Chapter 1

Semantic Interoperability Gives IT Meaning

KEY TAKEAWAYS

- Semantic interoperability is an enterprise capability derived from the application of special technologies that infer, relate, interpret, and classify the implicit meanings of digital content, which in turn drive business processes, enterprise knowledge, business rules, and enterprise application interoperability.
- The third wave indicates that competitive advantage for companies in all industries will be largely driven by efficiency in information sharing.
- Most strategic, long-term barriers to efficient information sharing are inadequately addressed by currently popular integration approaches.

Innovation in information technology (IT) is often represented as a series of waves. Going back to the 1950s, various technology waves appear with large amplitudes. These waves include the mainframe, the network, the mini-computer, the personal computer, and the Internet. Then there are technology waves with smaller amplitudes—the disk drive, the computer terminal, the relational database, the graphical user interface, the laser printer, markup languages, and component software.

Viewing technical innovation in this manner is useful, but its simplicity often obscures much of the real story. Rarely does just one technology make for a revolution. More often it is a combination of technologies that create a new platform that then creates fertile ground for an even larger set of innovations. Often, these technologies are not even new. Famous authors and musicians often joke that they became overnight successes—over the course of ten years. The same is true with technology. Almost all "breakthrough" technologies are typically preceded by years of research and prior iterations. It is typically the combination of technologies that creates the large-amplitude waves that set off a host of other advances. These advances come not only within technology but also in technical development with the advent of new design and engineering processes; in business with the development of new business models and new products and services; and—when technology touches the consumer—in social dynamics and interactions.

Adaptive Information, by Jeffrey T. Pollock and Ralph Hodgson
ISBN 0-471-48854-2 Copyright © 2004 John Wiley & Sons, Inc.

The most recent period of innovation, the Internet revolution, featured at its foundation the Internet with its IP protocol and network topology, which had been in existence for over 30 years. And although the HTTP protocol was a big step forward from other file and data transfer protocols, the use of the World Wide Web only really took off when low-cost web servers and free graphical browsers became available.

Even then, however, technical advancement was aided by a host of other enabling technologies—many of which were independent efforts that happened to be at the right time and in the right place. These technologies included innovative search algorithms that provided immediate utility to the web and advanced computer languages such as Java that enabled greater platform independence for component software. Additional progress came through of the use of rapid development techniques that fused information and interaction design closely with engineering and programming disciplines. The open source movement contributed to the pace of advances by providing thousands of innovators with the basic tool sets for working and creating new products and services. The Internet revolution was not one technology or one protocol, nor was it just about technology. New thinking was applied to almost all areas of business and many areas of social interaction. And although it is amusing to look back at all the talk about the "New Economy," business and society have been forever changed by a simple and yet revolutionary set of component technologies and protocols.

The personal computer revolution was a similar story. It started with a relatively simple and programmable operating system and extended to computer screens, disk drives, floppy drives, and laser printers. Novel consumer programs such as spreadsheets, word processors, and databases appeared, soon to be followed by desktop publishing. A vast number of companies sprang up to create and sell add-on hardware components that further extended PC capabilities. New distribution systems emerged, first in the form of small computer stores, then evolving into mail order, office supply stores, and computer megastores. Ethernet networks and file sharing were key components in making personal computers standard office equipment in the workplace and fueled the second and larger trajectory of the PC wave.

Whereas consumers' lives may not have changed as much as they did with the Internet, no one can ignore the change that personal computers brought about in the workplace. New business structures arose, and roles and responsibilities were drastically redefined—all of this because of a revolutionary platform that increased access to information and to information tools.

DOES IT MATTER ANYMORE?

The Harvard Business Review published a controversial article entitled "IT Doesn't Matter" in May 2003.[1] In this article Nicholas Carr made the bold claim that America's great IT industry no longer had any strategic value to offer corporations.

[1] *Does IT Matter? An HBR Debate.* Carr, Nicholas, June 2003.

Noting that the "data machines" of the past 50 years are limited to the core functions of "data storage, data processing, and data transport," he builds a case that software will no longer be a source of competitive advantage in business.

He is mostly right. That is, about current popular tools and capabilities.

Traditional integration tools (EAI, B2B, BPM), and packaged applications (ERP, CRM, PLM) for that matter, fall into Carr's category of "data machines." However, he misses the important point that software capabilities are a moving target—constantly evolving in response to changing business demands. As you will see throughout the course of this book, emerging technology approaches like semantic interoperability will drive new opportunity for business advantages by enabling companies to strengthen their own unique differentiators in the marketplace.

Because semantic technologies—technologies that enable data, process, and service meaning to become machine processable—offer a new paradigm for managing information inside the corporate infrastructure, not just offering new technical function points that can quickly become commodity, the application of these technologies will deliver highly tailored competitive advantages among companies who adopt them.

So, although Carr may feel that each technology offers the same advantages to companies who invest and adopt it, semantic technology is not just another me-too software application. *Rewiring the infrastructure that delivers information throughout the enterprise is sort of like enhancing the human body's central nervous system—it doesn't turn everybody into clones with the same capabilities, it enhances the characteristics of who you already were.*

However, a significant portion of Carr's thesis is dead-on accurate. Today, the vast majority of popular packaged software applications, off-the-shelf integration and middleware products, and information technology approaches are largely commodity products that offer incremental value to corporate consumers. Indeed, the IT industry had reached a point of diminishing returns because nothing was changing the rules of how technology worked. IT was simply about function points, interfaces, and custom coding requirements.

Actually, the whole notion of *information* technology was a misnomer to begin with; as Carr implies in his paper, we probably should have called it *data* technology!

Carr, like other skeptics, is standing at point A in Figure 1.1 and forecasting the inevitable commoditization of data processing technology,[2] which will lead to point B. However, the rise of new technologies and approaches has already resulted in new rules and a new curve—information processing. Emerging technology approaches discussed in this book, such as the application of semantic tools, technologies, and methodologies, will deliver us into the technology world beyond data—toward a real information technology industry.

[2] This S-curve example is similar to Clayton Christensen's (Innovator's Dilemma) concept of sustained technical change within a value net—in this case we are referring to the constant innovation in the IT industry as a whole, and suggesting that strategic competitive advantage remains possible by out-maneuvering competition and adopting new technology at the right time.

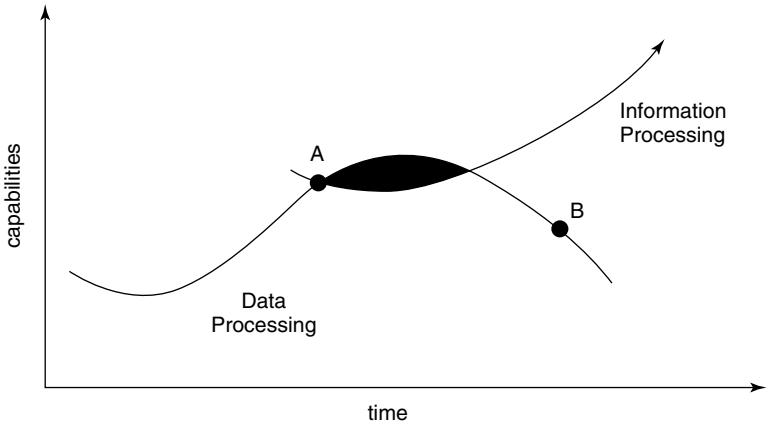

Figure 1.1 Sigmoid curve of IT processing capabilities

SEMANTIC INTEROPERABILITY: 15-YEAR OVERNIGHT SUCCESS

Semantic interoperability may initially sound downright foreign to some casual readers. However, it is grounded in long-established concepts of information theory and enterprise software practices. It will become the next 15-year overnight success in the information technology industry. The discipline of data semantics has broad and far-reaching applications in a number of software solutions. From work on the Semantic Web to emerging database structures, the concepts underlying the use of semantics in digital systems are having an immense impact on businesses industry wide.

Definition: Semantic Interoperability

Semantic interoperability is a dynamic enterprise capability derived from the application of special software technologies (such as reasoners, inference engines, ontologies, and models) that infer, relate, interpret, and classify the implicit meanings of digital content without human involvement—which in turn drive adaptive business processes, enterprise knowledge, business rules, and software application interoperability.

Semantic Interoperability vs. the Semantic Web

The term "semantic interoperability" is not interchangeable with the term "Semantic Web." Much of the work on the Semantic Web is focused on the ambitious goal of allowing relatively ubiquitous and autonomous understanding of information on the Internet. Semantic interoperability, on the other hand, represents a more limited or constrained subset of this goal. More immediate—and many would say suffi-

cient—returns can be gained by using semantics-based tools to arbitrate and mediate the structures, meanings, and contexts within relatively confined and well-understood domains for specific goals related to information sharing and information interoperability. In other words, semantic interoperability addresses a more discrete problem set with clearer defined end points.

It may be an overgeneralization to say that this problem subset fits the classic 80–20 Pareto principle, but nevertheless, the benefits in solving challenges within this subset are immediate and cannot be denied. A more apt differentiation between the two might be that semantic interoperability represents a 3- to 5-year vision whereas the Semantic Web represents a 5- to 10-year vision. The concepts surrounding semantic interoperability share many common terms, technologies, and ideas with the Semantic Web, but they also contain principles that have been developed and proven separately from Semantic Web initiatives. As semantic interoperability gains visibility, though, it is quite likely that there will be further blending of the concepts, protocols, and tools of the two.

Insider Insight—Stephen Hendrick on Why Semantics Now

Semantics are important because they help us cope with diversity. Diversity runs rampant in the Information Technology (IT) industry where new languages and standards are continually being introduced but rarely retired. Until recently, most vendors and organizations weren't concerned with this mushrooming diversity. However, as IT has progressively crept into most dimensions of our lives, the software asset base has expanded rapidly creating a software complexity crisis due to the large number of disparate standards, applications, languages, and datasets that were created without an eye toward interoperability and integration. The burden created by the software complexity crisis now means that most organizations spend more time maintaining and integrating existing applications than building or implementing new applications.

XML and Web services are considered an important first step in addressing interoperability. While there is no doubt that these technologies and their descendants are constructs that will aid in solving the software complexity crisis, their existence has simply brought us face to face, and far sooner than we expected, with the hard problems posed by diversity and the challenges of integrating disparate applications.

Addressing and resolving issues of diversity and integration involve building meaningful relationships between objects. Building relationships applies to both data objects and process objects. The value in building these relationships is the ability to coherently link together all aspects of the value chain that previously were isolated by virtue of their age, architecture, or implementation. The need for more structure in IT development tools is a phenomenon that has been taking shape for two decades. The widespread adoption of relational and object-relational databases as well as the pervasive shift to object-oriented programming languages specifically identifies the importance of being able to define intricate object-based relationships and hierarchies. Recent industry experience with objects and relationships is however simply a prerequisite for addressing the far more complex task of building granular relationships between complex objects comprised of many heterogeneous databases and systems. The quid pro quo is that our legacy

Continued

of widely disparate heterogeneous software assets in their current form can only be transformed into a coherent integrated array by first establishing what relationships exist in a semantically structured way. However, the power which results from a large scale unification of IT assets—first intra-company and then inter-company—will revolutionize IT and expand B2B and B2C relationships by several orders of magnitude. As this unification continues, it will dramatically simplify the way each of us relates to this far more expansive and semantically rich IT environment, laying an extensible foundation for building even better semantically structured relationships in the future.

Stephen D. Hendrick
Vice President, Application Development and Deployment
International Data Corporation (IDC)

Differing Uses of the Term "Interoperability"

The use of the term "interoperability" has become increasingly popular over the last several years, and its seemingly sudden appearance in the nomenclature has led to several differing uses and definitions. *Interoperability at its base level means using loosely coupled approaches to share or broker software resources while preserving the integrity and native state of each entity and each data set.* Common variants in use include data interoperability, Web Services interoperability, and information interoperability.

The term "data interoperability" is commonly used in conjunction with the communications frameworks used for connecting differing sets of emergency workers, such as first-responders. First responders are typically defined as the public service agencies—specifically local police, fire, and emergency medical, but also including federal and state law enforcement, emergency management, public health, National Guard, hazardous materials, search and rescue, and other specialized disaster recovery agencies or units. Although the information discussed in this book will have some relevance to this issue, specifically with respect to the access to various data sources, the primary problem set is one of mediating between data spectra and radio and data communication protocols and not necessarily addressing semantic conflicts.

"Web Services interoperability" primarily refers to ensuring compatibility between various Web Services protocols and frameworks. The rapid pace of innovation in the Web Services space reflects the efforts of companies to gain a competitive advantage over others which means that the ubiquity promised by web service interactions is under constant risk. As with data interoperability, this book may have some relevance but the issue regarding transport interoperability is more one of market forces as well as the correct implementation of web service specifications such as SOAP, WSDL, UDDI, XML, and XML Schema.

Information interoperability is the core subject of this book: Data alone is insufficient for successful software application interoperability—semantics turn data into information, and, in turn, the interoperability of semantics enables widespread inter-

operability of software components, processes, and data. At its core, information interoperability is the ability to address and mediate the logical nature and context of the information being exchanged, while allowing for maximum independence between communicating parties. This means addressing syntactic, structural, and semantic differences within structured, semistructured, and unstructured data. The larger definition encompasses greater transparency and more dynamic communication between information domains regardless of business logic, processes, and work flows.

Throughout the course of this book we will examine how this emerging discipline will drive business innovation through strategic value, and we will provide a detailed analysis of the technologies and techniques embodied in the emerging discipline. To make a complete case as to why this technology is important, we must explain the need, present forces at play in business, and identify the major barriers to strategic progress in information technology.

THE THIRD WAVE

We are at the cusp of what has been called the "the third wave"—an information society. Ten thousand years ago, the first wave, the agricultural revolution, launched a society no longer based on hunting and gathering, moving to one based on farming and a less nomadic lifestyle. The second wave, based on manufacturing and production of capital, resulted in urban societies centered around the factory. The new era, based on information, brainpower, and mediacentric communications is taking shape now. The future will reveal the profound social, cultural, political, and institutional effects this shift will have. The foundation of this new wave rests squarely on our capabilities to make information sharing ubiquitous and meaningful.

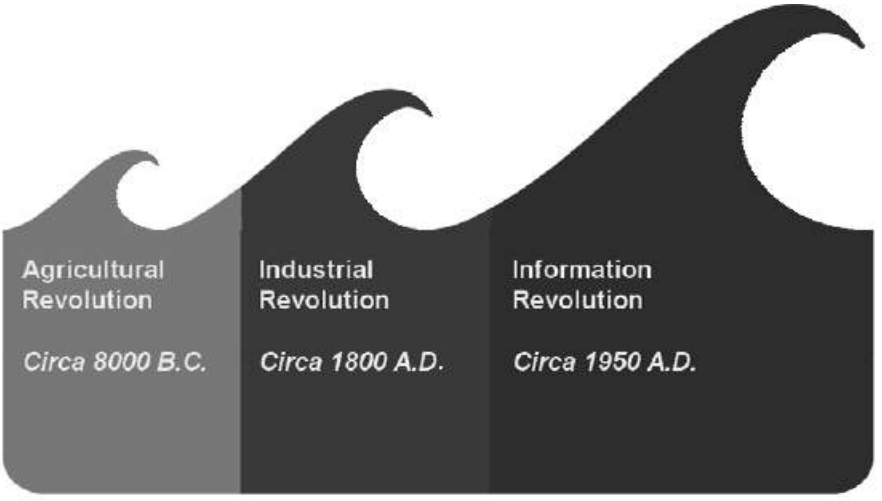

Figure 1.2 The third wave, an information revolution

Competition has been a crucial part of society for our entire history. There is no reason to think that the forces and motivation for competition will diminish along with the rise of the third wave. However, there is growing evidence that the fundamental *drivers* of competition will change as the basis for productivity changes. The decisive factors of production have shifted. In the past, land and capital were the dominant factors of production. Today, knowledge (of patents, processes, formulas, skills, technologies, customer information, and materials) is as important to businesses as land and capital were two hundred years ago.

IT can be used to improve competitiveness in a myriad of ways. Unfortunately, many of the applications of IT are quickly becoming commodities. E-mail connects us. Enterprise resource planning (ERP) systems monitor and store our company financials, human resource data, and supply chain information. Product data management (PDM) systems enable us to produce cars, airplanes, and other heavy equipment with greater efficiency and higher quality. The ubiquity of IT has reduced its impact on corporate strategy because companies have access to the same technologies, making it ever more difficult to use it in differentiated ways. Nonetheless, information used in computers is what keeps the world of business moving.

Strategies for Competitive Advantage

Today's marketplace is the most competitive and dynamic in all of history. Businesses are vying for competitive advantage, indeed survival, on many fronts. Systems in the value chain make businesses more efficient and reduce operating costs when they work together collaboratively—eliminating labor-intensive work and speeding responsiveness. Change is the only certainty in business. Changing environments create the need for businesses to adapt more quickly in order to stay ahead of their competitors. One way to meet this challenge is collaborative IT systems, which can enable businesses to change processes and procedures on the fly—thereby further differentiating themselves against their competitors. These differentiating factors can be leveraged to drive customer loyalty. By making a business's outward-facing demand-based systems more collaborative, corporate leaders can ensure that their customers always have the best service, information, and insight possible. Finally, innovators across industry will leverage collaborative technology to drive market expansion—growing the opportunity with new services and offerings that can expand their customer base.

Improved competitiveness can be accomplished in three primary ways:

- Improve cost savings by driving down expenditures
- Improve products, services, and organizational structures
- Create strategic advantage by exploiting new opportunity

Information technology can be a crucial enabler for each of these approaches. Despite IT's relative ubiquity in the marketplace, businesses can learn to leverage IT in unique ways that add to a core differentiation of the business.

The IT industry is in a slump. Recent innovations in technology, particularly technology available to business through IT vendors, have not offered innovations

that shift businesses core capabilities in differentiated ways. Too often IT offers "me-too" functionality that is not substantially different from what competitors already have applied.

When competitive pressures drive business decisions, technology is more often a competitive barrier than a competitive enabler. Take, for example, three common business tactics used to drive competitive advantage:

- Merger and acquisition activity
- Corporate consolidation
- Accelerated IT spending

Far from sure things, these tactics are rife with challenges for businesses seeking competitive advantage. As the following sections point out, there is more than a little irony in the fact that businesses are both a master and a slave to the forces, successes, and barriers that information technology presents.

Insider Insight—Dr. Rod Heisterberg on E-Business Transformation

In 1995 something wonderful happened, something quite remarkable—a true phenomenon. During that year the whole world unilaterally adopted the same set of Information Technology (IT) infrastructure standards to share business data, the Internet. No governmental edict or channel master command forced this technological paradigm shift. In 1994, the Internet was still an arcane technology primarily known and used by academia, computer scientists, and the military. By 1996, businesses were being established and were generating revenue from commercial buying and selling transactions by leveraging the enabling technology known as the World Wide Web. This seminal event has forever changed the way we do business—now, electronic business (e-business).

Circa 1999 market share leaders in aerospace, consumer packaged goods, and high tech electronics industries had transformed their e-business models beyond electronic commerce (e-commerce) into Collaborative Commerce (c-commerce). The *Internet Encyclopedia*, recently published by John Wiley & Sons, defines Collaborative Commerce as a strategy for the next stage of e-business evolution with business practices that enable trading partners to create, manage, and use data in a shared environment to design, build, and support products throughout their lifecycle, working separately to leverage their core competencies together in a value chain that forms a virtual enterprise.

Sustainable competitive advantage may be realized by adoption of c-commerce strategies and business models. Rather than simply exchanging procurement transactions as with e-commerce practices, leading enterprises are executing c-commerce strategies to share intellectual capital with their trading partners working as a value chain that provides a competitive advantage for the development and distribution of their products. These collaborative business practices are enabled by semantic interoperability, which is fundamental to realizing e-business transformation. The convergence of business process reengineering and Internet technology that spawned e-business during the 1990s has set the stage for reengineering the resulting management decision-making processes to promote c-commerce as the dominant business model of this decade of the 21st century.

Rod Heisterberg, PhD
Senior Consul for ManTech Enterprise Integration Center

Mergers and Acquisitions

Irrational exuberance and hyperinflated expectations during the 1990s drove capital markets to new highs and increased the desire of business leaders to acquire new profit centers. Even though the number of announced transactions has been lower since then, activity is still at a comparatively high 1996 level and projected to remain fairly steady in the years to come.

Business activity surrounding mergers is important to the discussion of semantic interoperability. M&A activity constitutes a great deal of impact on IT systems for all the parties involved in the transaction. Frequently, mergers have financial plans that set out a projected savings based on efficiencies derived from the synergy of the combined entities. These efficiencies rely on the successful integration of a diverse set of IT systems that typically include financial systems, human resource systems, manufacturing systems, and other mission-critical enterprise systems. Rarely are these business or technical synergies achieved.

When the companies do merge, there is a period of time during which they are legally one entity but actually operate internally as two or more discrete businesses. Typically this situation is maintained for only a few quarters while the business units work through the logistics of integration.

Frequently a number of difficulties arise with the process of integrating IT systems between the newly wedded companies. Solving these issues typically presents problem solvers with one of three IT-centric solution choices:

- **Systems Integration**—Massive spending goes toward the integration of disparate systems so that they can effectively use each other's data and processes.
- **Systems Consolidation**—Extensive effort is expended to move mission-critical data from one system to another. Sometimes, such consolidations are not entirely possible when there are significant differences in the information stored in two or more systems.
- **Systems Fragmentation**—Let each system continue to function in the manner in which it was designed: If it ain't broke, don't fix it. This usually results in unnecessary duplication of human resources and system resources and significant complications with data access and knowledge dissemination

Perspectives: Dozens of ERP Systems per Enterprise?

Consider a large electronics manufacturing company that underwent five merger and acquisition transactions in 18 months, and several more over three years. Its information technology infrastructure was a target-rich environment for consolidation and integration. With over eight financial packages, five human resource systems, and four major product management systems it was having serious trouble realizing the efficiencies that had been promised during the M&A transactions. Different brands, such as Oracle, SAP, PeopleSoft, MatrixOne, Agile, and version numbers R3 vs. R5) further convoluted its system architecture.

Situations like this are more common than many would think.

The bitter truth behind these expensive alternatives is that they are repeated for every new merger or acquisition. Constant change is the speed at which business operates, yet current technologies require months of time and massive expenditures to keep up!

Businesses that can respond to changes more effectively will most certainly earn a competitive advantage in their marketplace.

Corporate Consolidation

Corporate consolidation—sometimes called downsizing or right-sizing—is a business reality with tangible effects on IT infrastructure. Large organizations often turn to business integration as a way to streamline processes, cut costs, and provide better value to customers.

Consolidating IT systems requires extensive analysis and development effort to unify disparate business rules, work flow, and data definitions. Today this is a very labor-intensive process that cannot be automated because of the rigid nature of most legacy systems—human experts must be involved to make sense of poorly managed technical infrastructures.

Businesses often grow in fits and starts. One consequence of this is the fragmented nature of their IT infrastructures. Consolidation efforts are frequently required to impose order on a long legacy of chaos. These efforts can work on smaller scales, but major problems can occur if too much is taken on at once. Significant systems consolidation efforts can stall out, and other solutions—like starting from scratch—are pursued.

Perspectives: To Consolidate or Rebuild?

A large personals and dating website grew rapidly in the early years of its formation—resulting in a fragmented and chaotic architecture for its systems. When undertaking the effort to consolidate systems, as a mature business, it encountered an entrenched architecture that could not be modified in a cost-effective way. A separate team was formed to begin the process of architecting new systems from the ground up—in parallel with its live systems.

Sometimes rebuilding is the best alternative.

Although the reasons for consolidation will no doubt change throughout the years, business drivers that encourage executives to choose consolidation will be with us for quite some time. Unfortunately, today's popular integration and consolidation technologies do little to make this process easy or repeatable.

Increased IT Spending, Increased Agility?

Popular wisdom during the late 1990s had executives spending more than 50% of their annual budgets on technology[3] hoping for increased productivity. A recent survey shows that only 10% of industries show a significant correlation between IT spending and productivity.[4] Yet despite evidence to the contrary, business leaders still envision information technology as a key enabler for achieving long-time goals of becoming a real-time, event-driven organization that can keep pace and even lead industry changes. Faster availability of information on market conditions, customer satisfaction, and product performance are still enticing goals.

In the past, typical cycles for market intelligence and other trends analysis took months and sometimes quarters to process and report on. But as information technology is increasingly connected to remote information providers (customers, products, market indices) businesses are taking advantage of the new information supplies to create real-time views into the health of their organization. The speed at which a company can acquire, process, and understand information about itself, its customers, its suppliers, and the market is in direct proportion to its relative advantage over its competitors. In other words, the better you get at understanding information in real time, the bigger your advantage can be over your competitors.

Ultimately, a commitment to using innovative information technology can be a key enabler for the agile business. For example, the 10% of businesses that showed positive correlation between IT spending and productivity effectively used information technology to drive significant innovations within their industry.[5] An agile business is a business that creates new market opportunities and first outs of downturns. Commitment to using IT innovatively is embodied in a variety of forms including staff training, technology initiatives, and corporate culture—not just a measure of IT spending. But today, the most significant barrier to effective use of IT in the business is nearly insurmountable: end-to-end development speed.

End-to-end development speed is the total time it takes from executive sign-off to deployed functionality. It encompasses the time used for analysis, design, development, and several levels of testing. For most medium and large initiatives, development speed is measured in months, not weeks.

Soapbox: The "My Product Does It All" Syndrome

One of our biggest peeves with the EAI industry is the insistence of some vendors that their software is the right solution for all applications. But EAI and Web Services tools rely on a software platform that requires development of new configurations, adapters, and message formats for every new "node" in their proprietary networks.

[3] *Technology Spending.* CIO Insight, June 2002.

[4] US Productivity Growth, 1995–2000. McKinsey Global Institute, October 2001.

[5] *Does IT Matter? An HBR Debate.* Hagel, Seely-Brown, June 2003.

For example, if the FBI becomes aware of a threat to the port of Sacramento, California it might need nearly instantaneous connectivity to a wide array of state and local IT systems that track shipping and persons information. If the FBI doesn't have access to that data in its own systems, making that data available could take weeks or months—hooking into various protocols, data formats, vocabularies, and query structures.

EAI vendors need to understand where their solutions fall short. Only something radically different from the same old EAI or Web Services solutions will make a difference for these kinds of real-time needs.

Suppose a major manufacturing company has a new business partner. The manufacturer wants to streamline processes—enabling each person in the supply chain to view and modify invoices inside his own systems, rather than having buyers, sellers, assemblers, and engineers work over the phone, fax, or E-mail. This is a pretty straightforward scenario that could take 6–12 months to accomplish inside most legacy supply chain environments.

For executives this delay often represents the hidden costs associated with partnerships, new suppliers, and outsourcing. This limitation can severely hinder the organization's competitive advantage by delaying the creation of new capabilities and economies of scale.

The bottom line is that the value of IT is not in the size of an IT budget. No matter how innovatively a business approaches IT spending, the barriers to rapid deployment of new capabilities remain a significant impediment to the business's ability to execute its core strategies.

BUSINESS AND APPLICATION IMPACTS

The impacts of the third wave will be both broad and deep. As more businesses go global and more businesses go on-line, information access and timeliness are crucial to successful strategy.

Key Industries and Vertical Markets

Industries affected by information technologies are not limited to software or computer manufacturers. Today, nearly all businesses are information intensive. From the automotive industry and its reliance on manufacturing information and the supply chain to shipping businesses with sophisticated logistics tracking programs, big companies use IT systems for all aspects of the business.

Information-intensive industries include:

- Automotive
- Aerospace and defense
- Heavy and light manufacturing
- Life sciences

- Finance
- Professional services
- Government (federal, state, and local)

Key Processes and Application Areas

Within a given industry, indeed within a given business, several processes operate more or less independently of one another. They may deal with customer management, supplier management, or product development. Each process can have dependencies on other related processes, but in and of themselves they usually span multiple, geographically disperse software applications that are united by a common domain space.

Information intensive systems and processes include:

- Supply chain management (SCM)
- Demand chain management (DCM)
- Customer relationship management (CRM)
- Collaborative product commerce (CPC)
- Collaborative planning, forecasting, and replenishment (CPFR)
- Product data management (PDM)
- Enterprise resource planning (ERP)
- Decision support systems (DSS)
- Industry exchanges and portals

It is an understatement to say that information is central to these processes. In fact, these business processes rely on the digital information that flows within their boundaries. However, current technologies that apply to these processes cannot provide real-time mechanisms for handling dynamic changes in the information. This means that most current IT infrastructures cannot fully cope with the complexity of their business environments or the volume of information within them.

THE INFORMATION EXPLOSION

The world we live in today is unlike anything ever experienced in the past. Our global economy runs on no fewer than 200 billion lines of legacy code,[6] and our world population produced over 5 exabytes of content in 2002.[7] This is equivalent to adding half a million new libraries the size of the United States Library of Congress full of printed information in one year: simply staggering. In 2002, so much content was created that it was as if every person on the planet generated over 800 megabytes of content herself.[8]

[6] Rekha Balu, "(Re) Writing Code," Fast Company, April 2001.

[7] How Much Information, Berkeley, 2003 (www.sims.berkeley.edu/research/projects/how-much-info-2003/).

[8] How Much Information, Berkeley, 2003 (www.sims.berkeley.edu/research/projects/how-much-info-2003/).

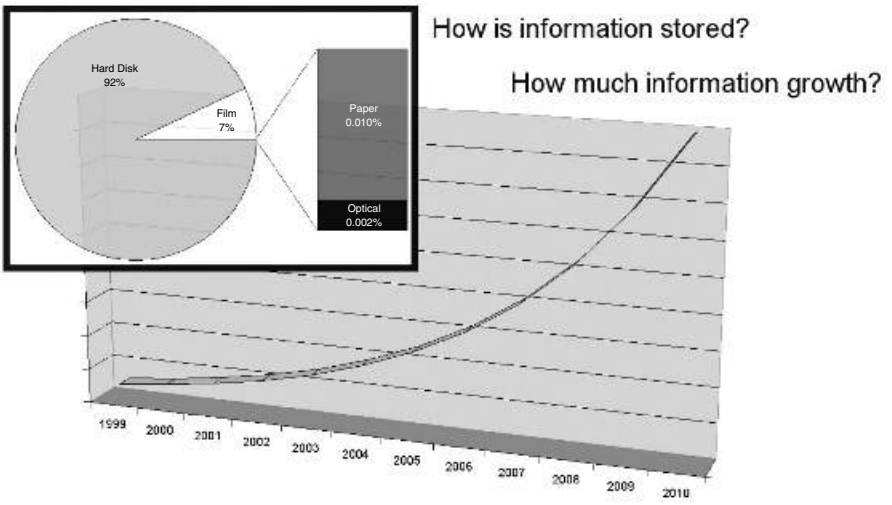

Figure 1.3 Information Explosion

In the three years from 2002 to 2005 more data will be created than in all of recorded history up to 2002.[9] Humanity's economic progress seems inexorably tied to how we create, disseminate, and use information about the world of business. Meanwhile, our tools for making that new information meaningful are severely outpaced.

Information Matters More Than Ever

As a response to this incredible growth, businesses are driving the evolution of software to support data *and* semantics together—which is actually the beginning of information. This capability is crucial to survival and growth. Businesses that effectively use the data they already have will reap the rewards, while those who flounder will be left behind. The interesting things about data and information that matter to most businesses are visibility and configurability of processes, business rules, and key data values. This need will drive development of newer kinds of dynamic information infrastructures. To fill this gap, a richer method for working with enterprise information must be created based on enabling the semantic interoperability of network services and components.

Ultimately, technology by itself doesn't matter to the business—but information does—and as long as information matters, the evolution of *information* technology will, too.

[9] Information Integration A New Generation of Information Technology, M.A, Roth. *IBM Systems Journal*, Volume 41, No. 4, 2002.

Data Rich, Information Poor

Nicholas Carr's article on the death of data processing should serve as a wake-up call to vendors and proponents of the same old middleware technologies. The focus of tools and technologies in the marketplace today is still on data and process. Without a dedicated effort to uncover the implicit meanings behind data and process, finite limitations to information utility will always exist. IT professionals have been too far down in the trenches looking at messages, events, and data to really see the forest for the trees. For decades we've been selling IT middleware products that are basically simple machines for moving things around. As they say, "garbage data in, garbage data out."

The hardest thing about moving on will be admitting to customers that the products they bought five years ago were just Band-Aids on a wound that really needs penicillin. Developing new technologies to move away from the datacentric software paradigm will be easy; many of them already exist. In fact, much of the software behind the semantic interoperability evolution has been developing in labs around the world for more than a decade.

Information Dominance

It is clear that information technology does still matter. But to really get competitive gains from information technology, businesses must also invest in technologies that make information more useful. In effect, this is an investment in information dominance over competitors. This goal can be difficult to justify to cash-strapped financial officers.

A delicate balance is needed between more traditional technology investments (like hardware, desktops, servers, ERP systems) and newer investments (like semantic interoperability, ontology development, and modelcentric development). A successful balance requires a targeted and creative application of IT budgets unique to individual corporate entities.

Balancing these priorities will lead to significant IT innovations and ultimately to competitive gains for the companies who invest in them. Information dominance is truly the key strategic objective for businesses. The struggle to achieve and maintain information dominance will play out as the crucial battle in capital markets. However, longstanding barriers to IT innovation will have to be overcome to really see substantial progress. It is these barriers, which stand in the way of this next great wave of technology improvement, that go unaddressed in the vast majority of vendor products on the marketplace today.

SYSTEMIC BARRIERS TO IT INNOVATION

The next great hurdle in enterprise computing is interoperability among the multitudes of core systems that enable businesses to function. In addition to the many known technical issues with certain middleware technologies, there are several

strategic issues that cut across all approaches to solving the interoperability challenge. These big issues are the most significant barriers to achieving an information technology infrastructure that will create competitive advantage indefinitely into the future. The same old approaches will run into the same roadblocks unless fundamentally new concepts are introduced for handling enterprise information.

The Speed of Change Barrier

The speed of change barrier is the gap between the rate of change in business and the ease of changing the IT infrastructure. The more quickly market conditions and organizational structures are changing, the more difficult it becomes to change the IT infrastructure. Therefore, when the cumulative effects of IT on a business enhance its agility, it becomes more difficult for that business to gain additional flexibility and agility. The closer an IT infrastructure comes to the actual "speed of change," the more effort is required to make additional gains. This behavior is crucial to understanding the business's ability to rapidly scale vertically and horizontally, meet new demands from customers, partners, and suppliers, and manage IT spending.

More than ever, business and governmental organizations are asked to move very quickly. Whether it is federal agencies responding to crisis or corporations forming ad hoc alliances to capture new markets, the demand for rapid, fluid, and frequent alignment of organizations is on the rise. Accordingly, IT infrastructures must evolve to meet these new demands, and a number of enterprise architectures have emerged as candidate solutions for this problem. Business pundits call these visionary solutions by a litany of different terms, but a common characteristic for all of them is their heavy reliance on technology as the key enabler for managing rapid changes in a business environment.

However, IT's actual capabilities leave business leaders wanting. In practice, the amount of effort required to take on even the most trivial of tasks is monumental. This is so because of a number of factors such as different corporate stake-holders with different interests and the sheer volume of customization and code-writing that must be done to accomplish any given enterprise technology solution.

One way to look at the problem is to attach figures to the costs associated with providing interoperability of information systems—and to determine a traditional cost/benefit trade-off. But, as implied above, a more accurate measure of the cost is to analyze how much opportunity is lost by *not* doing the integration—now and in the future.

This notion of real-time scalability is best understood by imagining the opportunities a large business has every day to streamline its value chain, connect to partners, listen to customers, and create new value. Hundreds of small opportunities get overlooked each day because middle managers know that building a business case and lobbying decision makers is a time-consuming effort with an unlikely payback—executive approval and a secured budget.

Now imagine that the business had an infrastructure that dynamically adapted to new requirements for communication, connections, and integration. Imagine that

it only took minutes to look up, connect, and establish interoperability, rather than taking months to go through design, development, and testing phases of most integration efforts. Costs per integration would be drastically reduced, and partners, customers, and suppliers could be brought closer to the business.

Experienced technologists reading the previous paragraph are likely to be laughing right now, because even if a technology existed that could dynamically adapt itself and maintain a real-time network infrastructure, there is another significant barrier to integration nirvana that is often the cause of long development cycles—entrenched enterprise system infrastructures.

Entrenched IT Infrastructures

Most large businesses do not have the luxury of modern packaged applications that already speak Web Services and XML. The dominant landscape today, and for the foreseeable future, is that of large legacy systems with thousands of business rules, data formats, and program architectures. These systems were not designed to share information.

Often called "silos of information," these legacy systems are written in a wide variety of languages including COBOL, C++, Fortran, and proprietary languages. More significantly, the processes and data definitions that these systems use are widely divergent and frequently understood by only a few within an organization.

These entrenched IT systems can be neither replaced nor connected together with ease. Inside large manufacturing businesses it is common to have dozens of incompatible systems, each responsible for a different product family. Sometimes the only cost-effective way to accomplish a basic level of integration is to perform swivel chair integration techniques over a sneaker net.

Definition: Analog Integration Techniques

Swivel Chair Integration—integration that is accomplished by a human being who sits in a chair that swivels between different computer terminals and keyboards and manually enters information into the different systems

Sneaker Net—the network between computer systems that is connected by human beings in sneakers, or some other footwear, running between them with new information

These entrenched IT systems represent tangible technical barriers to widespread availability of truly dynamic information sharing. However, even if these technical challenges are solved, there are still more significant barriers.

Fuzzy Data, Fuzzy Process

Even for businesses that have the luxury of modern systems that speak Web Services and XML, a serious and expensive barrier to dramatic improvement exists:

a severe lack of explicit knowledge about business processes and associated corporate information. Process and corporate knowledge is frequently intangible, difficult to locate and associate value to. This slipperiness of enterprise information and processes is the most difficult problem to understand and fix.

The specific meaning of business processes and information inside a company remains black magic to outsiders. Employees steeped in their own culture and practices are responsible for creating and maintaining the entrenched IT infrastructure of the business by using what they know of the folklore of the organization to determine how data is structured, processes are defined, and business rules are implemented.

To illustrate this, consider that there is no single "best" way to model data for IT systems. Given a parts inventory system for the same shop floor and the same parts, fifty different data modelers would model it in fifty different ways. The modeler's own background, combined with his assumed understanding of the organization, results in differences with the models. Although many patterns for good data modeling are understood, even the smallest of subtle differences in the models can result in absolute incompatibility among them.

Even with packaged applications, culture and folklore of the company influence how they are used. Supply chain systems, financial systems, product data management systems, and all other ERP packages require significant amounts of customization to the business rules, data models, and application code. ERP vendors encourage this customization because they know that every business, every organization, every process is different. However, the result of this customization means that even identical versions of an ERP application installed at nearly the same time for two different organizations will not automatically interoperate—customized integration is required.

Perspectives: The Customized Packaged Application

By the late 1990s the federal government had made significant steps toward modernizing its IT systems. Multiple instances of SAP R3 were deployed by several Department of Defense organizations. However, like many large businesses, the government quickly realized that it had another significant challenge on its hands—none of the customized SAP systems could natively talk to one another.

Despite the fact that the versions were the same, the amount of customization required to implement each SAP system made it impossible for them to communicate with each other.

This spawned additional federal programs.

It is difficult to overestimate the severity of this little-recognized truth. Every engineer who participates in the creation, configuration, and deployment of an application contributes to that application's uniqueness. Yet each bit of uniqueness must be accommodated for that system to communicate effectively with other systems.

Humans deal with fuzzy logic very well. We can walk into new situations, find patterns, react, and adjust as we go without completely halting the entire communication process. Unfortunately, IT systems cannot do this. Their architectures are based on the precision of 1s and 0s in expected sequences. This is why an SAP R5 ERP package cannot understand an XML document with an unknown schema. It's why an application developer might use the SAP IDE to write an XSL/T script to convert the unknown type into a known type.

Writing code to accommodate these kinds of situations is the standard response from IT vendors. Sometimes they might include graphical tools, perhaps with some drag-and-drop functions, to assist with writing that code—but the final result is always some type of custom script or compiled object to accomplish an integration task.

Each company's folklore, culture, fuzzy process, and data combine to create the single most challenging barrier to IT evolution.

The "H" Factor

Humans themselves are the single largest bottleneck to adaptive IT systems. Humans must be involved in the finest of details because of the hands-on nature of resolving data and process meanings. We are still the bottleneck for time, costs, and quality. The foundational elements of computing are still driven procedural behaviors that are codified by human programmers. This means that software will continue to answer the questions of how data gets processed and what it will do with the data—but it will not answer the question of *why* it is performing a given set of logic or *why* it should convert data this way instead of that way.

Handling the question of why is crucial in computing because it gives the IT system the opportunity to automate some things on its own. Shifting the burden of programming enterprise interoperability routines from humans to machines can make a dramatic impact on all of these barriers to strategic innovation in IT. For software to handle the questions of *why* it must begin to interpret the semantics of the data it is processing. Semantics are the foundation of all other substantial improvements in information technology and the key to unlocking information from data.

SUMMARY AND CONCLUSIONS

IT matters now more than ever. Persistent challenges to business demands greater flexibility and dynamism. Challenges ranging from the volume of new enterprise data, to the inherent rigidity of custom-coded applications and the increasing demand for real time business change require new ways to manage the information inside the technology. We've seen how key industries and business mergers, acquisitions, and consolidation continue to demand more agility and more flexibility at greater speed. In short, the need for an adaptive enterprise infrastructure persists. Thus far in the book we have examined the systemic business pain points that drive the need for better technology. In the next chapter we will take a look at some traditional solutions that, despite the hype, fall well short of solving these business needs.