than what is currently in the literature. This study can be extended by modifying or rethinking the propositions based on what was uncovered by the case. The first proposition states that a systematic methodology will facilitate change efforts.

Fannie Mae uses the University Model to develop people, align classes to business objectives on the job, and enable equitable evaluation based on training. The model is also used to keep technical training flexible. A natural extension of this study is to explore the impact of these human and technical factors on change management.

People development concentrates on two dimensions—human and technical. Human development encompasses behaviors, personality traits, and attitudes. Technical development encompasses skills, knowledge, and competencies. For real change to occur, Fannie Mae has surmised that people must understand why change is important to the business. Hence, IT training is not limited to basic C++ or SYBASE. A curriculum is designed with a Fannie Mae slant; that is, it reflects what people will do back on the job. It also includes conflict resolution, creativity, communication, and teamwork training to develop the "soft" side of its technical people.

Personality tests are given to people to help customize training to the individual. Fannie Mae also believes that behavior can be shaped if it hires people with the right attitude. Technical development focuses on skills, knowledge, and competencies. Skill and knowledge development is pretty straightforward because the model helps match training with business. The second proposition states that top management support is critical to successful change. The case supports this proposition.

The CIO is the champion and visionary of training transformation at Fannie Mae. He hired Gus to help him implement his vision. The president is supportive of the University Model and budgets generously for IT training (approximately $7 million per year). Two other important factors are risk-taking and mistakes. Top management is happy if training hits the mark 20% of the time because of dramatic increases in people productivity. This means that 80% of the time mistakes are made. Allowing mistakes as a natural part of the learning process encourages risk-taking and innovation. Researchers can augment this study by gathering more data about each factor.

The third proposition states that a strategy-driven approach will facilitate transformation. Transformation at Fannie Mae focuses on its training paradigm because it believes that people are the key to success. The idea is to train information system (IS) professionals so that they can deliver whatever work, services, and products the business needs. Training works with business managers to develop a curriculum that is strategically aligned with the business needs of the enterprise. Exploration of training–business partnerships and strategic alignment of the curriculum can further transformation research.

The fourth proposition states that holistic management will facilitate change. The case brings out three factors that make up holistic management—equitable assessment, customized training, and adoption of a holistic approach. Holistic assessment is when everyone is evaluated based on the contribution to the enterprise. It is not subjective or trivial. It is performance based. Evaluation at Fannie Mae is based on the customized training a person receives, how the person uses it to deliver what the business needs, and best practices.

A group of managers and peers decide on what the best practices are for each job. Evaluation forms are uniform (except for management evaluation), which helps people perceive them as fair. Training is customized to the individual and what he/she must perform back on the job. Finally, Fannie Mae embraces a holistic management approach to IT training. The philosophy is top–down. Top management communicates to people that business value and understanding is important for everyone. People are rewarded for enterprise thinking, innovation, and creativity. The University Model guides the design of an aligned curriculum with business needs across functional areas and the enterprise. Human development is at the center of the philosophy. People do the work and can therefore make or break a holistic approach. The fifth proposition states that knowledge is created and retained by people.

Fannie Mae is "betting the farm" on this philosophy. Top management invests millions of dollars in people development. The goal is to create an environment that rewards value and encourages risk-taking. It is based on the notion that people are the conduit of information sharing. They create information and knowledge, pass it on to others, and retain what they believe is useful to them. "Knowledge is a tricky thing. It is not tangible. The inferential engine inside [a person's] head generates what we want. We don't know how it works, but we know that as we develop and challenge our people, information sharing and knowledge creation dramatically increase" (G. Crosetto, personal communication, October 29, 1999). Even though Fannie Mae is totally committed to the idea that people retention translates into knowledge creation and retention, it puts a lot of pressure on management. "It is much harder to manage autonomy than merely telling someone what to do. We have to allow people the freedom to be creative, but we have to make sure that what they create is valuable" (E. McWilliams, personal communication, October 29, 1999). "We are really managing chaos. Controls are minimal as compared to an autocracy. The time and effort we put into developing the people system is daunting, but the benefits are amazing" (G. Crosetto, personal communication, October 29, 1999).

Failure to adapt to a changing financial climate and complacency toward adopting a SOE model led Fannie Mae down the road to destruction and placed it among one of the major contributors to the 2008 Wall Street collapse fiasco. Surveying government and commercial enterprise customers for a number of years, and from having been directly engaged in related topics for more than 15 years, the authors have observed the following tendencies, in companies such as Fannie Mae, which have contributed to their downfall. First, these companies have deficiencies and omissions from a government and commercial data strategy. Second, the companies make advances in commercial enabling technology with gaps in application and implementation. Third, while there is continuous investment in information technology on a large scale, the

increasing pressure on government and private enterprise for better use of scarce resources, as invested in information technology, can be improved by commercial off-the-shelf technologies that can accelerate adoption of smart data strategy. Finally, the companies forget that there are requirements for management and technical training, in order to keep the enterprise competitive.

Now, let us see how companies such as Honeywell and IBM apply these general considerations by developing a strategy that will address the following issues. It is generally accepted that "integrating" the elements of an enterprise is a good thing. Honeywell has integrated people with change and used smart data to accomplish this. It has produced corresponding benefits by developing process maps and continuous improvement. It is generally recognized that "information," whereby information = data facts + meaning, and information + context = knowledge, is an asset that can sometimes become noise. Therefore question everything and demand team ownership. It is generally accepted that information technology—infrastructure and software—has evolved into something that is useful and essential while costly, unwieldy, and unmanageable.

Honeywell realizes that IT is a necessary, but not a sufficient, enabler. It is generally accepted that "interoperability" among data, processes, interfaces, applications, taxonomies, policies, and social networks is highly desirable, including the latest semantic interoperability. Therefore Honeywell realizes that execution is the real difference between success and failure.

Case Study: Honeywell Business Process Reengineering (BPR)

We compare Fannie Mae to another example of an organization that embraced the smart data, SOE paradigm: Honeywell. In the case entitled "A BPR Case Study at Honeywell" [23], we embarked on a case study to explore one organization's experiences with radical change for the purpose of uncovering how it achieved success. The organization we examined was Honeywell Inc. in Phoenix, Arizona. From the interview data, we were able to devise a set of 10 lessons to help others transform successfully. Two important lessons stand out above the rest. First, execution of a carefully developed change plan separates the high performers from less successful BPR projects. Second, recognition that dealing with change is difficult and complicated is not enough. Top management should make change management a top priority and communicate the change vision across the organization.

From the case study, we developed a set of general lessons. The case experience allowed us to speak in-depth with people involved in enterprise transformation, which should make the lessons more practical.

Lesson 1: People are the key enablers of change.

Business processes are complex, but process mapping offers a comprehensive blueprint of the existing state. The blueprint enables systematic identification of opportunities for improvement. IT is complex, but vendors, consultants, and system designers can create models of the system. In contrast, people are unpredictable. They cannot be modeled or categorized universally.

However, people do the work and therefore must be trained, facilitated, and nurtured.

Lesson 2: Question everything.

Allowing people to question the way things are done is imperative to change. Fail-safing provides a systematic approach to effectively question the status quo. People are encouraged to question the existing state.

Lesson 3: People need a systematic methodology to map processes.

Process mapping is the mechanism used to map and understand complex business processes. The systematic nature of the process mapping methodology keeps people focused and acts as a rallying point. Moreover, process mapping provides a common language for everyone involved in the project.

Lesson 4: Create team ownership and a culture of dissatisfaction.

Once a team perceives that they "own" a project, they tend to want to make it work. It becomes "their" project. In addition, management should encourage people to be dissatisfied with the way things are currently done. However, punishing people for complaining about ineffective work processes is an effective way to promote the status quo.

Lesson 5: Management attitude and behavior can squash projects.

If the managerial attitude remains that of "command and control" and/or management's behavior does not change, transformation will most likely fail. Success depends on facilitative management and visible and continuous support from the top. When Honeywell got its new president in 1996, the attitude toward criticism changed dramatically. The new president was not as accepting of casual criticism. Criticism of the status quo had to be based on well-thought-out ideas and presented with the logic behind the thinking. This drastically reduced the complaints about existing processes without justification.

Lesson 6: Bottom–up or empowered implementation is important.

While support from the top is critical, actual implementation should be carried out from the bottom–up. The idea of empowerment is to push decisions down to where the work is actually done. Process mapping and fail-safing are two systematic and proven methodologies that help support empowered teams.

Lesson 7: BPR must be business-driven and continuous.

Process improvements should be aligned with business objectives. Process mapping, fail-safing, and teaming should be based on what the business needs to change to become more successful. In this case, effective communication of ideas from top management throughout the enterprise is imperative. In addition, organizations should be wary of the "I've arrived" syndrome. Change is continuous and is never over.

Lesson 8: IT is a necessary, but not a sufficient, enabler.

IT is not a panacea. IT enables BPR by automating redesigned processes. However, information is for people. People work with people to produce products for other people. In addition, people need quick and easy access to quality information to help them make good decisions. Therefore IT needs to be designed to support the business and the production of products to be effective.

Lesson 9: Set stretch goals.

Goals should be set a little higher than what the team believes they can accomplish. Since teams have little experience with the new paradigm, goal setting will tend to be based on the past. Project managers should work with the team to help them develop stretch goals.

Lesson 10: Execution is the real difference between success and failure.

The Honeywell case introduces four powerful mechanisms to facilitate enterprise change. However, real change will not happen without a plan for change and aggressive execution of that plan. We believe this is where most organizations fail. We believe that execution fails in many cases because organizations are not willing to dedicate resources, time, and energy to the effort.

Structural Elements

Three structural elements are shown in Figure 1.4. These elements comprise a stable, technology-independent, structural description of the architecture. They represent a technology neutral description of the attributes or essential features that an integrated enterprise will exhibit. They do not describe an end state, that is, "do this and you will have an integrated enterprise."

As structural elements, they fall into the category of architectural features: each having associated design characteristics; each having varying, measurable degrees of performance, which, in implementation, can be traded off against other enterprise design and performance criteria (i.e., software system engineering).

Smart Data

Smart data is data that has been invested with explicit semantic content through formalization of metadata (by definition, characterization or modeling process). The term is intended to be broad in that "smartness" is a measurable quantity by degree, not a state of being. In other words, there are degrees of rigor, precision, accuracy, structure, and abstraction to which data can be formally described. It is intentionally broad in the sense that it applies to "data," being the grist that is crunched by applications, and to metadata, models (data, business process, and others), metamodels, and even SOA contracts.

Smart data is the product of a rigorous and published process that describes actionable information flowing across the enterprise. Actionable information includes data that is descriptive of events, phenomena, materials, processes, procedures, actions, applications, structures, relationships, or the data itself. The more structured the descriptive process, the more that data is reduced to descriptive relationships among fundamental abstract elements and the smarter the data.

Smart Grid

Smart grid is the physical infrastructure and protocol routine for intercommunication within the enterprise. Smart grid differs from a conventional interconnect structure in

that it includes substantial pre-engineering of enterprise interfaces. The smart grid technology of today corresponds to SOA implemented through shared language such as web service protocols. This is in contrast to point-to-point or data hub architectures. These latter configurations should be considered dumb grids because they require substantial engineering of specific enterprise interfaces. Smart grid is technology neutral and defined independently from SOA or web services implementation. Characteristic of smart grid are:

- Shared interconnection network architecture with common entry and messaging methods
- Message management capability to ensure reliable delivery of data from originating source system to intended receptor system with notification of failure and recovery
- Information assurance controls that prevent intentional or unintentional corruption of the enterprise communication process
- Sufficient resources in the form of directories, routers, and so on to support the interconnectivity requirements of the enterprise

Smart Services

Smart services are synonymous with semantic services that are shared resources, configured as web services, which are considered assets available to the enterprise as a whole, regardless of physical ownership or economic model. Enterprise global repository (EGR) is essential to the SOE. EGR is a structured resource that provides access to metadata, process and data models, metamodels, and other constructs that describe the enterprise, including processes and data. The EGR is a build-time resource from which system developers access existing models for use in formalizing their own data and to which they post their completed products. The EGR is also a run-time resource, accessed as a web service, from which semantic translators access data models and the active translation of data from system to system takes place.

Three Engineering Disciplines

Three disciplines interact and underlie the structural elements. Enterprise integration (EI) technology gains will provide increasingly robust capabilities to implement the specific solutions within the SOE structure. As EI technology progresses, increasingly robust capabilities will be installed into the EI environments. Simultaneously, the structural element provides the conceptual background for technological advancement.

Data Engineering Data engineering is the practice of developing and documenting semantic content for enterprise data throughout the enterprise life cycle. It is an ongoing discipline that addresses new types or forms of data whenever in the life of the enterprise they are introduced. By investing in data engineering, the enterprise

will see a reduction in data conflicts in enterprise operations. In addition, it will enable more cost efficient data reconciliation processes and products, which in turn can extend the "reach" of the enterprise into increasingly marginal (low-volume) applications.

Basic to the data engineering discipline is the need to semantically characterize all classes of data relevant to the enterprise (all data that is passed system to system) through a formal process. This process is practiced through various methods of formalization of metadata: for example, definition of data elements, information modeling, and metamodeling. It can be practiced at four levels.

1. **Unmanaged, Ad Hoc.** Data is defined locally on the fly.
2. **Systematic Data Definitions.** Data is defined on an element-by-element basis. This is basically the conduct of a data inventory, which is the necessary first step to higher levels of data engineering. In a bottom–up fashion, data elements are rigorously defined using a prescriptive process such as the International 11179 Standard. Accuracy and synchronization with physical sources and instances of data are largely manually achieved, as is configuration control and governance supported by a variety of data repository types, such as databases and spreadsheets.
3. **Formalized Information Modeling.** Data accuracy and synchronization are achieved at a higher level of abstraction through information modeling of the underlying metadata. Information modeling more fully describes data/metadata by describing the relationships between data elements as well as defining the data elements themselves. This increases the semantic content of the data enabling the interoperability of data by means of semantic mediation engines. Configuration management at this level requires a more complex data repository to maintain the models as well as data element definitions.
4. **Metadata-Driven Information Integration.** A model-driven architecture prevails for the enterprise. Supported by both a design-time metadata repository and run-time integration engine, this level facilitates enterprise integration by application and system developers as well as during operations. Both are achieved through extensive data engineering in a top–down fashion such that metadata and information models are integrated through metamodeling and similar means.

Grid Engineering Grid engineering is the discipline that develops and evolves the smart grid architecture for an enterprise, consisting partially of adapting the generic SOA to specifics of the enterprise. This consists of selection of integration tools, processes, and standard protocols, such as the Web Services Description Language (WSDL) and Simple Object Access Protocol (SOAP). It also establishes the rules of engagement criteria for applications and systems to participate in the enterprise. It also prescribes the method for integrating legacy systems and applications into the enterprise.

With respect to the enterprise, selecting protocols and then broadcasting the selected protocols to system developers accomplishes grid engineering. However, grid engineering also entails the development and enforcement of an enterprise information assurance and security strategy. Future web service protocols that directly address this topic may simplify this task. Grid engineering also includes defining the service layers that connect enterprise systems to the grid. The attributes of a smart grid will not be attained through protocols alone. Every system that participates in the enterprise must provide functional capability as structured layers of services.

Grid engineering can be practiced at one of four levels.

1. **Unmanaged, Ad Hoc.** Legacy applications and corresponding business intelligence are integrated into the enterprise locally on the fly.
2. **Encapsulation and Objectification.** Legacy applications are adapted to a process of encapsulation and objectification. A service-oriented adaptor stack provides the "face to the enterprise." This stack includes translation and security engines. Modeling of information and process steps objectifies data and processes. Objectified data is maintained locally to the application. Configuration management is exerted locally mostly by application developers.
3. **Capture of Business Intelligence.** Business intelligence is captured in local applications and systems in a comprehensive fashion. A local content repository is used to store and manage business intelligence. Common local functions, such as troubleshooting and recovery, are selectively exported to the enterprise. Configuration management is exercised at the system level by system integrators.
4. **Externalization of Business Intelligence.** Business intelligence is captured out of local applications and systems in reference to enterprise level models and forms and exported to an enterprise repository. This enables asset mining by enterprise users. Through enterprise modeling, all business intelligence is derived from comprehensive unified models, resulting in a model-driven architecture. Configuration management is exercised at the enterprise level.

Process Engineering Process engineering is the practice of designing and documenting enterprise processes. Done properly, it enables process improvement while maintaining enterprise operations. At its most basic level, it involves development of the rules of interaction for the processes that comprise the enterprise and their enforcement. These rules of interaction, which apply to processes as well as data, are similar to the business process rules of today, but only apply to process characteristics that impact the execution of other processes within the enterprise.

Rules that constrain the internal operations of a business process application are excluded from the category of rules of interaction. Rules of interaction deal with formality in describing the outcomes of processes, not constraining how they work internally. Their purpose is to provide a human and machine interpretable characterization of what the process does so that other process designers can accurately anticipate the result as they design their process.

Process engineering can be practiced at one of four levels.

1. **Unmanaged, Ad Hoc.** Processes, legacy and otherwise, are integrated into enterprise business operations locally on an ad hoc basis.
2. **Business Rule Standardization.** Processes are integrated into the enterprise business operations through business rules that tightly constrain the manner in which the process is built within application software. Business rules are stored in various designer accessible repositories, registries, and spreadsheets. Accuracy and synchronization are issues that have to be worked out on an interface-by-interface basis. Manual governance (through meticulous checking of design rules) and manual configuration management prevail.
3. **Process Modeling.** Enterprise business processes are modeled with the process models stored on a generally accessible repository. Application developers generally post and subscribe to the repository. Process integration largely takes place at design-time and configuration management is exercised at the system level.
4. **Outcome-Driven Processes.** Enterprise business processes are modeled abstractly using enterprise level models as reference points. The enterprise business process repository enables process integration through orchestration or choreography engines at run-time as well as through design-time methods. Through enterprise modeling, all business processes are derived from comprehensive unified models, resulting in a model-driven architecture. Configuration management is exercised at the enterprise level.

Enterprise Integration

Enterprise integration has progressed from point-to-point connectivity, data hub connectivity, service-oriented architecture, enterprise service bus, and enterprise information interoperability manifesting in the service-oriented enterprise. Figure 1.5 also illustrates the combining of net-centric enterprise services and semantic interoperability aggregating into SOE, the complete paradigm.

Your world is full of artifacts representing each of these paradigms: some point-to-point, some hub-and-spoke, and some implementations toward SOE. Unfortunately, progress is a mishmash. Point-to-point is much more costly than hub-and-spoke, and hub-and-spoke architectures are more costly than SOE. The combination of conflicting legacy causes the enterprise to lope toward a state of entropy.

The U.S. government addresses enterprise-scale integration and performance optimization with different programs and with attempts at sharing knowledge, although the U.S. government is far from having a cohesive strategy that integrates the thinking and powers of three branches of government. New leadership must recognize the need for a common strategy and focus on America's public resource management that produces a common result—a secure and prosperous nation.

- Embodiment into functions
- Embodiment into outcomes
- Embodiment into things

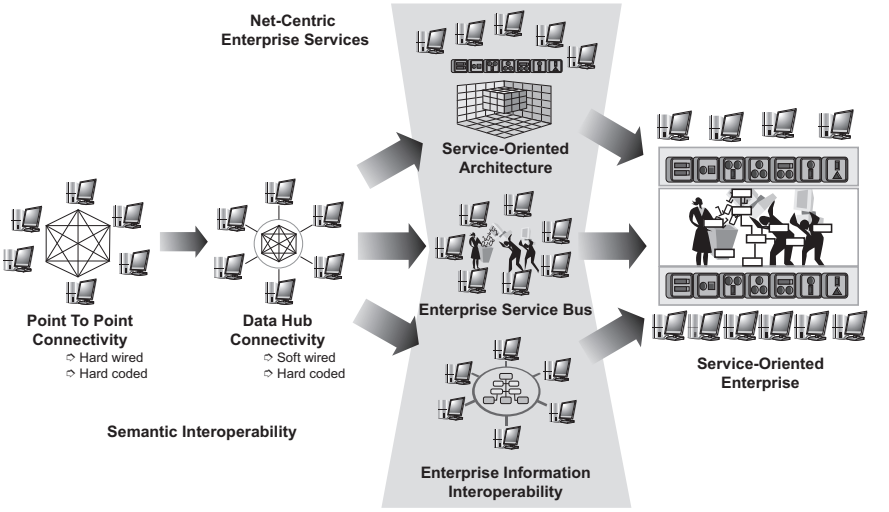

Figure 1.5 Evolution of enterprise integration strategies.

U.S. Army LTG General Jeffrey Sorenson, G-6 CIO, described the current system of systems strategy manifest in a program called Future Combat Systems (FCS) with the intention to produce a lighter weight vehicle that embodies all of the open interoperability characteristics with the ability to deliver line-of-site and beyond-line-of-site lethality to the enemy. According to Boeing, the prime contractor, "the Future Combat Systems (FCS) program is an Army modernization initiative designed to link soldiers to a wide range of weapons, sensors, and information systems by means of a mobile ad hoc network architecture that will enable unprecedented levels of joint interoperability, shared situational awareness and the ability to execute highly synchronized mission operations."

According to Wikipedia, "system of systems is a moniker for a collection of task-oriented or dedicated systems that pool their resources and capabilities together to obtain a new, more complex, ''meta-system' which offers more functionality and performance than simply the sum of the constituent systems. Currently, systems of systems is a critical research discipline for which frames of reference, thought processes, quantitative analysis, tools, and design methods are incomplete. The methodology for defining, abstracting, modeling, and analyzing system of systems problems is typically referred to as system of systems engineering."

System of systems (SOS) is different from service-oriented enterprise (SOE) in that the context for SOE is, from beginning to end, the enterprise. SOS presumes standalone or ad hoc systems being brought together, leveraging metasystem characteristics to produce something stronger than the parts. However, we might argue that better results might come from end state designs that begin and end in the enterprise context.

"The U.S. Army could spend more than $300 billion to purchase and run its Future Combat Systems (FCS) family of armored vehicles, drones and communications networks during their expected multidecade life, according to a cost estimate prepared

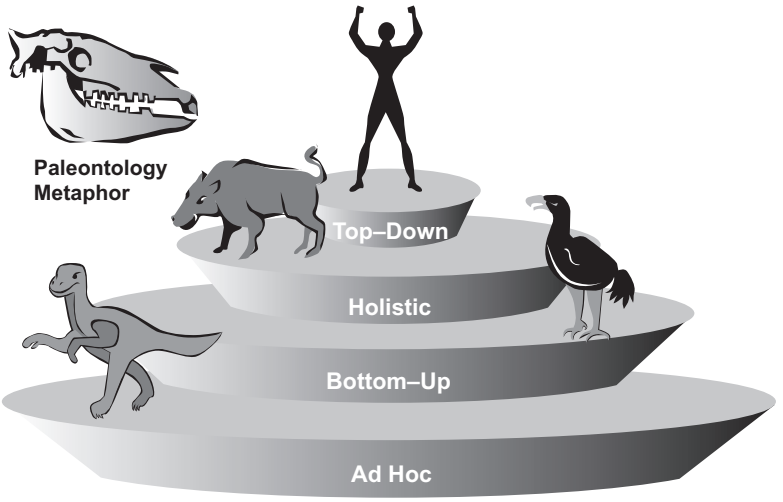

Figure 1.6 Levels of enterprise engineering.

by the U.S. defense secretary's Cost Analysis Improvement Group (CAIG)." Already, the program is woefully behind schedule and ahead in cost.

Figure 1.6 illustrates four progressions beginning with ad hoc, bottom–up progressing to holistic, and finally top–down. The references begin with dinosaurs and birds to warm-blooded mammals and finally humankind. The advance of animals from dinosaur reptiles, birds, four-legged mammals, to humankind paralleling ad hoc, bottom–up, holistic, and top–down remains an interesting metaphor. That's our sense of humor.

1.5 BREAKING FROM TRADITION WITH IMPROVED RESULTS

To break from tradition, executives must do some things differently in their pursuit of enterprise performance optimization. Ever present is possession of high value for data.

CEO Smart Data Strategy Work Breakdown Structure

What specifically are executives to do?

[A0] Optimize Enterprise Performance.

At what point in the order of business do CEOs or department secretaries address data strategy and attention to enterprise data? Observing the incoming Obama administration, we witnessed (1) attention to data details and (2) attention to staffing executive positions with those who understand processes and rules and can interpret data.

Propelled by events, incoming executives must grasp an operational environment that is fraught with problems and catastrophes. They must address strategy, tactics, and operational management issues in parallel, doing so with constrained bandwidth and capacity for change and improvement.

The national economy is collapsing. The nation is at war on two or more fronts. U.S. auto manufacturers are unprofitable and sales trend downward as products are uncompetitive. These issues are nontrivial, so how should executives factor attention to data strategy?

The CEO or president provides a brand of management and a set of guiding values and principles that are a part of what we call the *leadership and integration architecture*, which IT professionals might call the *control architecture*.

The leadership and integration architecture is a complete management system that contains process activities, inputs, controls, outputs, and enabling mechanisms needed by management to manage.

Did Obama have time to formally establish the leadership and integration architecture? He had a running start, but until his subordinate team is approved and on the job, it really can't be completed. Herein lays a fundamental flaw in our government process. One can attempt to manage intuitively, projecting personal charisma and the like, but that will only go so far in creating an illusion of management. Tim Geithner is an example of a consequence of being overwhelmed with too much, too fast. Look around at other departments and you will see DoD is in better shape because it has a carryover executive.

Part of the enterprise management system is structurally intact in mature enterprises, government, and commercial industry. It is inherited, though subject to change and improvement. Changing the structure requires deliberate planning, as often changing without a plan may result in a condition worse than before the change.

Therefore having a management approach is essential to undertaking the enterprise leadership position. Sooner than later, it is time to address the responsibilities strategically while managing day-to-day tactically.

[A1] Define Enterprise Outcomes.

Defining enterprise outcomes is a top priority as all resources and processes are committed to producing them.

Outcomes may be classified into the following types:

- Plans
- Results

  Problems solved

  Needs satisfied

  Opportunities exploited

  Products produced

  Services delivered

For our purposes, a plan contains a series of steps or milestones for guiding achievement accompanied by a schedule and resource allocations. How are plans different from processes and activities?

Enterprise processes and their component activities are recurring capabilities that may be engaged or applied to achieving plans. Plans are intended to address the production of results such as problems solved, needs satisfied, opportunities exploited, and products produced—all with associated measures and metrics.

Plans include attributes such as volume, rate, frequency, milestones, labor allocations, burn rates, and other resource allocations and utilizations resulting in costs, and accounting for assets, and value production.

Data are input to plans as well as input to processes. Data are outputs of plans and processes. Data are evidence of accomplishment.

[A2] Define Enterprise Controls.

We prefer calling enterprise controls the *enterprise leadership and integration architecture*. It accounts for management processes and their associated controls that permeate the organization as controls on subordinate processes. To a large extent, controls shape the brand of management.

As stated before, controls on government enterprise contain all of the laws and regulations as well as policies and guidances that may appear as memos and other communiqués such as visionary plans and programs.

One definition of *program* is "a system of projects or services intended to meet a public need." That is a complex idea. First, the existence of a program represents management's decision to commit considerable resources to produce results from an effort of considerable scope and scale. Therefore identifying a program as a part of a plan might constitute a statement of leadership intention, commitment, and control as it defines budget allocation and time constraints.

A strategy may be executed as a program whereby to accomplish it requires special resource commitment and management focus.

Smart data strategy would manifest programmatically by the following:

1. Executive policy statement emphasizing the importance of operational entities having a data strategy that contributes explicitly to the enterprise data strategy as defined in the policy.
2. Using smart data strategy as a catalyst for change and improvement as an enterprise initiative.
3. Adding and implementing smart data strategy enablement including adding new skill sets and new technologies.

[A3] Identify and Define Enterprise Inputs.

Enterprise inputs include accounting for all resources, capital, and materials that are consumed and transformed by the enterprise into intended outcomes. Under the smart data strategy approach, all resources would be accounted for by tracking their use and transformation through enterprise processes, concluding with classification into assets, results, cost, waste, and noise.

The activity of accounting for resource use, transformation, and results requires explicit data tracking and management. Refined data tracking can happen in the background, although at any time, information can be made explicit to support monitoring, planning, problem solving, decision making, sense making, and predicting.

Smart data strategy is specific in anticipating what data will be needed, when, where, and for what purposes, as well as who will have access based on credentialing and privileging.

[A4] Define Enterprise Processes.

Enterprise processes are modeled for all core functions. They are attributed to the lowest level necessary with respect for accounting for inputs, controls, outputs, and mechanisms.

[A5] Attribute Enterprise Processes with Enabling Mechanisms People and Technology.

Attributing enterprise processes is a deliberate and precise effort to assign people and technology enablement to accomplish the work needed to produce desired and required outcomes.

Herein lays a great opportunity. Improving processes and improving data quality (i.e., making data smarter) will improve automation and provide significant opportunity to reduce the requirement for large numbers of analysts. In addition, smart data will provide executives with the opportunity to refine resource deployment with greater precision.

[A6] Define Enterprise Metrics and Feedback.

Information measures are what are being measured and metrics are the specific units of measure and evaluation. All process metrics are accumulated in an enterprise performance repository, which is the basis for providing planned feedback to management and for further analysis.

Applying the Integrated Definition (IDEF) modeling technique, the enterprise management optimization process appears as in Figure 1.7. Omitted for simplicity are such things as capital and material inputs.

For enterprise management, there is typically a biannual annual (on-year and off-year) cycle. In the U.S. government, planning, budgeting, and funding cycles extend for multiple years with overlying processes and cycles that prescribe certain management products.

Since many commercial enterprises are dependent on government customers as prime contractors, subcontractors, and suppliers, their internal processes mirror their customers.

In government and commercial industry, certain management products are required as imposed by laws, regulations, or requirements from sources of capital. Requirements vary depending on organization type. For instance, public companies must comply with Sarbanes Oxley. Nonprofit organizations must comply with regulations governing their type while government organizations have federal requirements. Private enterprises may be thought to have more latitude, but they are constrained by sources of capital and by laws

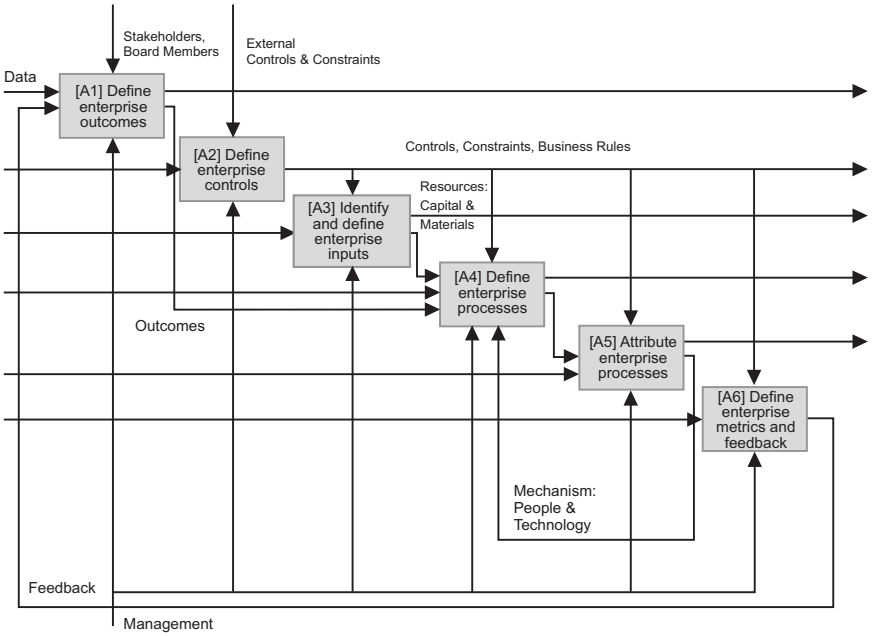

Figure 1.7 Optimize enterprise performance management process [A0].

for taxation, for instance. In a competitive environment, certifications emerge as self-imposed and imposed controls.

Where is the data in this diagram of management performance optimization **?** Data is input to and output from all processes. Data represents or measurably describes every control and every mechanism.

What data do executives need to determine to achieve the outcomes that the enterprise should produce? In an existing private enterprise, the nature of the business will suggest certain outcomes necessary to satisfy customer needs in return for which the business produces a profit. For a new business, the entrepreneur will anticipate customers and needs based on certain data from primary and secondary research. Even for an existing enterprise, primary and secondary research about customers and competitors is a continuous process.

For government enterprise, certain legislative givens identify outcomes, and, in addition, the executive branch of government working in concert with the legislative branch will generate new specifications for outcomes in response to constituent needs and wants.

In a large and complex enterprise, outcomes aggregate from subordinate processes. Outcomes specified by executives at the top of the enterprise stimulate organization responses at various levels and activity.

REFERENCES

1. Michael D. Shear and Anita Kumar, "Obama Picks Technology and Performance Officers," *Washington Post*, April 19, 2009.
2. National Institutes of Health, Glossary. http://science.education.nih.gov/supplements/nih4/technology/other/glossary.htm.
3. Max Stier, "Challenges for the New Chief Performance Officer," *Washington Post*, January 7, 2009.
4. Al Ries, "Metric Madness: The Answer to Mathematical Failure Seems to Be More Math. If You Run a Company by Numbers Alone, You'll Run It into the Ground," *Advertising Age*, May 4, 2009.
5. Jim Finkle, "IBM Plans Cloud Computing Services for 2009," Reuters, April 23, 2009.
6. Marcus Cake,"Applying Web 3.0 to Financial Markets and Economic Development," http://www.marcuscake.com/.
7. Bob Bragdon,"The Value of Data," *CSO Online*, April 1, 2008.
8. Michael C. Daconta, "Designing the Smart-Data Engine," http://74.125.95.132/search?q=cache:P34xt0K_BVgJ:web-services.gov/Designing%2520the%2520Smart-Data%2520Enterprise.doc + Designing + the + Smart-Data + Enterprise.&cd=1&hl=en&ct=clnk&gl=us.
9. Michael C. Daconta, Leo J. Obrst, and Kevin T. Smith, *The Semantic Web*, Wiley, 2003.
10. "AIG Annual Report 2006," http://www.ezonlinedocuments.com/aig/2007/annual/HTML1/aig_ar2006_0004.htm.
11. Tim Berners-Lee,"Worldwide Web Consortium," http://w3.org/Consortium/.
12. Larry Clinton,President, Internet Security Alliance, testimony before the House Subcommittee of Communications, Technology, and the Internet, May 1, 2009.
13. http://www.opensource.org.
14. http://www.ichnet.org/glossary.htm.
15. Sid Adelman, Larissa Moss, and Majid Abai, *Data Strategy*, Addison-Wesley, Reading, MA, 2005.
16. Steve Martin, *Picasso at Lapin Agile*, 40 Share Productions, 1996.
17. Terje Totland, Norwegian Institute of Science and Technology, Knowledge Systems Group, 1995 p. 1172.
18. Donald Hall, Wilbert Bailey, and Jim George, A Service-Oriented Enterprise Strategy for Enterprise Integration, *Business Integration Journal*, Vol 2, p 38–42, July/August 2005.
19. E. Altman, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy," *The Journal of Finance*, 23, 589–609, 1968.
20. J. A. Rodger, Ch. 9: Utilization of Data Mining Techniques To Detect and Predict Accounting Fraud: A Comparison of Neural Networks and Discriminant Analysis. In Parag C. Pendharkar (Ed.) *Managing Data Mining Technologies in Organizations: Techniques and Applications* Idea Group Publishing, Hershey, PA pp. 174–187, 2003.
21. D. B. Yoffee, and M.A. Cusumano, "Building a Company on Internet Time. Lessons from Netscape." *California Management Review*, **41**, 38–28, 1999.
22. D. J. Paper, R. Chang, and J. Rodger, "Managing Radical Transformation at FannieMae: A Holistic Paradigm." *TQM Magazine*, **14**(4), 475–789, 2003.

23. D. J. Paper, J. A. Rodger, and P.C. Pendharkar, "A BPR Case Study at Honeywell." *Business Process Management Journal*, **7**(2), 85–99, 2001.
24. Paul L. Francis, Director Acquisition and Sourcing, GAO Testimony Before the Subcommittee on Armed Services, House of Representatives, Defense Acquisitions Key Considerations for Planning Future Army Combat Systems, March 26, 2009.