The Seybold Report℠

Vol. 3, No. 15 • **November 17, 2003**

A SEYBOLD REPRINT

This material appeared in *The Seybold Report* (ISSN 1533-9211), Volume 3, Number 15, November 17, 2003.

Technology Strategies

The Next Wave: Revolution in Process and Content

Business gloom in the post-dot-com years may not have prepared you for the next round of innovation, but it's coming. A new wave of digital automation is building, one that rests on a new technology platform, will radically alter our content and publishing processes, and will fundamentally change the economics of professional communications. In short, the next wave of publishing technology has begun.

Over the past six months, Seybold and Project10X have been conducting a joint research program examining the next generation of content, printing, and rich-media communication systems. The study has been exploring key issues and drivers that are shaping business directions, processes and technology infrastructure, and the strategies that companies are pursuing. This three-part series presents a progress report on the research to date.

In this first article, we examine the past, present and future of publishing technology, discussing technology changes in the context of major conceptual advances and investment cycles. We then introduce a framework for interpreting future developments and forecast the economic benefits and ROI that we expect to see from the network-services stage of publishing technology.

Next-Wave Publishing Technology: Revolutions in Process and Content

By Mills Davis and Mark Walter

Ending three years of doldrums following the dot-com bust, content and media technology is beginning a new cycle of investment and innovation that will fuel tremendous changes in publishing. In this three-part series, Project10X and Seybold Consulting examine this next wave of publishing technology and the coming revolutions in process and content that will transform the industry.

Last spring Project10X and Seybold embarked on a joint research program to scout the road ahead for next-generation content, printing and rich-media communication systems. Throughout the summer and fall, the study has been gathering viewpoints and perspectives from individuals in companies with significant content and media interests, including corporations, agencies, marketing firms, creative services, media publishers, media producers and printers, as well as technology vendors, technical service providers and associations that serve these businesses.

Looking across these different segments of publishing, the study has been exploring key issues and drivers that are shaping business directions, processes and technology infrastructure, and the strategies that companies are adopting to pursue their goals. This three-part series presents a progress report on the research to date.

Figure 1 summarizes historical and forecasted economic vital signs for the communications industries, according to media merchant bank and research firm Veronis Suhler Stevenson (www.veronissuhler.com). In our research, we have extended the concept of communications industries to include, in addition to commercial content providers, a range of enterprises, media producers and technology-solution providers, as depicted in Figure 2.

For the past three years, business conditions have been difficult for companies with content and media interests. Markets have fallen. Capital projects have been put on hold. Companies have been keeping their heads down and concentrating on reducing their operating costs.

Over the past year, however, there have been leading indicators of recovery in the marketplace. In some segments, capital spending is increasing. It's still uneven, but this is what we would expect in the early stages of an economic upturn. It's a sign that it's time, if not yet to commit, most certainly to get ready.

While companies with content and media interests have been hunkering down, the technology landscape has transformed steadily. We see signs that a new wave of digital automation is building. We could be on the cusp of the next major shift: a new technology platform, revolutions in process and content, and a fundamental change in the economics of computing and communications.

Some of this is not new. The industry has been through platform changes before: desktop publishing, for example. Over the course of a decade or so, personal computers drove down the cost of typesetting a page by a factor of ten, and then did the same for color reproduction. Similarly, the advent of the Web dramat-

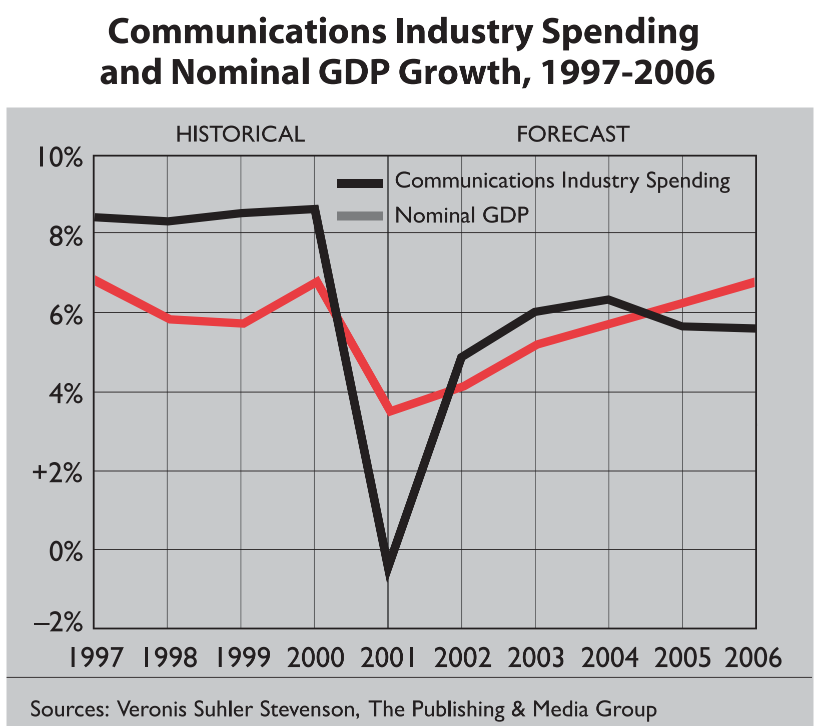


Figure 1. Communications industry spending and nominal GDP growth, 1997–2006. This graph summarizes historical and forecasted economic vital signs for the communications industries, according to Veronis Suhler Stevenson. Digital products and services are growing at a healthy rate, but still make up a small fraction of sales in the industry. Sales in other areas are flat or growing slowly.

ically altered the economics of distribution, and subsequently prompted wide-scale adoption of Web-enabled technologies in business as well as publishing. In both cases, inexpensive technology democratized publishing, opening new markets, attracting new investments and changing the way organizations communicate.

There will be platform changes, new media and new forms of content. Change will be faster than before, and its impact on those with content and media interests will be even more far-reaching than previous technology waves.

But what is new is that this next wave of change looks to be much bigger—in economic terms perhaps 100 times larger. There will be platform changes, new media and new forms of content. Much of this will be driven from outside of our traditional definition of the publishing and graphic arts industries. We also believe that change will happen faster than previous waves and that its impact on companies with content and media interests will be more far-reaching than previous revolutions that have swept through our industry.

The purpose of this three-part series is to summarize our findings so far about the next waves of publishing technologies and to profile what we see as the coming revolutions in process and content. The changes we foresee will affect not only publishing, but all of the communications industries. The series is organized as follows:

Part 1: Next-wave publishing technology. In the article that follows in this issue, we examine the past, present and future of publishing technology. We discuss the development of publishing technology in the context of major conceptual advances and investment cycles and present a framework for interpreting future developments in relationship to previous technology waves.

Part 2: Revolutions in publishing process. Next-wave technologies will transform publishing strategies, processes, platforms, media channels and infrastructure. A key feature of this landscape will be new kinds of processes—integrated network-based publishing processes that span end-to-end value chains, deliver content across multiple media channels and are implemented using shared-resource architectures. The transformations are taking place now. We'll illustrate new directions in communication with examples drawn from the research.

Part 3: Revolutions in content. Next-wave publishing will ultimately be about more than just new processes and media channels. It will also be about creating new forms of content that will open profitable new publishing markets and enterprise opportunities based on transforming content into semantic webs of knowledge assets. Our third article will explain the characteristics of this wave, explore the implications of the shift to semantic webs and present potential scenarios for businesses with publishing, media and communication interests.

Figure 2. Content and media communications landscape in 2006.

Content and Media Communications in 2006

Commercial Content Providers — $760B

- Broadcast
- Entertainment
- News
- Professional
- Business
- Consumer

Media Producers — $600B

- Creative
- Content Management
- Media Production
- Media Distribution

Enterprise Content — $750B

Agriculture	Construction
Energy	Financial Services
Government	Health
Manufacturing	Natural Resources
Retail and Wholesale	Services
Telecommunications	Transportation

Technology Solution Providers — $620B

- Platform
- Network
- Application Packages
- Consumer Electronics
- Technical and Professional Services

Sources: MILLS•DAVIS. US DOC, Veronis, Schuler & Associates, CAPV, Jupiter Communications, Next Quarter, KPMG

Part 1: The Next Wave of Publishing Technology

Taking a historical perspective on technological change affecting our industry, this article explains the forces that are driving the next wave of publishing technology and the impact it will have on the way we build systems. We profile current and future stages of publishing technology, then present a conceptual model of the technical architecture and the life-cycle economic benefits and ROI that we can expect to see in the coming network services stage.

Long waves of innovation

Looking back over the past two centuries, major conceptual advances that power economic growth seem to occur about twice a century. Joseph Schumpeter, an Austrian-born economist, noted long waves of industrial activity in the 1940s. More recently, Merrill Lynch analyst Norman Poire sketched out a diagram *(see Figure 3)* that illustrates Schumpeter's concept. We've added a "you are here" overlay to Poire's diagram to indicate the current intersection of waves of innovation that are propelling publishing technologies today.

When traced back to the Industrial Revolution in 18th-century England, Schumpeter noticed that waves of innovation ebbed and flowed every 50–60 years. Each fresh wave had brought with it a "new economy" that led to investment and excess, followed by a shakeout—but, ultimately, as *The Economist* concluded two years ago, left the world a richer and better place ("A Crunch of Gears," *Economist*, Sept. 29, 2001).

The chart shows six long waves. Inventions in cotton-spinning, iron-making, and steam power propelled the first boom. It lasted from the 1780s to the 1840s. The second wave arrived with innovations in steel-making and railways, lasting for half a century before running out of steam around 1900. Electrification and the internal-combustion engine powered the third 50-year wave. The fourth industrial wave was launched in the early 1950s on the back of petrochemicals, electronics, computing and aerospace. The fifth wave, distributed intelligence, started in the 1970s with the precursors of the Internet. It continued with the adoption of client-server corporate networking, and rapidly accelerated following the introduction of the World Wide Web. In the wake of the dot-com shakeout, this wave is shifting into a new growth gear that will power future publishing technologies. That's right: Far from being over, the current wave has probably another 35 years to go. Meanwhile, a sixth wave is forming that will be powered by nanotechnology, bioscience and clean energies.

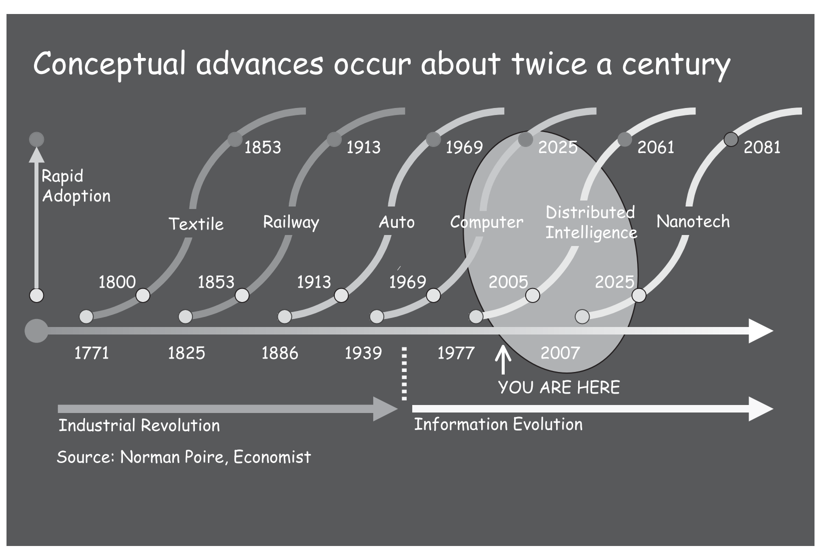
New surges of economic activity tend to play out in four distinct phases. The first phase is a period of rapid innovation as practical applications of seminal inventions emerge. The next phase brings rapid growth as successful participants—whether in cotton, railways, motorcars, electrical goods or petrochemicals—enjoy fat margins, set standards, kill off weaker rivals and establish themselves as leaders of the pack. (In the IT space, we might think of Cisco, Intel and Microsoft as leaders today; but will they continue their dominance during the next wave?) In the third phase, the market matures and the dominant firms hunker down for slower growth. The final phase is a short and sharp decline that occurs when the next set of technologies start jostling for the attention of investors.

IT and COM investment cycles

Within the past 50 years, investments in computers or, more generally, information technology, have also occurred in cycles. Figure 4 maps these cycles of technology investment, using data from Thomas Pisello's book, *Return on Investment of Information Technology Providers*. Over the past 40 years, these cycles exhibited dramatic growth in the amount of new investment they attracted and the shortening of the recovery time between major cycles. New cycles are occurring more frequently, and their impact on the economy as a whole, as well as publishing, is becoming wider and deeper.

According to industry researchers at Gartner Group, Alinean and elsewhere, the next cycle will target network services, rich media, mobility and knowledge technologies. If the pattern continues, the next investment cycle, which is starting now, will dwarf the previous ones in size and scope. At \$6–\$15 trillion dollars over the next ten years, it will be three to five times larger than the dot-com investment boom that followed the invention of the World Wide Web. We believe that companies involved with content, publish-

Figure 3. Long waves of innovation. Today we are at the intersection of three major innovation advances: one nearing its end, one that will continue another 10–20 years, and one that is just starting. These innovation waves spur enormous investments and radically alter the economics of affected industries. As with the computer wave, the current one is affecting virtually all industries.



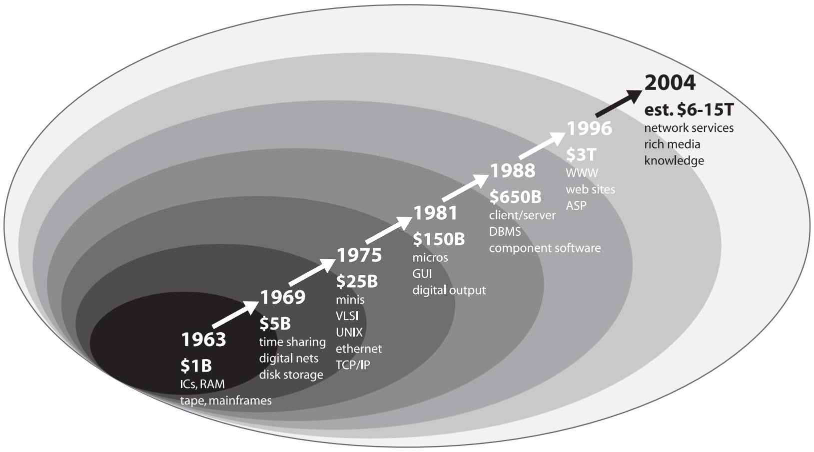


Figure 4. Investment cycles. The next cycle of information technology is under way. If the pattern continues, it will dwarf the last cycle, which prompted the dot-com boom.

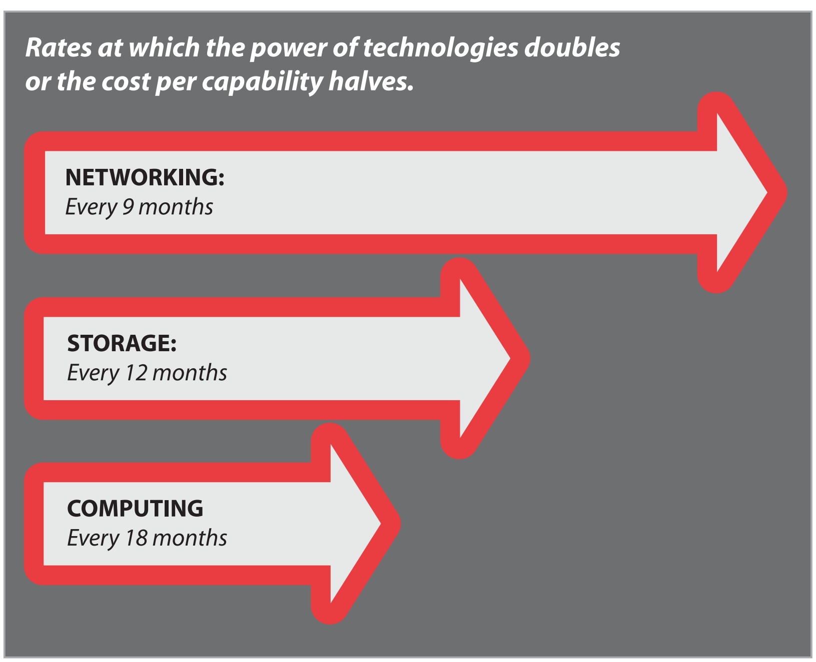
ing and media communication will capture a significant portion of this new investment.

Technology power curve. Figure 5 shows that speed or capacity is doubling (or, more or less equivalently, price is halving) for network, storage and computing every 9, 12 and 18 months, respectively. This means that individual companies and supply chains gain the potential to become vastly more productive, efficient and profitable than they are today. Also, since networking power is doubling twice as fast as computing power, it seems all but certain that high-performance solutions will be built to exploit communications bandwidth and distributed resources.

Figure 5. Technology power curve. Next-generation systems will take advantage of a distributed architecture, because networking power is doubling twice as fast as computing power.

Publishing technology development stages

So, the good news is that long-wave innovations, near-term investment cycles and technology power curves are propelling publishing technology forward. But in what direction are we headed? What sort of innovations should we expect?



Just like computing in general, publishing has had its own stages of development marked by seminal advances in technology. Figure 6 summarizes current and emerging stages of publishing technology and highlights the key directions and innovations of each stage.

The column headings of the chart identify five technology stages. The first three columns—*desktop, server* and *World Wide Web*—are current stages that are still evolving. The second two columns—*network services* and *semantic web*—are emerging stages that will shape publishing profoundly during the coming decade. These stages overlap and continue to evolve, building progressively on each other (from left to right across the diagram). Desktop publishing didn't go away when server-based publishing arrived, and Web publishing continues to evolve, even as network services arrive. As we will see, the dominant force across all stages is the evolution of *distributed intelligence*.

The rows of the chart summarize directions and innovations associated with each stage:

- *Technology directions* are discussed in terms of their economic drivers, the focus of development during each stage, the resulting benefits and the key points of impact. In each stage, the markets, methods, media and economics of publishing are changed.

- *Innovations* of each stage are discussed in terms of platform changes, new forms of content and new media channels. Platform innovations enable economic breakthroughs. New forms of content (such as new digital formats) enable "killer apps" that, in turn, open new publishing markets and change the dynamics of market leaders and laggards. New media channels, venues and devices extend the reach and relevance of publishing and multiply returns on content.

Previous waves. Of course, publishing technology has a much longer history of developments and innovations, dating back to the invention of writing, which we have not included in this chart. Also, we have not included any 20th-century waves of publishing technology predating the personal computer that were presented in Seybold newsletters during the 1980s and 1990s. These included:

- Wave 1 — Hot metal linecasters

- Wave 2 — Phototypesetting

- Wave 3 — Dedicated (mainframe- and minicomputer-based) publishing systems

Since these technologies have all but disappeared from the marketplace, this chart picks up the story

FIG. 6: PUBLISHING TECHNOLOGY DEVELOPMENT STAGES *Stages overlap and build progressively, left to right.*

Wave		CURRENT WAVES			NEXT WAVES	
Stage		**1. Desktop**	**2. Server**	**3. World-Wide Web**	**4. Network Services**	**5. Semantic Web**
DIRECTIONS	Driver	10X reduction in page costs. Content goes digital.	5–10X shift in cost to share information internally between systems and process steps by applying database technologies to content, media and publishing workflow.	Greater than 100X improvement in economics to deliver digital page content/media globally via networks.	5-10X shift in speed and total process costs.	Greater than 100X shift in the economics of knowledge vs. information. Knowledge = theory + information.
DIRECTIONS	Focus	(a) Digital representation of content (text, table, graphic, image, video, audio & interactive data forms)(b) digital tools for content creation, mastering, and rendering(c) devices for input (A2D) and output (D2A) of content (d) single-user platform.	Data representation separated from applications. Computer-orchestrated workflow. Content management (using databases). Cross-enterprise applications. Client-server architecture.	Inexpensive, WWW global connectivity and access to content and media through linked Web sites.	Process representation separated from applications. Shared resource architecture for process integration (A2Ai, B2Bi, B2Ci). Any and everything over IP. Any time, anywhere content delivery. Proliferating media channels & devices.	Knowledge representation separated from content. Knowledge assets in semantic form. Knowledge stacks integrating public, commercial, and proprietary knowledge assets. Semantic media (products & services). Semantic web services. Semantic grid.
DIRECTIONS	Benefits	Individual productivity gains and cost reductions from applying digital tools to manual processes.	Group productivity and business efficiencies resulting from shared data. Enterprise applications (e.g., ERP, workflow, CRM, SCM, EAI, etc.).	Increased revenue, reduced time to market, closer connections to content and customer through Web publishing, e-commerce, and e-business.	Reduced time, effort and cost: (a) to product; (b) to market; (c) to deploy, operate and maintain solution; and (d) to change.	Individual, group, and enterprise super-productivity. New value creation for knowledge-rich content assets. Quantum gains in performance for learning, teaching & communicating aspects of knowledge work.
DIRECTIONS	Impact	Opens new markets for personal, corporate, and professional (commercial) publishing. Massive democratization of media creation.	All-digital workflows drive document management and conventional publishing of print and non-print media.	Opens global markets for WWW-based content. Threatens traditional publishing value chains. New processes for network content delivery, Internet publishing, and peer-to-peer super-distribution of content (e.g., Napster). Internet business models.	Publishing processes have new boundaries: (a) embedded content delivery as part of a larger enterprise or customer process, (b) multi-channel, rich-media content delivery, (c) on-line, interactive games and mobile entertainment, (d) digital libraries.	Semantic tsunami. Revolution in content. opens high value-added markets for knowledge commerce, semantic-based knowledge work, knowledge-based process, and semantic computing.
KEY INNOVATIONS	IT/COM Technology	Single-user application (workstation to personal computer as platform)	Corporate multi-user applications	Global point-to-point information	Integrated, distributed, federated processes (n-player, n-application, n-location, n-tier)	Declarative knowledge-based processes (N-ary)
KEY INNOVATIONS	IT/COM Technology	Local area network	Wide-area network	Global public network	Network service grid	Semantic grid
KEY INNOVATIONS	IT/COM Technology	PC application/application plug-ins	Application server interfaced to desktop client	Web server, Web application servers interfaced to browser client	Service entry points (requestors) and routes among servers and clients	Knowledge server & apps
KEY INNOVATIONS	IT/COM Technology	Platform components, scripts	Middleware	Web server scripts	Process connectors & Web services	Semantic web services
KEY INNOVATIONS	IT/COM Technology	Platform GUI (single user)	Multi-user desktop client GUI seprated from server	Multiuser desktop client GUI delivered in Web browser	No GUI for network services; custom multiuser client GUIs delivered through Web browser	Semantic browsers
KEY INNOVATIONS	Content Forms	Basic data forms and content representation	2D data relationships (tuples)	Globally linked distributed information	Globally-linked distributed processes	Globally linked, semantically-related, distributed resources
KEY INNOVATIONS	Content Forms	Application- and data-specific local file repositories	Application-specific central repository inside organization talks to dedicated clients	Application-specific central repositories talk to Web	Multiple repositories inside and across organizations talk to each other and processes via Web services	Repositories and processes related through ontologies
KEY INNOVATIONS	Content Forms	Platform file system	RDBMS, RO-DBMS	Web site, search engine	Business process mgmt. system	Knowledge operating system
KEY INNOVATIONS	Content Forms	Content formats	Data schemas	Hypertext, hypermedia	Process schemas	Knowledge stack / layers
KEY INNOVATIONS	Content Forms	Unicode, character (command line), text (email, word process), tables (spreadsheet), graphics (2D, 3D), images, audio, video, compound documents, animation, interactive games	Relational tables, object structures, proprietary metadata	HTML linked files, search engine tagging	DOI, BPML, BPEL, Templates, XML-based data structures and metadata representation	RDF, ontology & knowlede-encoding standard, N-ary deep structures of theory, content, and information
KEY INNOVATIONS	Media Forms	Conventional & digital print	Print-on-demand, variable data printing	E-books, personalization & customization of print	Digital paper, network printing	K-books (digital & hybrid)
KEY INNOVATIONS	Media Forms	Stand-alone interactive media on PCs	Server-delivered interactive multimedia to desktop	Multimedia delivered via Web	Voice, video, interactive multiuser multimedia over Internet	Semantic web services delivered anywhere
KEY INNOVATIONS	Media Forms	PC screen (GUI), game station	Terminal and PC screen (GUI)	Hand-held devices (PDAs)	Digital appliances (interactive, video phone)	K-phones, personal knowledge assistants (PKA), ambient intelligence
KEY INNOVATIONS	Media Forms	Hard disk	CD-ROM	DVD	Flash memory	Quantum memories (?)

Figure 6. Publishing technology development stages. The stages overlap and continue to evolve, but each introduces new content forms that open new markets for start-ups that unseat the leaders of previous stages.

Source: Project10X

with the appearance of desktop technology, which we referred to as "The Fourth Wave" in 1987 *(see The Seybold Report on Publishing Systems, Vol. 17, No. 9)*.

Desktop stage. The desktop stage started in the 1980s. It was characterized by desktop publishing, but at its essence it meant building publishing systems from off-the-shelf components—a model made possible by personal computers, graphical user interfaces, local area networks and laser printers. The desktop stage has also been called the era of "aesthetic computing," because it introduced full-color display of content and graphical user interfaces.

Technology development at this stage focused on:

- ***Digital representation of content,*** such as text, tables, graphics, imagery, audio and video, as well as page-definition languages such as PostScript and PDF. It established data formats for individual categories of content.

- ***Application- and data-specific local file repositories*** and platform-specific file systems.

- ***Digital software tools*** for creating content, mastering and output rendering. The desktop era produced a string of killer apps for working with different forms of content.

- ***New devices*** for input capture (analog-to-digital conversion) such as scanners, graphics tablets and mice; as well as for output (digital-to-analog rendering) such as color displays, digital speakers, digital laser and ink-jet printers, digital film and video recorders, digital imagesetters, platesetters and direct-to-press imaging.

- ***Single-user platform,*** operating system and graphical user interface.

- ***Proprietary workflows*** for conventional and digital printing. Stand-alone interactive media on the personal computer.

The economic driving force during the desktop stage was a tenfold reduction in page costs as content went digital—first in black-and-white, then in color. Individual productivity gains and cost reductions came both from working with content in a digital form and from the transition away from proprietary workstations to affordable personal computers. Economic breakthroughs during the desktop stage opened new markets for personal, corporate and professional (commercial) publishing as well as consumer games. The democratization of media and publishing technologies resulted in professional-quality tools at a fraction of what they once cost. For example, 20 years ago a dedicated Scitex color-retouching system cost around $1 million; today a G5 Macintosh and Adobe Photoshop does the same work and costs significantly less than $5,000.

Server stage. The server stage starts in the late 1980s. It is characterized by client-server architecture and the application of database technology to content, media and publishing workflows. The economic driving force during the server stage was an overall five- or tenfold shift in the cost to collaborate and to share information internally within a business, among internal systems and process steps, by applying database technologies to content and publishing workflows.

Technology development at this stage focused on:

- ***Client-server architecture,*** which included application servers, middleware and multi-user desktop client GUIs separated from the server and interfaced over proprietary local, enterprise and wide area networks.

- ***Separating data representation*** and management from applications so that content and data could be more readily shared between applications.

- ***Document and content management*** using databases and proprietary metadata, in which an application-specific central repository inside the organization talks to dedicated clients.

- ***Computer-orchestrated digital workflow*** enabling group productivity increases, for example, by creating automated linkages between steps that streamlined production and reduced sources of error.

- ***Print-on-demand*** and variable-data printing. Server-delivered interactive media delivered to the desktop. CD-ROM media with proprietary desktop user interface.

- ***Corporate multi-user applications*** such as enterprise publishing, ERP, CRM and SCM, which promised to digitize and integrate a wide range of internal functions and processes with significant savings.

Desktop innovations had empowered the individual user. Server-stage technology broadened the scope to the workgroup, department and business-unit level. Desktop applications had opened new markets by redefining what it meant to be a professional publishing "user." In contrast, server technology retooled existing corporate markets in order to unite the best of both desktop and mainframe systems. It paired the flexibility, interactivity and affordability of desktop technology with the scalability, control and power that had previously been the purview of centralized mainframe servers, but was now available more affordably through server technology.

World Wide Web stage. The World Wide Web stage, which began in the early 1990s, is both larger than and substantially different from the previous stages. It is characterized by inexpensive, global connectivity and access to content and media through linked Web sites using universally supported standards and protocols for the Internet. The Web replaced proprietary online networks and their associated products, and created an open and global electronic-publishing platform that included both the underlying network and the user interface.

Technology development at this stage focused on:

- ***Internet build-out*** to provide progressively higher bandwidth, global, point-to-point, digital information exchange. Also, simplification of internal distribution of content through Internet protocols as standards for local area and enterprise networking, not just wide area networks.

- ***Browser-to-web-server architecture*** that includes multi-user desktop client and GUI delivered in a Web browser. Scripting languages link Web servers to application servers and databases.

- ***Web software tools*** including page authoring, site and content management, search engines, as well as e-commerce, e-business and business-to-business application suites.

- ***Internet publishing*** and peer-to-peer superdistribution of content (*e.g.,* Napster). Multimedia delivered via the Web, including e-books and Web-driven personalization and customization of print. Internet-enabled hand-held devices (*e.g.,* PDAs, cell phones); new processes for network content delivery.

- ***Internet business models***. Professor Michael Rappa, NCSU, has documented nine basic categories of business models on the Web, including more than 40 significant variations.

The economic driving force during the World Wide Web stage was a greater-than-100-times improvement in the economics to deliver digital page content and media globally via networks. The innovations of previous stages had digitized and streamlined the process of producing printed publication. The World Wide Web stage transformed both the economics and experience of content and media distribution. This opened huge audiences and markets for WWW-based content.

With the arrival of the Web, online publishing was no longer the sole province of specialty providers (such as Lexis/Nexis or AOL). The Web and its browsers proved sufficiently robust and flexible to handle a wide variety of published content. Suddenly everyone—from journals, newspapers and magazines to corporate brochures and newsletters, catalogs, technical documentation and encyclopedias—had a common platform for publishing electronically.

While Internet publishing threatened to undercut traditional (print-based) publishing value chains, what fundamentally tipped the balance of published communication from print to electronic was the combination of dramatically lower distribution costs with a common, software-based delivery platform. Just as desktop publishing had done previously, Web-based content delivery struck just the right balance between good-enough formatting and an unbeatable price. Benefits have played out in multiple ways: increased revenue via the Web versus previous channel mix, reduced time to market, closer connections to content and customer via content-based e-business.

Network-services stage. The next stage, which we call "network services," is already emerging. It is characterized by a shared-resource architecture, integration of content publishing with core business and supply-chain processes, and multichannel, rich-media content delivery. The economic driving force during the network-services stage is a five to ten times shift in speed and *total process costs*—that is, it affects anyone and everyone involved with the process, wherever they are. These advances are on top of the gains already attainable through the innovations of previous stages. As we shall see in Part 2 of this series, the network-services stage is essentially a revolution in process.

Technology development at this stage is focused on:

- ***Process representation*** separate from applications in the workflow.

- ***Shared-resource architecture*** for process implementation, including business-process management (BPM), Web services, service grid and Internet standards.

- ***Bandwidth on demand***, with anything and everything over IP (using Internet protocols for voice, data, images, documents, video, games, etc.).

- ***Rich media***, mobility, proliferating media channels and devices, and "anytime, anywhere" delivery of content and media.

- ***Create once, publish everywhere*** content workflows.

- ***Total process integration,*** including application-to-application integration (A2Ai), business-to-business integration (B2Bi), business-to-customer integration (B2Ci), and content and publishing integrated with enterprise processes.

To put the technology of this stage into a framework, in the section below we will outline the technical architecture that characterizes this stage.

In the network-services stage, the boundaries between traditional publishing and other processes will become more porous through the use of outsourcing and service-oriented architectures.

In the upcoming Part 2, we will examine how the network-services stage will shift boundaries between traditional publishing and other processes (such as authoring, prepress, printing, distribution and fulfillment as well as core enterprise functions) internally and across the value chain. Boundaries will become more porous through the use of outsourcing and service-oriented architectures. Printers, for example, are more and more repositioning themselves as media producers, and reaching both upstream into the creation cycle and downstream into packaging, fulfillment and distribution. Meanwhile, from both inside and outside the industry, Internet and telecom providers, IT platform vendors, publishers, media producers and other service providers all are rethinking their services and repositioning themselves for a global networked economy.

Another trend we'll take up in Part 2 is that media (that is, the formats of communication) are shifting. The number and variety of channels continue to proliferate, putting stress on older, media-specific processes that don't take into account the need for cross-platform, integrated communication. Whether the business setting is corporate marketing, professional publishing, consumer entertainment or government, there is a growing need for process architectures and service-provider relationships that enable companies with content-media interests to respond to the opportunities quickly, economically and flexibly. For commercial content providers, the key to higher margins will *not* be to just become more efficient at replicating current content processes and products; the key to higher margins will be to become adept at (a) changing processes in order to respond quickly to these changes in media formats, and (b) delivering content digitally in ways that respond to the information needs of individual customers and systematic processes.

Semantic-web stage. The semantic-web stage has just started. It is characterized by two fundamental developments that we will explore in more detail in Part 3 of this report.

First, in this stage, semantics (the meaning of something) gets encoded separately from content. This results in a new form of content. As we have seen in previous technology stages, digitization of new forms of content typically results in new "killer apps," new categories of output and new markets.

Second, semantics gets encoded separately from process. This results in a new kind of platform and a new form of process. As we have also seen in previous technology stages, fundamental advances in platform and process architecture (for example, desktop digitization of content, client-server separation of data representation from applications, WWW global networking and browser-based content distribution, and network-services separation of process representation from applications) lead to grand shifts in content and media economics.

Technology development at this stage is focused on:

- *Semantic web*—moving from simple HTML linkages to machine-interpretable tagged relationships between resources. The goal: a web of globally linked, semantically related, distributed resources.

- *Ontologies*—development of semantic-web resources called ontologies that organize concepts and their interrelationships in ways that facilitate machine reasoning and inference. The goal: resources, repositories and processes semantically related through ontologies.

- *Semantic-web services*—developing shared semantics that enable disparate systems to discover a Web service and understand what it does, how it works and how to access it.

- *Semantic grid*—developing shared semantics to facilitate multi-participant dynamic specification, allocation and persistent management of distributed computational resources.

- *Knowledge technology*—commercial standards for encoding "all knowledge" in semantic form; knowledge engineered as a declarative structure of information plus theory into a semantic web that creates a knowledge asset; conversion of commercial, public and proprietary knowledge-rich content into knowledge assets; separation of knowledge assets from applications; semantic content management (overcoming the complexity issues with RDBMS technology); and semantic processing (hardware solutions for handling massive declarative structures, *i.e.,* semantic webs).

- *Semantic tools*—"Learning" tools for converting content to semantics and amassing knowledge.

These “agents” extract semantics from any and all forms of content, and encode it in a knowledge document or a knowledge set. Knowledge-worker tools enable individuals and teams to work with ideas in semantic form and to integrate knowledge assets together. Semantic browsers (or “knowing” tools) view knowledge and related content, following reasoning paths or answering questions. Knowledge engines form part of knowledge-based computing and application processes. Communicating and teaching tools translate from semantic form to language, picture, simulation and other content forms.

The economic driving force in the semantic-web stage is a 100-fold shift in the economics of knowledge (as contrasted with information). As we will discuss further in Part 3, we see this stage as a revolution in content that will redefine industries. It has huge ramifications, not only for individuals and organizations with content, media and publishing interests, but also for the global competitiveness of our entire economy. For commercial publishers, for example, the changes contemplated in the semantic-web stage envision significant new market opportunities that build firmly on existing customer relationships, content assets and subject matter expertise. Content-provider knowledge assets will enable individuals, corporations and governmental organizations to leverage learning, knowing and communicating aspects of their knowledge work, thereby creating superefficient processes, super-productive individuals and groups, and organizations truly able to compete on the basis of knowledge superiority.

The architecture of the network-services stage

In the previous section, we presented a long view of the driving forces shaping the industry in order to profile current and future stages of publishing technology. Now, we want to take a shorter view, focusing on the immediate road ahead.

Figure 7 depicts a new framework for information technology and communications that is emerging for the network-services stage. It will significantly improve the economics of building and operating integrated publishing processes. This new framework is based on four key technologies:

- Business-process management
- Web services
- Service grid
- Internet standards

Business-process management. Business-process management (BPM) is a new enabling technology for implementing and integrating business and production workflows for enterprise content, printing, publishing and media communication. How does it work?

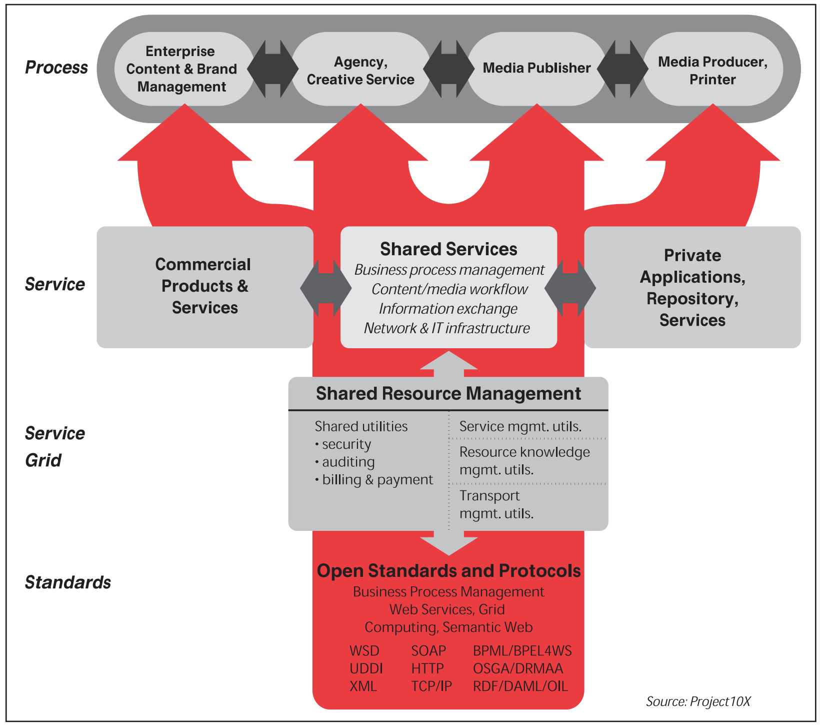


Figure 7. Shared-resource architecture. A grid of Internet-accessible utilities—based on standards—underlies the XML-based Web-services and business-process management advances that characterize the network-services stage.

First, BPM provides a language for defining processes of all kinds. The key innovation is that BPM separates the representation of processes from application logic. The behavior patterns that make up a process are treated as a new data type that can be expressed in an XML schema.

Second, BPM applies equally to all kinds of processes. For example:

- Material, information and business processes
- Internal, external and value chain-wide
- Centralized, decentralized and federated processes
- Manual and automated processes
- Processes built with commercial products and services, legacy building blocks and new purpose-build components.

Third, processes are modeled, executed, monitored and optimized—all in the same environment—using a process virtual machine, called a *business-process management system* or BPMS. This system was one thing that was missing from the attempts at business-process reengineering during the client-server era. The BPMS provides a single system with which to model, execute and change processes in real time, and as often as necessary. This means that companies can both make incremental improvements to processes and undertake larger re-engineering efforts using the same technology. Also, it means that everyone in the supply chain (*e.g.,* customers, providers, supply-chain partners) can interface and integrate their processes together directly and economically, without having to incur the extra cost and inflexibility of trying to implement a "standard system." Also, by combining BPM with Web services, businesses can integrate their legacy systems economically to provide functions needed in their new processes and workflows.

The development of BPMS technology is analogous to the separation of data representation and management from software applications in the client-server stage of technology.

The development of BPMS technology is analogous to the development of database-management systems. With databases, the separation of data representation and data management from the data-using application fueled the growth of client-server enterprise and workflow applications. This time, it is process that is being separated from applications, and this change will greatly improve the economics of building, operating and maintaining integrated processes.

Web services. Moving down the stack, the next technology is Web services. Web services are the new building blocks of processes that will be used both within and between applications as well as within and between organizations. Across networks, everything that happens is a service.

Web services may be thought of as loosely coupled applications that can be exposed as services and easily consumed by other applications using standard Internet technologies. A Web service is identified by a URI (universal resource identifier) whose public interfaces and bindings are defined and described using XML. Its definition can be discovered by other software systems. These systems may then interact with the Web service in a manner prescribed by its definition, using XML-based messages conveyed by Internet protocols.

Services in a service-oriented architecture may be exposed and shared as private, public or commercial Web services. Further, these shared services are arrayed as a vertically layered service stack consisting of:

- *Network and IT infrastructure*—These services provide scalable, reliable and managed transport, storage and hosting functionality across all necessary private and public infrastructure choices.

- *Information exchange*—These services provide secure, virtual access to distributed knowledge, content and information resources.

- *Content-media workflow*—These services provide functions through which value-chain partners collaborate to plan new content and brand assets, originate and manage content, and produce and deliver media products across multiple channels.

- *Business*—These services provide business process provisioning (outsourcing), operations and support, enterprise and value-chain management, and portal services.

Service grid. Service grid is the next technology. It enables ongoing management of distributed resources and the service interactions between them across a process life cycle. From an IT perspective, the service grid is metadata-driven middleware. A service grid consists of a set of utilities that make it possible to discover, integrate with and orchestrate delivery of shared services as part of an end-to-end process. It is analogous to an electrical power grid that everyone can tap into.

Internet standards. In a shared-resource architecture, all levels of the IT and communications infrastructure are implemented using universally supported, open standards, protocols, and management utilities for the Internet. All of these standards are based on XML and existing Internet standards and technologies.

Figure 8 depicts the emerging standards stack. The stack has six layers. The top three layers identify process standards. The first layer identifies industry process templates, such as CPFR (collaborative planning, forecasting and replenishment) and SCOR (supply-chain operations reference model) for supply-chain integration, and STP (straight-through processing) for financial services. These are supersets of business-process management. The second layer identifies standards for business-process notation.

The third layer identifies standards for implementing business-to-business, as well as intra-business, processes and workflows. For example, the United Nations e-business standard, ebXML, deals with establishing trade agreements and modeling integrated business-to-business processes, including sourcing arrangements (via BPSS). BPEL4WS (business-process execution language for Web services) and XPDL (the Workflow Management Coalition's standard for workflow modeling) are proper subsets of BPML. Also, we believe that by Release 2.0 of CIP4's JDF and JMF (job-definition format and job-messaging format), this standard for defining print products and specifying production workflows and messaging between graphic arts applications will also become a subset of XPDL (and BPML).

Life cycle benefits

What is the principal advantage that a shared-resource architecture can provide when compared to previous technology stages? It gives any business greater choice and control over how it manages its content, information technology and communications investments. Each company can weigh the relative value of making, buying, renting (outsourcing) or sharing (*e.g.*, open-source) its needed functionality, and can determine the most desirable combination of benefits, costs and risks.

Figure 9 depicts the value of the architecture graphically in terms of return on investment (ROI) from information-technology projects. The diagram depicts hypothetical ROI curves for two implementation scenarios. Both curves show cumulative returns across the life of the investment. Time flows from left to right. The amount of return, positive or negative, plots vertically. Obviously, the most desirable ROI curve is one that (a) requires zero capital investment, (b) begins to produce positive returns almost immediately, (c) produces substantial positive returns in a reasonable timeframe, and (d) speeds up and drives down the cost of additional related projects that produce even greater returns.

The upper curve depicts a shared-resource scenario based on open, vertically layered network, content and business-process services. The company has a choice of options—make it, buy it, host it or rent it—for provisioning the components of the integrated solution. This scenario assumes that capabilities are provisioned as shared services. One or more service providers host these capabilities, which are accessed and managed through the service grid. Web services are used to interface and integrate legacy systems, databases and toolsets into the new solution. The result is an open, standards-based, architecture-driven content and business workflow that embraces and extends existing capabilities, and runs equally reliably behind the firewall, at the edge, within the network or across the network.

The lower curve depicts a do-it-all-yourself scenario. We're all familiar with internally developed solutions. Content and business workflow is custom-integrated from off-the-shelf software products. The IT and communication infrastructure is built internally.

How do these curves for these two scenarios compare across the life cycle of the hypothetical project?

Innovation phase. The first part of the curve depicts the innovation phase. Every project begins at zero, with nothing ventured and nothing gained or lost. The slope of the graph during innovation is negative; expenditures come before returns. This is the time of greatest risk and greatest exposure. A key question in evaluating this phase of a project is: How deep into its pockets is the company being asked to go?

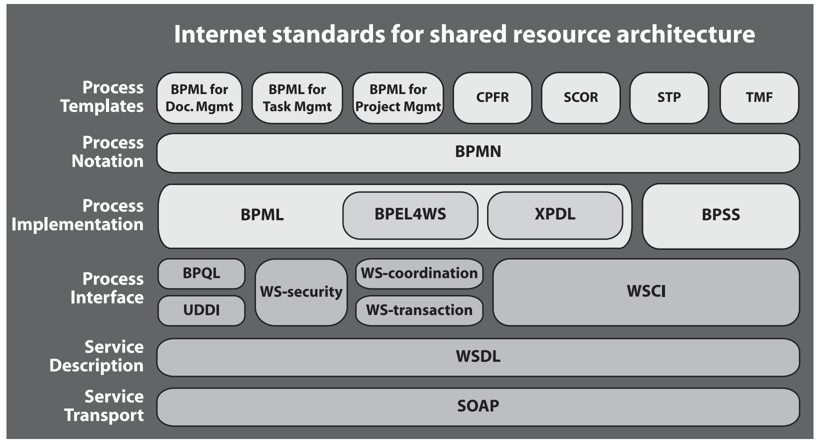


Figure 8. Internet standards for shared resources. The six-layered stack shows the relationship of standards that will play pivoltal roles in building the shared-resource architecture. The lower three layers identify standards for executing processes via Web services. The fourth layer deals with process interfaces, discovery of Web services, security, coordination and orchestration of services, and related transactions. The fifth layer deals with Web-services definition. And the sixth layer deals with service transport across networks (*e.g.,* using SOAP, the simple object-access protocol).

During the innovation phase, the shared-services scenario incurs relatively low setup costs. No capital expenditure is required for new facilities, equipment or software licenses. It is possible to host the application integration, workflow development and pilot testing within the network, reducing the time and resources to innovate. Workflow is largely assembled from pretested building blocks. The good news is that shared content-management and workflow services, business-process services and network services are pre-integrated using open standards, so that both process-

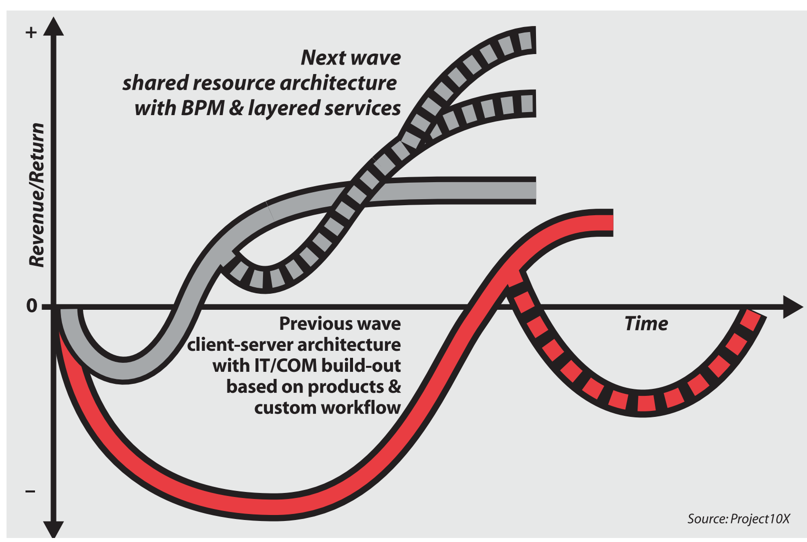


Figure 9. New value equation for return on investment. In the next stage of publishing technology, content systems built under a shared-resource architecture will have a better ROI (top) than those built under the typical scenario of business-workflow applications built with off-the-shelf software products (bottom). The solid lines represent the initial innovation and deployment phases of the project; the dotted lines represent subsequent growth phases that spring from the initial project. The shared-resource architecture shows a better ROI curve because it incurs fewer setup costs, costs less to maintain, and makes it easier to introduce follow-on projects that extend the original project to a wider scale.

development and switching costs are minimized. Open, vertically layered, shared services deploy rapidly.

During the innovation phase, in contrast, the do-it-all-yourself scenario incurs higher setup costs for facilities, equipment and software licenses. Workflow development and pilot testing require significant time and resources. Deployment of the application across a proprietary network infrastructure (especially if it is global) is difficult and time-consuming.

Operations phase. The second part of the curve depicts the operations phase. Solution deployment and initial operations frequently overlap. As operations are phased in, the slope of the curve slows its decent, levels out and begins to rise. This is called the inflection point of the curve. Unless the project is an abject failure, returns start to accumulate. When benefits exceed operating outlays, the curve turns upward. When cumulative returns equal cumulative investments, you reach the break-even point. The key question in evaluating this phase of a project is: How long will the company have to wait for positive returns?

The curve continues to rise ("in the black") so long as the benefits exceed the operating costs. Eventually, though, the benefit stream will slow—the market shifts, the business environment changes, the equipment wears out—and the curve flattens and begins to fall. The key question in evaluating this phase of a project is: What is the maximum positive benefit (or upside)?

During operation, the shared-services scenario requires less overhead for facilities and staffing. The IT burden of maintaining and supporting users is outsourced. Service-grid principles simplify points of accountability and management of quality-of-service. Security and robustness of the deployment is much easier (and less expensive) to ensure for mission-critical workflows. Since transport, storage and hosting are outsourced, off-site backup and recovery are less costly to provide.

Meanwhile, the do-it-all-yourself scenario requires significant overhead for facilities and staffing. The IT burden of maintaining and supporting users is borne internally. In this scenario, maintaining high-quality service may be problematic or expensive, especially for integrated solutions that need to be secure, robust and interfaced with customer or supplier systems. Off-site backup and recovery, for example, is costly to provide on a one-off basis.

Growth phase. The third part of the curve depicts secondary and tertiary enhancement projects that build off of the content solution established by the primary project. The measure of performance that is relevant here is the ratio of added cost to added value. A good ROI curve would enable these projects to begin in a timely manner and be funded by positive returns from the base project. A key question in evaluating this phase of a project is: What is the total upside for related projects that can be funded from the proceeds of this project?

During the growth phase, a solution built with shared services should prove easier to scale up and scale out—adding new users, locations or storage capacity. The shared-resource architecture is designed to facilitate best-of-breed substitutions, integration of new capabilities and extensions to embrace legacy applications. As a platform for growth, shared services provide greater integral value and potential leverage to drive secondary and tertiary ROIs. For example, start-up costs for enhancement projects can be minimal (since development and pilot efforts can be hosted in the network), and deployment is easily supported. Since shared-service solutions are built on standards, switching costs are low. This minimizes risk and helps preserve the value of investment because the entire enterprise-content solution can, at any time, be moved in-house to run in the company's data centers.

During the growth phase, the do-it-all-yourself scenario tends to be harder to scale, due to architectural rigidity and long provisioning cycles. Also, it tends to provide a poorer platform for growth, making it harder to leverage investments. To evolve the solution, new setup costs arise (*e.g.*, for facilities, equipment, software upgrades, etc.) to support development and pilot testing. Deployment inevitably is harder.

To sum up, this comparison illustrates some of the implications of different choices of architecture and building blocks for constructing integrated enterprise-content solutions and commercial publishing systems that become available during the network-services stage. In our example, a shared-services scenario requires less capital investment, incurs smaller risk exposure, breaks even sooner and provides greater integral value for future growth. By comparison, the conventional scenario requires more capital investment, incurs greater exposure, has a longer time to break-even and provides less integral value to support future growth. The good news, of course, is that a shared-resource architecture allows companies to mix

Coming Next: Part 2

Our second installment in this three-part series will present case studies and stories that illustrate the conceptual framework introduced here in Part 1. Part 2 will examine how new publishing technologies are (and will be) working out in the marketplace. We'll discuss 30 key characteristics of content and media publishing in the network-services stage and illustrate new directions with examples drawn from our research. Case examples will highlight different facets of the new publishing landscape. For example, some examples will showcase new publishing strategies; others will illustrate new value-chain relationships. A key focus will be new process models and product (and service) platforms. Some cases will feature new media and multi-channel delivery; others will illustrate infrastructure changes. TSR and match these approaches in whatever ways make best sense for their business.

Summary of Part 1

This article has highlighted findings from new research by Project10X and Seybold Consulting into next-wave publishing technologies. The findings suggest that the industry is at an important juncture—one at which previous stages of innovation continue to play out even as a new one begins.

First, we identified long waves of innovation and investment that are driving technology development. We noted that research from several quarters indicates that the industry is at the cusp of a new cycle that is likely to dwarf the dot-com boom.

Second, we examined five stages of current and emerging publishing technology: desktop, server, World Wide Web, network services and semantic web. We profiled the driving forces and key technology innovations in each stage. We discussed platform changes that lead to new processes, redeployments of functionality and economic breakthroughs that restructure markets and fuel growth. We described the development of new (digital) forms of content leading to “killer apps” and new markets. Also, we talked about the emergence of new media forms, which lead to distribution breakthroughs, media substitutions, market extensions and multiple returns on content.

Third, we presented a conceptual framework for the new IT and communications architecture that is emerging in the network-services stage. This framework recognizes the emergence of a key enabling technology—business-process management—and relates it to the rise of Web services or, more generally, service-oriented architectures. We discussed the improved economic options this brings to the tasks of building, operating and evolving next-wave publishing systems.

Our intent in this first part of the series was to describe the conceptual framework that emerged from researching technological changes in our industry. The framework applies not only to publishers, but also to every business with significant interests in content, media and digital communication—agencies, media producers and corporate enterprises. As you develop and refine your own technology strategies and plan projects for 2004, the framework serves as a model for placing your own projects (and those of competitors) in a larger context, namely, how they will help your organization make the transition into the network-services stage of publishing technology. **TSR**

About the authors

Mills Davis is managing director of Project10X. A researcher, visionary and industry consultant for more than two decades, Davis focuses on improving profitability through process innovation, technology investment and strategic partnerships. Previously, he directed the Digital Roadmaps Project and the HiFi Color Project. He can be reached by e-mail at mdavis@project10x.com or by phone at 202-667-6400.

Mark Walter is senior editor of *The Seybold Report* and the lead Seybold analyst on the Next-Wave Publishing Technologies project. Contact him at mwalter@seyboldreports.com.

The Seybold REPORT℠

Vol. 3, No. 20 • **January 31, 2004**

Publishing Strategies

Next-Wave Publishing Technologies, Part 2: Revolutions in Process 3

The first part of our series on next-wave publishing technologies *(Vol. 3, No. 15)* presented an architectural framework for visualizing how publishing technology is evolving. Picking up where we left off, this issue takes a penetrating look at the impact that these technologies will have on corporate and commercial publishing businesses and the processes that support them. Our study outlines repercussions in four areas, with industry examples illustrating how the broader trends play out in specific industry applications.

Content strategies. What are the implications of these technologies on business and content strategies? We found strategists in both corporate and commercial media sectors tackling globalization, market and media penetration and e-business opportunities.

Media platforms. Digital media are proliferating at ever-increasing rates, making it harder for companies to optimize their workflows for specific media. Recognizing key media characteristics, and exploiting them, will be critical to executing next year's strategies.

Product platforms and infrastructure. Executing content strategies across multiple media will require a supporting infrastructure for designing, producing and delivering digital content products and services. What does it look like, and who's building one?

Value chains. Who does what in the workflow, and how will their positions be different than before? Our research found that boundaries are blurring and roles are being redefined, as existing supply chains tighten and new ones emerge.

ALSO IN THIS ISSUE:

Print Workflow

DEPARTMENTS

For news between issues, visit:

Seybold365.com

Seybold
SEMINARS PUBLICATIONS • CONSULTING

Business & Editorial/Production Office
428 E. Baltimore Ave., PO Box 644
Media, Pennsylvania 19063, USA
phone: (610) 565-2480
fax: (610) 565-4659, 565-3261
web: www.seyboldreports.com

editor: Peter E. Dyson (phone x122)
e-mail: pdyson@seyboldreports.com
managing editor: Patricia S. Evans (x125)
senior editor: Mark Walter (x126)
executive editor: George Alexander (x142)
contributing editor: Steve Edwards
correspondents: Kurt Wolf
e-mail: kwolf@uitikon.ch
contributors: Laurel Brunner, James Felici, Hans Hartman, Matt McKenzie, Bill Rosenblatt, Ron Roszkiewicz, Andy Tribute, Bernd Zipper, David Zwang
production manager: Edmund H. Rozecki (x118)

Customer Service
phone: (800) 325–3830 or (610) 565–6864
fax: (610) 565-1858
e-mail: pubsvcs@seyboldreports.com
accountant: Shirley Gay (x 112)

San Francisco (Seminars) Office
795 Folsom St., 6th floor, San Francisco, CA 94107-1243
phone: (415) 905-2300
fax: (415) 905-2FAX
publisher: James Smith
editorial advisor: Craig E. Cline

The Seybold Consulting Group
The Seybold Consulting Group advises both users and vendors on an individual project basis. Please direct inquiries to Mark Walter at (610) 565-2480, x126, or mwalter@seyboldreports.com.

No Crisis in Digital Printing

Although some suppliers targeted the wrong market and are losing money, others are doing quite well.

Kurt Wolf's interesting article in *The Seybold Report* of January 19th *(see Vol. 3, No. 19)* is focused on the digital-printing technology suppliers who are addressing the traditional graphic arts market. Indeed, some of them (like Heidelberg or MAN Roland) were not as successful as they wanted to be.

However, some remarks might be useful. I think the headline provokes a deep misunderstanding. Digital printing is not in a crisis; it's a booming new business! And we have no crisis of digital-printing technology itself. The main players, such as Canon, HP, Kodak Versamark, Konica, Ricoh, Xerox, etc., are earning more money (each of them!) with digital printing than all the traditional players in the offset market put together.

The point is, new technology by itself cannot build a new market. Technology development in the graphic arts business is not ahead of the market. It always follows the needs in the communication and consumer market. That means that at the point that communication became digital—pushed by the Internet—printing technology had to become digital as well. There is no other way.

Digital technology is changing business models. And digital printing is not comparable to a normal, commercial-printing production job. It's a new business for a new kind of service provider, or it could be a do-it-yourself-business within the corporation as well. That's a big difference and the real paradigm shift in the marketplace, which was not recognized by the traditional suppliers. The result: Vendors chose the wrong marketing strategies by choosing the wrong target groups.

In my opinion, Kurt's focus is a bit to narrow, because he reduces digital printing to a minor graphic arts application. Digital printing is the leading application in business and consumer publishing, at least at certain levels of speed, quality and investment. The benefit of digital printing for corporate communication is extremely high on the business side. That is because digital printing is the perfect way for corporations to satisfy their need to address their customers in a personal and individual way. It's all about interactivity.

To repeat: There is no crisis. True, there are some suppliers who are making mistakes or losing money. They missed the boat. But there are others, the majority, who are doing well.

Andreas Weber
weber@digitaldruck-forum.org
Digital Printing Forum,
www.digitaldruck-forum.org

Oris Color Markup debuts

CGS adds Prolatus utility to line

CGS Publishing Technologies has adopted the Prolatus Markup tool, bringing remote markup and approval into its color proofing family. CGS will price the tool, which it's calling Oris Color Markup, at $4,000.

You may recall that Prolatus Markup uses a digitizing tablet to capture reviewers' annotations on a proof print. The annotations are then transmitted to a central server, which retransmits the marks to all participants in a collaborative session. It can also capture an approval signature on a contract proof.

This is the third component of the Oris remote proofing system. The system is based on the Color Tuner color-managed RIP. Last fall, a Certified Proof tool was released; that software verifies calibration of all printers in a distributed system. Still to come, we're told, is a database that will allow trend analysis of the device characteristics at each site in a system.

Peter Dyson

The Seybold Report (ISSN: 1533-9211; USPS 020-953). Subscriptions are available in the U.S./Canada for $595 per year and Internationally for $640 per year (24 issues mailed semimonthly).

POSTMASTER: Send address changes for ***The Seybold Report*** to: Seybold Publications, 428 E. Baltimore Ave., Media, PA 19063
Periodicals Postage Paid at Media, PA and at additional mailing offices.

Next-Wave Publishing Technology Part 2: Revolutions in Process

BY MILLS DAVIS AND MARK WALTER

The second of a three-part series that examines next-generation publishing technologies, this article explains their characteristics and presents industry cases that illustrate the concepts and theoretical model introduced in Part 1.

In Part 1 of our series *(see Vol. 3, No. 15)*, we traced the recent evolution of publishing technologies and introduced a framework for interpreting future waves of innovation in our industry in relationship to previous technology stages, some of which are still evolving. We then outlined the architecture of the next stage, Network Services, which has already started. We explained the benefits of its service- and process-oriented architecture to the life cycle of technology investment, operation and return on investment.

In this article, we look at how next-wave publishing technologies are transforming publishing processes. As the industry moves forward, the nature of publishing is changing along two axes at once. From the top down, we're changing what publishing is: creating new products and strategies that respond to changes and opportunities in the market. At the same time, from the bottom up, we're changing how we go about publishing—adjusting the processes and systems we use to create, produce and distribute content, and redefining the value chains for business and product lines as well. Of course, these two axes are synergistic: A change in strategy often precipitates changes in infrastructure; and adoption of a new technology can pave the way for more efficient processes or new content products.

In this article, we'll start with a discussion of next-wave strategies and a look at the characteristics of emerging new media. We'll then present examples of product and service platforms that illustrate the architecture needed to achieve that vision and conclude with a look at how the strategies and architectures affect the entire value chain in different industries.

Content Strategies

Content strategy is a relatively new term in our industry. It refers to our overall game plan for leveraging intellectual property for financial and strategic advantage. Developing a content strategy has become a corporate necessity, because effectiveness and efficiency of digital communications have become competitive facets of business in every sector of our economy.

In the 1980s and 1990s, during the desktop and network stages of publishing technology, most publishing activities focused on specific end-products, and strategies were concerned with the best ways to acquire, develop, produce and sell those products. We optimized our workflows for specific products that were tied to specific media—ads, newspapers, books, catalogs, magazines and broadcasts. The arrival of CD-ROMs prompted reference publishers to embrace a cross-media content strategy, but even when their strategy was validated by the arrival of the Web, most of the industry clung to the single-media approach, treating the Web as a discrete medium deserving its own strategy, content and systems.

Today, the number of media formats has grown to the point where the need for cross-media content strategies is more obvious. Companies can point to the time and costs associated with reworking content for different media. In response, in the wave that's occurring now, companies will focus on managing components of their content products, breaking them down to a more granular level and separating content from media and products in order to gain efficiency across media. At the same time, they will begin to apply systems to managing process change. Why? Because markets will demand that content be delivered and packaged in both old and new forms, on ever-shrinking development cycles, by parallel, overlapping and short-lived processes.

The content strategy themes of the next wave are interrelated. Markets are becoming more global and media formats are proliferating, both of which put pressure on an organization's ability to adapt content. At the same time, companies will employ a variety of strategies to increase brand loyalty with customers and improve their ability to integrate content systems with back-office systems, both inside and outside of the corporation. In most cases, the content strategies require that you have an inventory of your content, a schema for classifying it and an architecture for organizing and managing it. As we'll see, these strategies are also greatly facilitated by the next wave of publishing technologies.

Globalization of content

There are two aspects to globalization that we'll note here. The first is the expansion of markets for content and media. The second aspect is the dispersion of business and production processes relating to content. Both aspects of globalization affect commercial and corporate content interests.

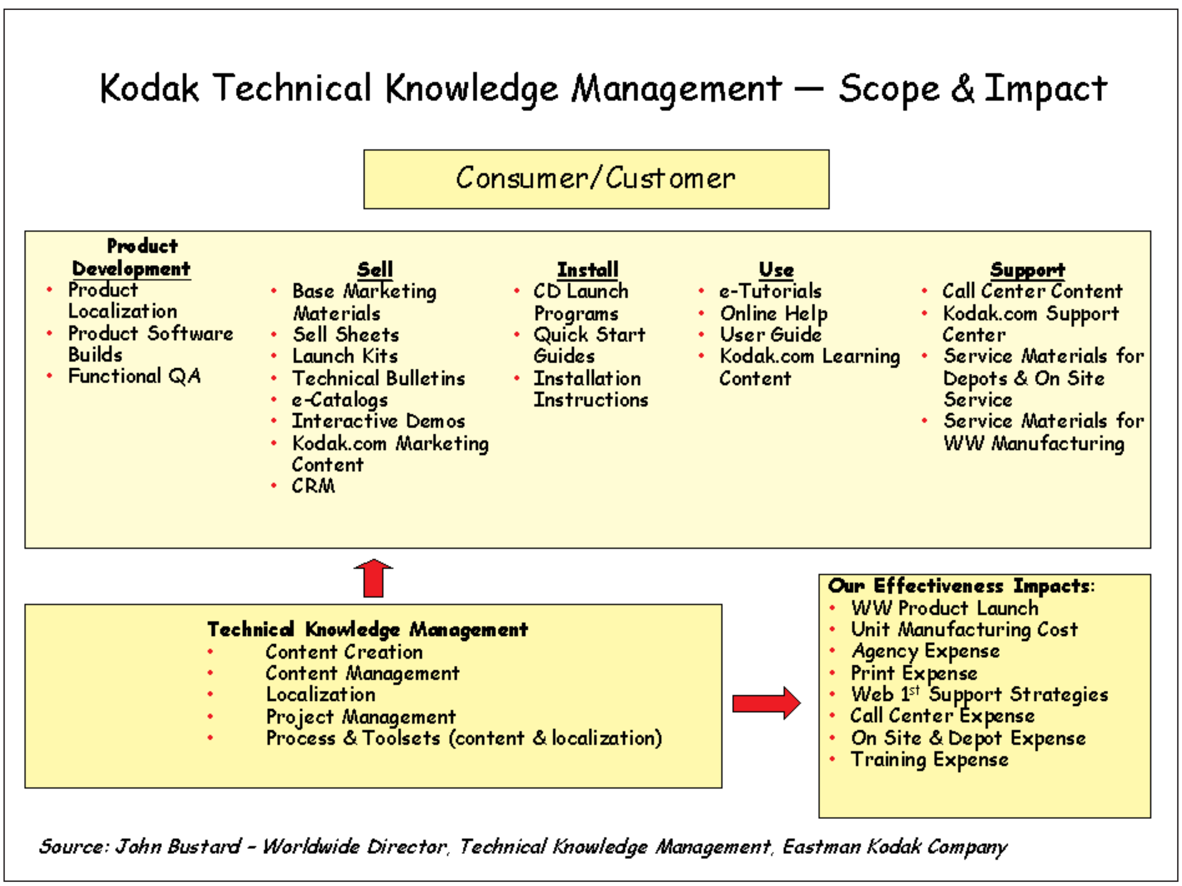
Market globalization. For many companies, 60 percent or more of the market for their products is outside of their nation. To effectively reach these markets, companies must localize product-related content for specific geographic regions, starting with pre-launch marketing materials and extending to advertising, promotion, sales collateral, support documents and Web sites. The inefficiency of managing these interdependencies manually is driving companies to implement Internet-based systems that manage content throughout the life cycle of the product (from R&D through launch, to in-service support and upgrade, to sunset), as well as end-to-end across the value chain.

John Bustard, director of worldwide technical knowledge management for Kodak, illustrated this trend and its potential scope and impact on a manufacturing organization in a speech last fall *(see Fig. 1)*. In support of the company's global sales efforts, Kodak's content strategy is to create once, localize once, store once and deliver the content in multiple ways, including print, Web, CD and e-mail. That strategy led Kodak to implement an SGML-based approach to authoring and translating technical support materials. John Deere, Caterpillar, Hewlett-Packard and many other manufacturers have similar strategies for their technical content, supported by SGML/XML-based content technology systems.

Process globalization. Regardless of where you sell your products and services, the content-centric processes behind them are moving, increasingly to different geographical regions. As we pointed out in December *(see "Publishing Workflows Go Global," Vol. 3, No. 18)*, for commercial publishers, it is increasingly common to find portions of their production process as well as their back-office processes located offshore. Everything from scanning and copyediting to retouching and printing are moving to new countries, especially Asia. For increasing numbers of companies, support is globally sourced as well.

In fact, any number of functions in the supply chain may be (re)deployed closer to end markets or wherever the costs to perform the function are more favorable. The principle of labor rate arbitrage (moving the work whereever the labor rates are lower) is not new. What is new is the degree to which a network-services infrastructure makes this practical.

Fig. 1: Globalization complicates content management. Kodak's strategy for managing technical content takes into account multiple delivery channels and localization issues over the life cycle of product information.



Maximize share of customers

In most businesses, it costs much less to sell to existing customers than it does to acquire new ones, so it is natural that organizations want to maximize their customer relationships. How do businesses go about maximizing their share of a customer, and how do those strategies tie into content?

Follow customer life cycles. One way is to meet more of the needs of customers with which you already do business. A strategy of the Thomson Corporation, for example, is to position publishing businesses and products across the life cycle of the practice of law. As each customer ages and advances in his career, his needs change. By developing a series of products and services appropriate to these different stages, Thomson's Legal and Regulatory group hopes to leverage both its customer knowledge and brand to retain the customer over an entire career *(see Fig. 2)*. Last year Wolters-Kluwer took a similar tack, restructuring its product-oriented health properties to align with customer segments.

Micro-segment. Another way to maximize share of customer is to micro-segment to the level of specific activities, tasks and decisions. McGraw-Hill, for example, has been taking this approach in the construction industry. Its products are positioned along the life cycle of the design-build processes of a construction project *(see Fig. 3)*.

One can see a similar strategy at work in trade press magazines focused on business professions and industries and in consumer publishing titles aimed at individuals. In both cases, the magazine brand and customer relationships create opportunities to introduce complementary Web-based services that zero in on specific tasks, such as researching investments, making a purchasing decision or planning a wedding.

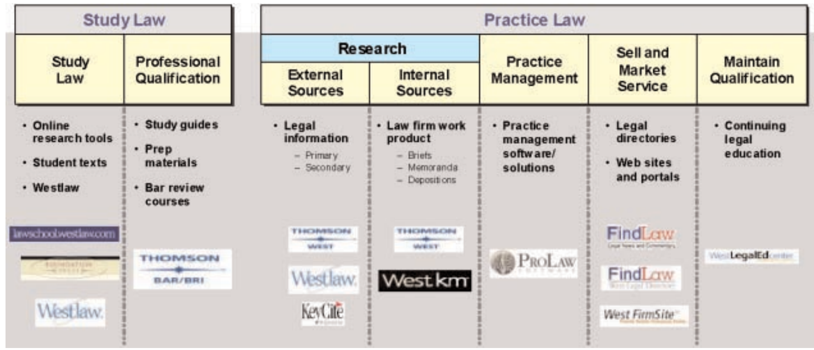


Fig. 2: Targeting the customer life cycle. Thomson has developed a suite of legal products that address the life cycle of the customer—in this case, learning, practicing and managing the practice of law. This strategy affords opportunities for content reuse across different media formats: newsletters, advisories, references in print and online, as well as use across different stages of practice.

At the same time that they segment by task, publishers are striving to embed their content products and services closer to the customer's process. Thus, the *Readers Guide to Periodical Literature* has morphed from a book that stays in a library's reference area to an online search screen that patrons log into from their dormitory or home. Taking it one step further, Factiva, Gale and Electric Library all run inside the Research Pane of MS Office—now you can search for answers while writing your report. Similarly, the service manuals for fixing very large and complex machinery—systems that control ships, jets or power plants, for example—are increasingly linked to the diagnostic tools that help identify the problem, and to the parts catalogs, ordering systems and inventory-management systems that help technicians get the parts they need to make repairs. Segmenting content at such a granular level increases its value to the customer, thereby deepening the customer relationship.

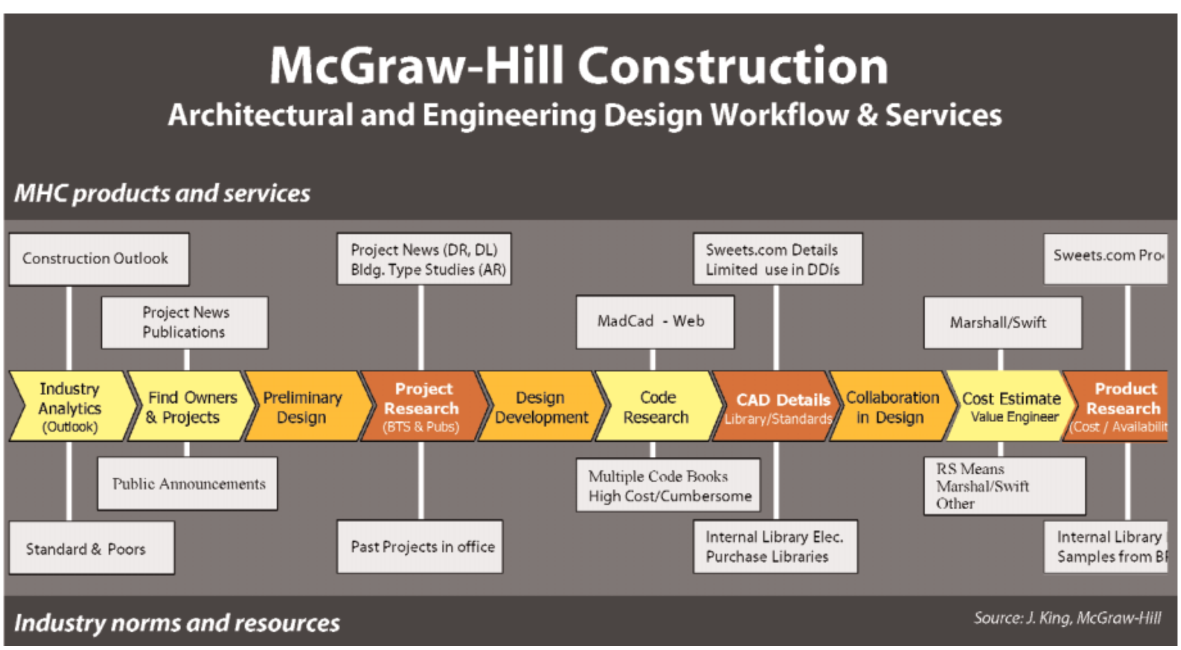


Fig. 3: Segment to process and tasks. McGraw-Hill's construction unit has focused on the life cycle of the construction process, pulling McGraw-Hill products from different business units into a cohesive service that targets different industry participants and addresses the information needs that arise at specific stages of a construction project.

Personalize, customize. Another way to increase value is to customize content for a group of customers or personalize it for individual customers. This strategy has been pursued on the Web for many years. Now we see it gaining traction in direct marketing through the use of variable-data printing. The more advertisers and direct marketers knows about the customer, the more specific they can make their appeal. Digital printing expert Frank Romano reports that response rates to print campaigns typically improve by 11 to 34 percent when they are personalized with information from a database. However, making the piece relevant to the customer may be even more important. According to Rab Govil, president of the Print On-Demand Initiative (www.PODI.org), "relevance is the largest lever that marketers can use to increase response rates." Comparing relevance to personalization, "we found relevance pulled response rates three times higher than just personalization," said Govil. Romano and Govil agreed that combining relevance and personalization yielded the best results. "Relevance with automation is even better," said Govil, because it makes personalization cost-effective on a mass scale.

Create ties to related products. Another way to increase sales from existing customers is to use knowledge of the customers' tastes and habits to sell them related products or services. Amazon set the Web retailing benchmark for this technique, showing customers computer-generated suggestions and recommendations based on their interests and purchases. The online business of Martha Stewart made a similar connection between this strategy and content technology—modifying its workflow so that staff could place tags into articles that indicate what category of products should be advertised on the same Web page as the articles. In consumer markets, it's not just movie studios that leverage content assets by selling toys, memorabilia and collectibles of the same brand. General Motors' licensing of its brand images is a multimillion-dollar business each year, and educational publisher Scholastic now sells music, videos, toys, puzzles, games and even stuffed animals based on characters of its most popular children's books.

This strategy of cross-selling extends to business markets as well. Business magazine publishers, such as Reed and CMP, for example, supplement ad revenue by offering complementary print, online and face-to-face services: advertising supplements, custom publications, seminars, trade shows, Web sites and direct-marketing services.

Maximize share of media

As companies promote brands to a global audience and strive to retain loyalty with customers, they find that they must communicate through a variety of media. "Maximizing share of media" is a strategy to increase return on brand equity and content investment by delivering the brand and related content to an audience through multiple channels as efficiently as possible, without sacrificing brand-identity requirements.

Disney. A good example of a company that works hard at maximizing its share of media is the Walt Disney Company. A consumer-oriented media company, Disney has content assets that it seeks to maximize by leveraging the brand in both traditional and new media.

Several years ago, Disney recognized that the "Internet" is no longer synonymous with the World Wide Web. There are an increasing number of mobile devices (phones, hand-helds, etc.) that are interesting platforms for extending the reach of its brands. Its objective then, as now, was to brand and position content across all viable media.

Derek Kerton (www.kerton.com), a consultant who specializes in wireless markets, worked with Disney to overcome one thorny challenge Disney faced in entering the wireless market: the presentation difficulties posed by different form factors of devices receiving content through different service providers. Kerton points out that many companies were content to treat wireless media as a derivative of the Web, even to the point of scraping screens off their Web sites as a quick and inexpensive method of developing wireless content. For Disney, which has always been near-fanatical about its brands, such an approach was unacceptable. It wanted brand consistency across different media platforms, so the content must be adapted, and in some cases changed, in order to give consumers the high quality of experience that they associate with Disney brands.

Disney's analysis of the wireless value chain, and its diverse and changing technology and specialized skill requirements, revealed that partnering was critical to achieving its objectives in speed to market, flexibility, brand integrity and sustainability over time. Its strategy for moving from wired to wireless markets depended on finding a variety of value-chain partners, including Motorola, Sega, Ericsson, AvantGo, Palm and Openwave. No one does it all.

Drive to all-digital workflows. Assuming that you want your brand and content assets in as many media as possible, how do you minimize the cost of content acquisition, production and distribution, while maximizing the return on your content investments? This is a strategic concern for both commercial media companies and corporate enterprises. Their strategies are to employ a combination of reuse, syndication, aggregation, value-added republishing and superdistribution.

The reuse strategy is a familiar one, and its repercussions overlap with those of the globalization strategies—namely, the trend toward capturing or creating once and distributing everywhere. The higher your level of reuse (for different media, products or markets), the higher your potential return on an investment in content management. Fig 4 illustrates in a simplistic way the scope of content sources and destinations that are driving organizations to make this investment. In conjunction with those content-management efforts, companies seeking to increase efficiency and economies of scale are shifting to all-digital workflows, from acquisition through delivery of their content.

Digital first, analog after. An important tenet of next-wave publishing is that digital products come first; analog ones follow as byproducts of digital masters. For everyone who has businesses that produce print, broadcast or film products, accepting this tenet requires a major overhaul in existing processes. In many cases, because processes were already optimized for analog media, organizations took the opposite approach—bolting a process for making digital derivatives onto the back end of their existing workflow. That is still very evident today in the electronic facsimiles of print products, which include everything from PDFs of corporate forms to electronic editions of books, newspapers and magazines.

Our research found that the graphic communications industry as a whole is midway through this process transformation. In commercial publishing, the reference market has largely completed the transition: The next print editions of *Encyclopaedia Britannica,* the *Oxford English Dictionary* and many other reference works from other publishers will be snapshots of the digital database that has become the primary product. Catalogs are not far behind. Using systems from Pindar, Flow Systems, A2i and others, distributors are constantly updating their electronic catalogs and then producing spin-off print catalogs from the master database. Journals are creeping toward this change, as are newspapers. Even magazines, where feature content is so media-dependent, are taking steps to publish some of their news first on the Web.

Reference and technical content was the first to make this transition in the corporate world, but the shift to digital-first is taking hold here, too. A CAP Ventures research study published earlier this month discovered that although print remains the most often used channel for corporate communications, nearly 40 percent of respondents reported that their content has moved to a workflow in which electronic delivery of content is first (*e.g.*, Web and e-mail), with other formats, such as print, following.

This shift in workflow is important for several reasons:

- *It improves time to market.* Delivery of digital goods over the Net is much quicker than delivery of physical goods via postal service, courier or shipping service.

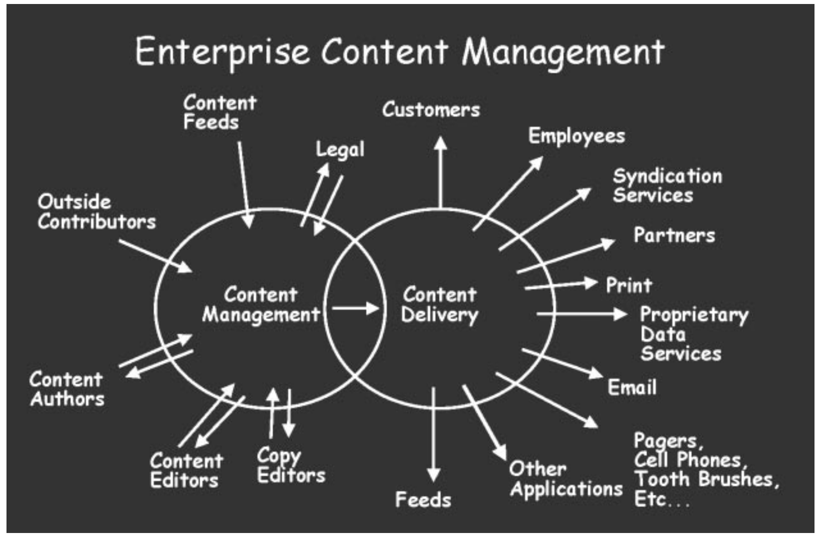


Fig. 4: Enterprise content management. The diagram depicts the multiplicity of sources, destinations and media that are now normal parts of enterprise publishing. To efficiently distribute to multiple media, all of the content acquisition, manipulation and distribution must be managed by interconnected systems.

- *It costs less.* Even when print is the end product, the use of digital masters can slash costs and inventory by enabling on-demand printing or distribute-and-print models. But when the end product is digital, production, the delivery and media-packaging costs are all lower than their physical counterparts. For corporations, where publishing is an overhead cost of supporting the business, lowered costs in communications enhance the profitability of the larger business. For commercial media companies that sell content, lowered costs can increase the profit margins of the goods being sold.

  However, when a technology that was once a barrier to market entry becomes widely accessible to the public, it can spark disruptive change. We see this unfolding now in scholarly publishing. Upset with rising journal prices and empowered by the democratization of Web publishing, academics have formed the "Open Access" initiative, a movement to redefine the value chain in a way that would severely diminish the role of the established journal publishers. Publishers are responding by focusing on adding value, a strategy we'll discuss below.

- *It supports new channels.* Third, digital workflows pave the way for delivering content in new channels—most, if not all, of which will be digitally driven.

Multiply returns on content

Once a digital workflow is in place, additional delivery options open up. We'll mention three: syndication, aggregation and value-added services.

Syndicate. Once the province of large commercial publishers and wire services, syndication is now a content strategy of industrial firms as well as smaller media players piggybacking on standards and tools devel-

KEY CHARACTERISTICS OF NEXT-WAVE PUBLISHING

Publishing becomes complex, multi-channel, rich-media content delivery, causing companies to embrace new strategies, platforms, infrastructure and value chains. Process becomes key to value, differentiation, efficiency and ROI for content.

	Themes	Characteristics and Comments
Strategy	Globalization and micromarkets	Product markets span multiple countries and demand content localization. Content-related processes disperse.
	Maximizing share of customer	Publishers target life-cycle needs of audience; micro-segment to the level of specific activities, tasks, and decisions; personalize and customize content; tie-in to products and services.
	Maximizing share of media	Create once, publish everywhere; shift to digital-first workflows.
	Multiply returns on content	Syndicate, aggregate, add value.
	Embrace e-business	Connect business systems to content; develop, test and refine Net business models.
Media Platform	Media channels proliferate	Broadband, mobility and wireless growth; anything and everything over Internet protocols. Keeping content flexible is key to adapting as new formats emerge.
	Media formats intermingle	Multifunction devices supplant single-function players.
	Code and content merge	Delivery platforms combine rich multimedia content presentation capabilities with Web-enabled programming interfaces.
Product Platform	Multidimensional	Supports multiple stages of content and product life cycles. Many participants; multiple channels.
	Integrated with e-business	Linked to internal and external business systems (ERP, CRM, fulfilllment) and portals; real-time production and packaging.
	Equipped for network collaboration	Supports Web services; distributed architecture enables both intra- and inter-enterprise collaboration.
Infrastructure	Service-oriented architecture	Built with Internet-centric standards, service utility grids, shared services, business-process management. Publishers have build-it, buy-it, host-it & rent-it options for building (and operating) next-wave publishing solutions.
	n-to-*n* integration	Integrates content-media value chain and orchestrates processes. Interoperability with: customer process, content sources, supply-chain providers,external business services, technology providers.
	Life-cycle process management	Unified, open environment for planning, provisioning, deploying, executing, managing and optimizing collaborative, multi-channel publishing processes.
Value Chain	Restructuring what gets done and who does what	Publishing becomes customer-centric, embedded into customer-facing process. Redraw boundaries between members of the value chain.
	Redeployment of resources	Consolidate or distribute workflow as appropriate. Global partnering for needed skills, technology and services.

oped for others. For a current example, consider RSS (Really Simple Syndication or RDF Site Summary, depending on which variant you use). Not only is it simple enough to be used by small commercial publishers (ourselves included), but thousands of bloggers now use RSS to syndicate their online diaries. Even more use RSS news readers to scan multiple Web pages without visiting each of them in turn. The popularity of this format and wide availability of these feeds creates new opportunities; recognizing and capitalizing on these types of opportunities will be crucial to discovering new businesses in next-wave publishing.

Aggregate. Acquiring content digitally lowers the cost of aggregating content sources so that customers can be offered a wider array of choices from a single source. This strategy drives many media mergers—Elsevier's acquisition of Academic Press and Harcourt's Health business; Wolters-Kluwer's acquisition of Lippincott, Springhouse and Ovid; or, more recently, *La Opinion*'s decision to join with *El Diario/La Prensa* to give advertisers to Hispanics better reach for a single media buy.

Aggregation and mergers are nothing new, of course. What is new is the proliferation of sources, and with that the ability to use networked systems to tie aggregated content feeds into Web-based applications. National Semiconductor, for example, aggregates the pricing and inventory data of its distributors into its extranet, so that customers looking to buy its chips can quickly see what distributors have a particular chip in stock and at what price.

A very different example is last year's ROSA project conducted by the Nature Publishing Group with funding by the Joint Information Systems Committee in the U.K. The project participants created Urchin, a tool for aggregating and filtering RSS feeds. The tool, which is now available as open source, pulls metadata into a database in order to facilitate new types of searches. In a review of Urchin, Ben Hammersly, an expert on RSS, described the difference that RDF metadata can have on searches. "For example, a simple text search system, looking at keywords, can do very little other than Boolean queries, such as 'genetics AND carrots,' but with an RDF-based system, you can make much more interesting searches, such as 'genetics AND carrots, published before January 1999, with an author who works in the U.K., and who speaks French.' That sort of search would be tremendously difficult without an Urchin-type system." Because of the growing multiplicity of information sources, Urchin-like tools will be increasingly in demand by both publishers and their customers.

Add value. A related tactic for adding value is to enrich the content itself. The value of categorizing content, for example, is so well proven that it has spawned a whole market of tools for automating the process. But, as the Urchin example above illustrates, there is still much more that can be done to enrich content with markup that will add value to customers.

Another value-adding strategy is to wrap services around the content, of which there are numerous examples across many industry sectors. Newspapers have leveraged their help-wanted ads by creating online career-building sites. Journal publishers have added alerting services and "virtual journals" that let customers build their own "issues" of articles that match user-defined queries. Successful online music retailers recognize that consumers value the ability to make custom CDs. Corporations are leveraging their brand assets by creating self-service extranet sites that enable subsidiaries or distributors to access promotional materials. A few corporations have taken this a step further, creating applications that enable affiliates to build their own customized sales materials based on corporate-branded content. As we look ahead to greater use of RDF for the semantic Web, we expect that adding value by enriching content and relating it to software-based services will be a key content strategy of next-wave publishing.

Embrace e-business

Preparing your content for digital markets is an important facet of the content strategy, but equally important in the next wave will be the business systems that support the sale and distribution of digital content. The strategy will be to establish a completely integrated, collaborative ecosystem that encompasses the complete value chain of content/media producers, owners, distributors, service providers and customers.

A key requirement to achieving this strategy is integrating internal business, management and execution processes. These can no longer be separate (or disparate) systems. Process integration requires integrating the metadata (or knowledge about) these various systems and repositories. That's one reason we're starting to see large media companies embrace enterprise content integration: They already have too many systems to consider replacing them all with one new system, but the strategy demands better integration across different parts of the business.

From the value-chain perspective, trends toward Web services and business process management are important steps toward supply-chain integration. They will lead to the emergence of integrated service platforms designed to facilitate make-it, buy-it, rent-it and share-it options for provisioning systems and services, as well as their effective integration and deployment. Another word for the integrated service platform is "service grid," which we'll illustrate in the Sony case later in this article.

Test Internet business models. For media companies, developing, testing and perfecting new business models will be key components of their next-wave publish-

Fig. 5: Gadgets as media players. Next-wave media devices and players will be more mobile, due to wireless networks.

ing strategies. The models will presume digital networks (both wired and wireless) for content delivery. Executing often requires retooling or partnering, which often introduces new players into the value chain.

Managing rights and embedding them into digital media will be an important aspect of these new models, but markets will not necessarily wait for digital rights management (DRM) technology contests to be resolved. For example, the traditional media business models—buy or subscribe—will be supplemented by "leasing" models that convey temporary, rather than permanent, access to the material. Our study found publishers already experimenting with this model: textbook publishers allowing students to rent an online book for a semester; scholarly publishers allowing researchers to search a collection for a day; libraries loaning e-books to their patrons; corporations granting contractors access to sensitive documents for only a specific period of time.

Digital rights management will emerge as an enabling technology that facilitates use of these new models. Unfortunately, until *de facto* or official standards are accepted, DRM technology will be its own battlefield. The market won't wait for this battle to be settled, but the fight over incompatible standards does dampen market acceptance. Today, this is as true for e-books as it is for online music. Even so, both of those markets are growing and represent important future revenue streams.

Summary

To increase sales and trim costs, those who own content look to maximize its value through various strategies. To maximize returns on content, they micro-segment markets and look for related customers and market needs that can be addressed, and they look for new ways to package and deliver this content using different media formats. These strategies create the need for content product platforms and a new infrastructure, and often lead to new partners as well. We'll turn to those considerations below, following a discussion of next-wave media formats.

Next-Wave Media Devices, Players and Platforms

One of the main trends of the network-services stage of publishing is the proliferation of media and venues for delivering content. These include print and non-print media, Internet and mobile delivery, events and services. As you develop your strategies, what assumptions can you make about the presentation media with which customers will ultimately interact with your content? This section outlines a few key characteristics that will be embraced by emerging media, where data formats, hardware devices and software players and applications combine to form digital media platforms.

A plethora of formats

Media platforms for delivering content-based products and services continue to proliferate, complicating efforts on the part of publishers to optimize processes for specific media. Our research indicates that no consolidation of formats is likely any time soon; therefore, the key to growing sales and profits will be how well publishers can adapt to and take advantage of these new formats.

Virtually all of the new content-delivery platforms that will emerge over the next five years will be driven by digital feeds. Even analog media, such as print, will increasingly be driven by digital devices. However, print and live events will remain important venues for delivering content; therefore, next-wave strategies and systems must support them.

Networked with wireless mobility

As media formats proliferate, they are evolving as well. One vector of change is bandwidth affordability, which favors multimedia formats. In the United States, growth of broadband services to the home (DSL and cable modem) grew by nearly 30 percent in 2003. Market penetration of broadband is even higher in some parts of Asia and Europe.

The ability to break free from the tether of the network cable is a liberating experience, one that is necessary if digital media platforms are to rival print, which for publications like books, magazines and newspapers, has proven quite capable of meeting human mobility requirements. Once digital media platforms obtain this portability, the inherent weaknesses of print—its static nature and inability to carry sound and video—will become more important as differentiators between analog and digital platforms.

A similar pattern, showing the value of mobility and how it can drive the shift from analog to digital platforms, can be seen in the telecommunications sector. In 2003, the number of mobile phones passed the number of terrestrial phones for the first time. Once the digital platform came close to matching the key performance criteria of its analog counterpart (in this case, service quality and reliability), the market quickly embraced the mobile alternative. Fierce competition among handset and service providers has fueled rapid development of rich-media capabilities within these phones, while terrestrial phones have remained listening devices, much the same as they have been since the invention of the telephone.

Multimedia in nature

Single-function analog devices (audio tuners, videocassette players, music CD players) are giving way to multi-function devices (phones that play songs and video, DVD players that have radio tuners, hand-helds with which you can read, play video games and watch broadcasts). The ability to mix media in a single device will drive demand for better content management: Media producers will need to not only organize specific types of content but also keep track of relationships and dependencies between objects of different types. E-mail marketing is now routinely multimedia, just as printed direct-mail advertising has been for many years. As marketers improve the granularity of their campaigns and enterprises strive to communicate with customers on a more personal level, their ability to automatically pull the right content for a specific message will become a competitive advantage. Right now that means pulling the right images as well as the variable text. Before long it will mean pulling the right sound or video clip, as well as the context from the CRM (customer relationship management) system.

That last point can be made for commercial media interests as well. In all of these devices, there will be opportunities to sell content-based services—messages, alerts, audio clips—as well as ads. Mobile-phone companies already view content as an important revenue generator. At Vodafone, for example, premium content represented 15 percent of service revenues last year, and that percentage is growing. The phone service providers—like Internet service providers—represent a new channel of aggregators-distributors with which advertisers, media owners and producers will want to partner. New value chains will be established, and, because of the multimedia capabilities of the device, new partnerships formed. But serving these new audiences will require managing content at a very granular level, with a rich set of metadata surrounding each content object.

Interactive, with code and content intertwined

Another vector of change is the convergence of content and code, which favors interactivity, intelligence and agency as integral dimensions of media delivery. Next-generation media platforms will not only play fixed content; they will adopt the Web convention of allowing code to interact with content. The result will be platforms for "rich media applications"—software applications that have much better graphics and media-handling capabilities than in the past. Because they could be programmed, computers have always excelled in their flexibility. Now, because communication, not computation, is the driving energy of the computer industry, those who make devices for media consumption are paying increasing attention to the aesthetics of media communication.

In 2003, for example, we saw increased acceptance of the Tablet PC and the expansion of Macromedia's Flash platform with the introduction of Breeze and Central. The Tablet represents a hand-held form factor optimized for reading and note-taking—Microsoft's acknowledgment that on-screen reading must be as comfortable an experience for users as it is for paper, if consumers are to start changing their reading habits. In parallel, Macromedia is improving the aesthetics of software applications, using its Flash MX format as the delivery vehicle for software applications that combine video, text, graphics, animations, XML data and chat, all inside a branded experience controlled by the application designer.

No consolidation of formats is likely any time soon; therefore, the key to growing sales and profits will be how well publishers can adapt to and take advantage of these new formats.

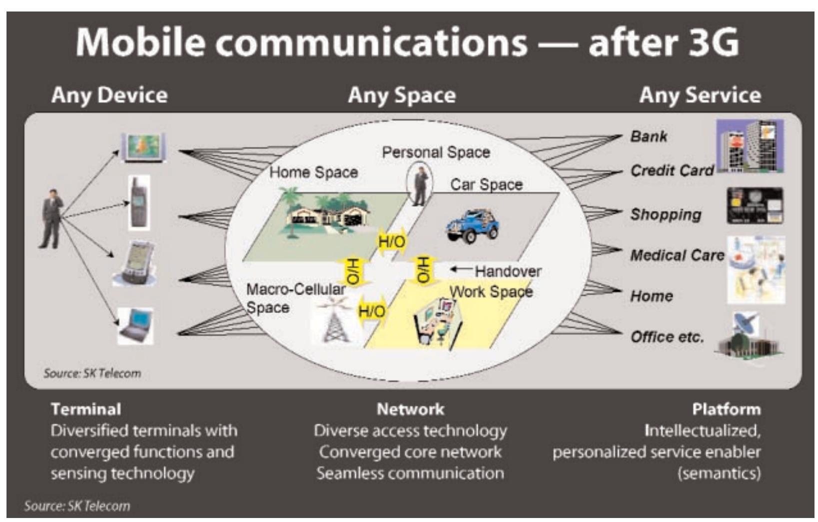


Fig. 6: Mobile communications—after 3G. This chart from SK Telecom depicts mobile communications after 3G (4G and beyond), where broadband converged wired and wireless services can connect any device, to any service, in and between any space, with service levels that start at 1–10 MB/second and go up in speed. The availability of broadband will have a dramatic impact on the platforms for media delivery.

A media platform example that's still a year or two from reaching the market is Microsoft's Avalon, the new presentation layer of Microsoft's next operating system, code-named Longhorn. Avalon blends user interface, document rendering and multimedia player into a single API. Its design goal is to unify code, content and adaptive layout services, so that media-rich applications (ranging from brochures or magazines to interactive games) can be presented on multiple digital devices and form factors.

As media-technology consultant Vin Crosbie put it, "convergence will not mean owning a newspaper and a broadcast station. Convergence will be what comes together at the device the consumer is holding in her hand." The devices will not only be mobile and networked; they will be able to mix multiple media, and they will commingle presentation and code. As we suggested above, digital rights management will be part of that code.

Content Product Platforms

To support content strategies that count on delivering packaged digital goods and services to these new devices, publishers (both corporate and commercial) will have to build repositories of flexible digital content. Surrounding the core repositories are associated services for designing products, producing them, packaging them, ordering them and fulfilling them. These repositories of content (and associated metadata) create the "platform" for managing the life cycle of digital goods—from design and creation through work-in-process to archive. If finished products are split into content components, and if proper attention is given to categorizing, labeling and normalizing them, then it becomes much easier to reuse those components for future products.

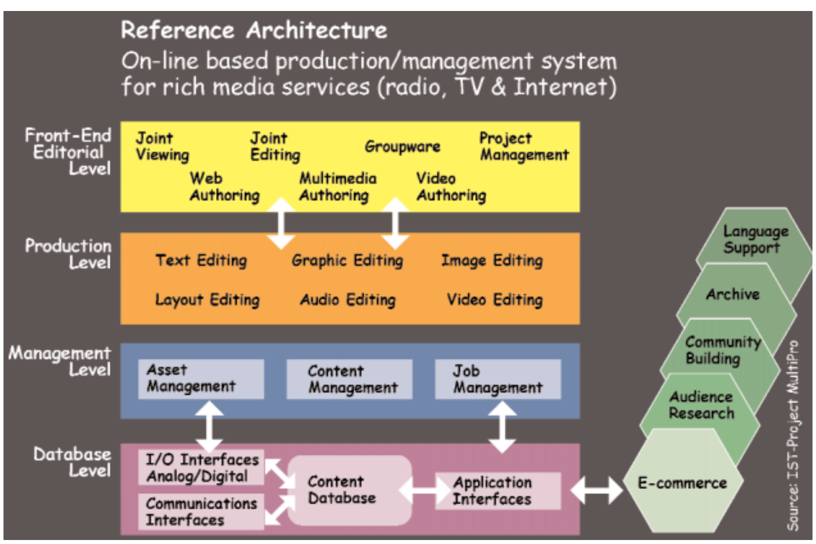


Fig. 7. Reference architecture for multimedia authoring. A simplified view of the reference architecture that IST developed for its MultiPro project illustrates the multiplicity of functions addressed by a complete multimedia authoring and production-management system for broadcasters. Realistically, building such systems requires integrating a variety of specialized tools and subsystems.

For most media companies, this shift represents a big change. Television producers think in terms of shows; publishers think in terms of books, editions or issues. In the past, all of the energy went into getting the product on the air or out the door, with little thought to how any of the pieces might be used in the future. Yet experience across multiple media segments has shown that some attention to reuse can pay significant dividends in shortening time to market and lowering production costs. As David MacCarn, the chief technologist of WGBH, aptly said, "We have to stop thinking about what we produce as the end product and start thinking about it as the first of potentially many derivative products."

Multidimensional

To support the media assumptions above, the product platform must facilitate creation, production and delivery across multiple channels. Channels include not only different types of media but also audience segments within a single media. That means that the databases that store raw digital assets will sort, label and store them at a more granular level than their bundled, final-product form. Labeling will be more complex, along multiple taxonomies oriented toward different markets, fields or classes of customers.

Broadcasters. A number of broadcasters are working on multimedia content-product platforms. In the U.K., the BBC has tested a new rendering engine that automatically adapts broadcast content for different types of carriers. In Europe, the MultiPro research project funded by the Information Society Technologies (IST) Programme of the European Union created a reference model *(see Fig. 7)* and demonstrated the feasibility of a multi-channel authoring and production system for broadcasting. In the U.S., WGBH, worked with Artesia, Sony and Sun Microsystems to define a reference architecture for digital asset management in support of multimedia authoring. (In addition to its television broadcasts, WGBH creates VHS cassettes, DVDs and about 40 percent of the content for the web site of the U.S. Public Broadcasting Service, www.pbs.org.) In addition to using the architecture in-house, WGBH worked with Sun to create a DAM R&D center in Boston that is open to other broadcasters as well. In Washington, the Discovery Channel is using its Artesia system to create a digital library of reusable footage that saves production costs, not only on distribution to multiple media but also in the production costs of new shows that can now consider previous clips instead of shooting all new footage.

Professional publishers. One professional publisher that has already implemented a content product platform is Congressional Quarterly. As the company name implies, the flagship publication of the publisher known for its coverage of the U.S. Congress was once published on a quarterly basis. But the advent of the Web in the 1990s accelerated customer demand for CQ news and information in electronic form. Each new product was developed discretely and took many months to design and launch. In 1999, CQ embarked on a $3 million project that created an integrated product development platform that would support digital products. The project consolidated 17 different product-centric databases into one logical repository of XML-tagged content. That core repository was then connected to a variety of systems for Web publishing and to Quark QPS for print production.

Today, more than half of CQ's revenues come from online products, which typically have better operating margins than its print ones. More importantly, CQ is now well positioned to respond to changes in the market. The editorial and marketing teams can rapidly prototype potential products, extracting sample content from the database and laying it out in a proposed context to show to potential customers or partners. The staff can also adjust the code on the Web site to change real-time customization and personalization based on individual customer requirements—in effect, implementing a just-in-time manufacturing model for content.

In CQ's experience, developing staff expertise is as critical as selecting vendors when building this type of platform. "Playing in this sandbox does require the right tools. But a first-class development staff is an integral part of that toolset," said Larry Tunks, who has fostered a multidisciplinary approach to project teams. "We've blurred the lines of expertise, but maintained crystal clear lines of responsibility," said Tunks. "We've also been successful, in part, because we have the collective buy-in of the executive, editorial, technical and sales/marketing teams."

At McGraw-Hill, Jim King, CIO and Senior VP of Information and Media Services, is leading an effort to create a similar digital product platform for content-based products and services. At the core is a digital-asset-management repository that is linked to an application and service platform, an "e-fulfillment platform" (for e-commerce and digital rights management), a customer-relationship management platform, and a common delivery platform *(see Fig. 8)*. Collectively, this core platform supports multiple product factories that develop intranet direct feeds, analytics services, custom solutions packages, data-mining services, pay-per-use packaging, syndication products, online product sites and offline publication products. King's plan is to implement this conceptual vision as a federated, service-based architecture. The digital-product platform will support McGraw-Hill's business lines, for example providing a common infrastructure across all of McGraw-Hill's content delivery to the construction industry.

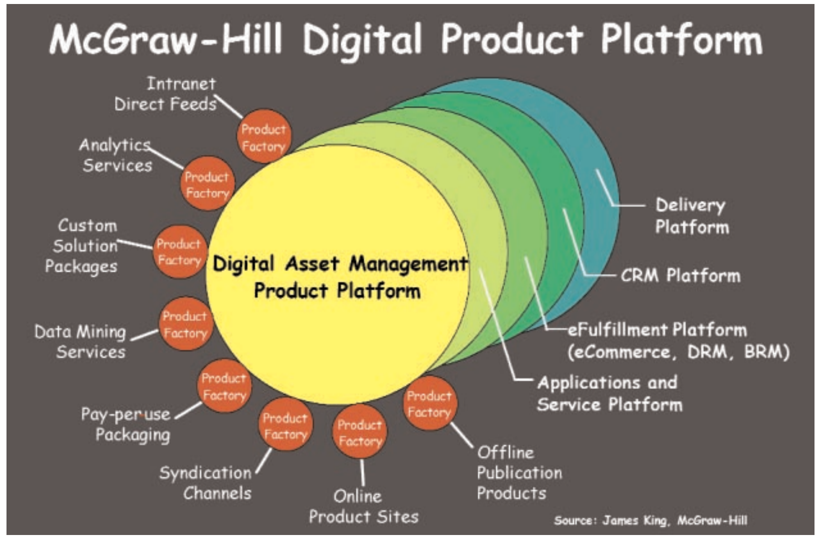


Fig. 8: Prototypical product platform. The core repository of digital content objects is accessed by production systems ("factories") that create products and is supported by business systems that manage customer relationships, fulfillment and delivery.

Other sectors. The same philosophy applies in many enterprises in other sectors of the global economy. In heavy manufacturing, such as large transportation equipment, most products are made to order, and documentation must be tailored to the configuration of the order. Unless there is a database of components at a sufficient level of granularity, there will be a labor-intensive process of modifying the manuals to fit the order.

In service sectors, including insurance and financial services, the inability of the organization to flexibly and dynamically change the documentation (in this case policies or reports) or its marketing materials often constrains its ability to deliver customized services to the client. A personalized multi-channel communication approach enhances customer relationships and improves the rate at which the sales team converts prospects into customers. For example, in a direct-marketing campaign for its training classes, Hewlett-Packard switched from conventional blanket mailing to a personalized campaign that combined personalized printed pieces with a personalized Web site. Its response rate more than tripled, and revenues were ten times the sales target.

Integrated with e-business

In the past, content management implied a close association with a publishing production process. In next-generation publishing, products will be built much more quickly—in some cases, on the fly in response to a customer inquiry. For example, a customer visiting a publisher's Web site is encouraged to search for items of interest in the online catalog. From the search results, the customer selects items to buy and orders them. The expectation is that digital goods will be packaged and delivered immediately.

Harvard Business School Publishing has built a digital publishing platform that works in this way. It pulls the metadata for the online catalog from a Documentum repository of finished goods into a Web-based ordering, fulfillment and delivery system. For Harvard Business School cases, HBSP applies Sealed Media digital rights management on the fly to the Acrobat file as a customer orders it from the Web site. By not applying the rights restrictions to the cases ahead of time, HBSP has flexibility as to the kinds of terms it is able to offer the customer. The terms can vary depending on the type of customer (professors versus students or business users), and pricing can vary according to usage restrictions, such as how long the customer has access to the case or whether she can print as well as view it.

Network-ready for collaboration

The platform will not only be Web-accessible, it will operate across multiple sites and support Web services; in other words, it will be able to communicate not only with internal systems but also with those of external partners and suppliers. It will support multiple participants, spread across multiple stages of the life cycle of content and products, fostering collaboration among different participants and organizations. Most vendors have already begun to retool their systems for this architecture; within two years we expect it to be commonplace functionality.

We found evidence of this characteristic to be quite strong in the digital asset management arena. General Motors, Chrysler, Warner Brothers, Coca-Cola and other corporations are taking DAM systems that once served a single department and making them accessible as services to the larger enterprise. The group managing the assets can collaborate with internal staff and outside creative agencies during creation of the assets while also streamlining the distribution of finished goods (logos, product shots, promotional documents) to subsidiaries, distributors and other product groups within the company. In the case of the Discovery Channel, they've extended their DAM system from archive of finished goods into a Web-accessible collaboration platform that enables staff at its headquarters outside Washington to review footage the same day it is shot on location in Asia.

Service Platforms and the Shared Resource Architecture

What sort of technical infrastructure is needed to achieve these process transformations? Next-generation publishing will be supported by the technical framework that we outlined in Part 1 as the "shared resource architecture." It's a platform built according to a service-oriented architecture (SOA) for content-intensive business processes and value chains. Ultimately, it results in content-product platforms as just described. But underneath there are important building blocks—services and utilities built on top of standards.

In Part 1, we outlined its components: standards, service utility grids, and process and application services. Now let's look at some case examples.

Standards

Standards are at the bottom of our model, because they are the foundation on which service architectures are built. Though unglamorous, the political work of standards-setting can pay dividends to the publisher.

Time Inc. A case in point is Time Inc., publisher of more than 60 magazine titles that reach more than 250 million readers worldwide. Time's publications receive content and advertisements from varied sources, including corporations, advertising agencies, creative services and authors. Editorial, design and production workflow takes place in multiple locations. Its supply chain for manufacturing and distribution of publications includes multiple prepress services, printers and distribution services. Its roadmap calls for connecting all portions of the publishing supply chain together in a collaborative, all-digital production workflow that will speed time to market, eliminate costs and improve profits for all parties.

The first element of Time's approach is an active commitment to open standards and industry specifications, which it seeks to make an integral part of an industry-wide service grid. The company has actively participated in magazine industry initiatives relating to color, digital workflow and business data interchange.

Also, the publisher has advocated use of color-managed PDF workflows and remote digital proofing for magazine workflows. Growing industry acceptance of this practice is displacing film and analog proofs (physical media) from the workflow. This represents hundreds of millions of dollars in potential cost savings across the industry as well as approximately a 50 percent savings in cycle time for affected stages of the process.

The publisher's initial application of network-shared services was to provide managed transport of advertising pages and magazine production masters between advertisers, ad agencies, its publications, print producers and distributors. Its primary motivation was a need for security, speed and reliable quality of service for transfer of time-critical content across the supply chain. After gaining some experience (and overcoming some culture shock), the publisher is gradually eliminating preparation and transfer of physical media from the workflow, resulting in significant hard cost savings and cycle-time reductions that enable the publisher to close editions later, and thereby extend its deadlines for advertising and late-breaking news.

The publisher next extended its use of shared services to include content-management services that integrate with the distribution network. It established a comprehensive content repository to support production workflow and delivery of digital magazine masters to external vendors.

To this platform the publisher is adding collaborative workspace services that provide user-ID-enabled access to digital assets, hot folders and content workflow services, such as file preflight and certification and remote proofing. These services can be accessed across all modes of linkage and include provision for disaster recovery.

Service grid and Web services

Large organizations are building grids of digital services, layered on top of standards, that provide a way to integrate business units and their applications. Like a power grid for electricity, these corporate digital utility service grids are always on, a resource that can be tapped into by internal business groups as well as external partners, suppliers and customers.

Sony Broadband. Sony's Broadband Services Company, a unit of the Sony Corporation of America and of Sony Corporation's Network Application and Content Services Sector (NACS), oversees broadband-related businesses and develops common e-business platforms for the U.S. market. It also creates new service businesses utilizing Sony's content and technology assets, and it explores business opportunities for wireless, Internet and digital services. In a year-long project that went live with its first service in February 2003, Sony Broadband Services created an architecture and platform based on Web services that neatly illustrates the service-grid concept *(see Fig. 9)*.

Broadband Services is a corporate function, supporting three Sony operating companies in the U.S. (Sony Electronics, Sony Music and Sony Pictures). Each of the business units has large development budgets to support its own unique requirements, but the three also have much in common.

The service-grid approach enables a large company like Sony to continue to distribute technology responsibility to individual business units while, at the same time, reducing the number of overlapping and redundant systems. Each business unit is allowed to keep its legacy systems and remains free to implement best-of-breed solutions specific to its business. But the shared infrastructure provided at the corporate level creates cost and time-to-market incentives for business-unit developers to leverage, rather than ignore, software deployed elsewhere in the corporation.

In Sony's case, the service grid also brought greater consistency to the customer-facing aspects of its e-businesses, which previously had been developed independently by each business unit.

Grid computing. As the name implies, the utility grid will leverage grid computing—an emerging technology that distributes processing loads across a network of computers. By distributing the processing load, companies can lower the capital costs of implementing high-performance computing. Grid computing may help smaller companies compete with bigger players in established media markets. As an example, AXYZ Animation (www.axyzfx.com), a Toronto-based animation studio with a few dozen employees, has been using grid computing since 2000 to produce renderings of its artists' drawings.

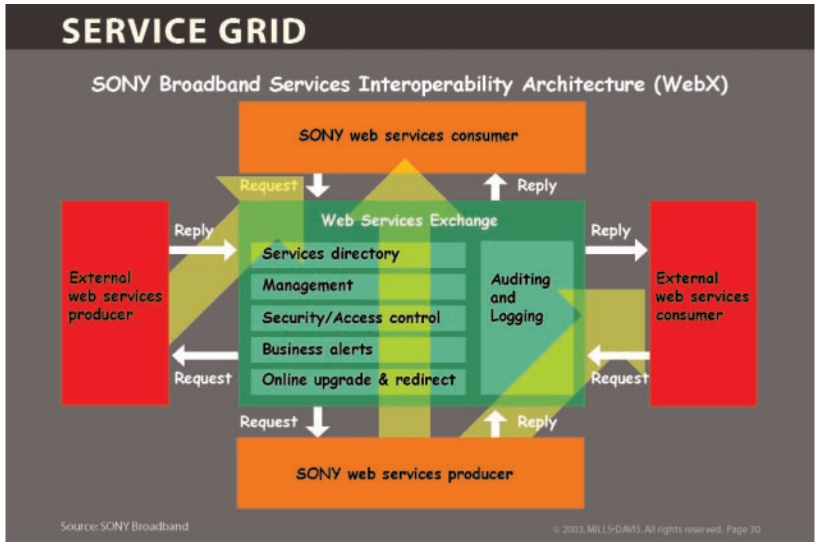


Fig. 9: Service grid. The Sony Web Service platform, code-named WebX, was deployed using Blue Titan and BEA WebLogic Server. Blue Titan Network Director sits between the network layer and the layer of applications. WebX provides a centralized set of utilities and services for managing resources and transport, as well as management of the services, including monitoring and versioning of services. WebX also provides content-based routing and message transformation, insulating business-unit developers from inconsistencies of schemas and protocols.

AXYZ creates special effects for television ads for big names, such as Disney, GMC and Kellogg's. Depending on the pace of the work and the deadlines, the number of frames to be rendered on any given day can vary tremendously. Predicting workloads accurately is also made more difficult by the creative process, in which the advertiser and its agency often request unanticipated changes. Because of the compute-intensive nature of the rendering process, rendering jobs would tie up animation work during the day and were frequently left to run overnight.

AXYZ customized Sun's Grid Engine to dedicate one processor of each of its dual-processor animation workstations to rendering. Now rendering takes place in the background all day long, in parallel with the work of the animators. The animators are more productive, and AXYZ is able to handle jobs that might otherwise go to its competitors.

Private Web services

In our model, above the core corporate service grid are Web-based services that enable specific business processes. These may be private, public or somewhere in between. Increasingly, we see the use of Web services enabling new ways for those who own content to connect with service firms, such as agencies, repro houses, printers and distributors. Point-to-point integrations are being replaced with architectures that support *n*-to-*n* integrations among internal and external systems, a prerequisite for adapting the business to new markets and new supply chains. In effect, even "private" services are no longer just in-house. A survey of 75 large corporations published by Forrester Research last fall found that almost three-fifths of those planning customer service initiatives and more than two-fifths of those planning supply-chain projects were using Web-services technologies.

The overall effect of Web services will be to make boundaries more porous. Enterprising businesses are already testing the model.

Butterworths. One example that went live this year was a private connection between Butterworths, the U.K.-based legal publisher, and Letterpart, its composition house, also in England. Prior to the new system, Butterworths followed the timeworn practices of book publishing: Editors marked up manuscripts on paper and sent the manuscripts to various compositors for keyboarding and page composition. Two weeks later, page proofs would arrive back from the compositor for review; changes would be marked on the pages and mailed or faxed back to the compositor. The cycle took weeks to complete, with delivery times eating up days for each publication. Butterworths switched to a digital workflow, based on converting the manuscripts upfront to XML. Today, editors submit jobs to the compositor over the Web from their desktops. The compositor, running a XyEnterprise composition system with Web services, receives the job, automatically composes it against the indicated style sheet, makes a PDF and returns the PDF to the editor's browser for review. The turnaround time for receiving initial page proofs went from two weeks to seconds, and subsequent corrections and edits are also dramatically faster to complete. The traditional publisher-supplier relationship is maintained, but now the composition facilities of Letterpart are accessible digitally, making an outsourced service behave like an inhouse resource.

Corbis. Another example comes from Corbis, the stock photography supplier. Last year, the company started using Web services to streamline the process of launching new Web sites. The payback has been a dramatic shortening of the time and lowering of the costs of developing new sites. Without the new architecture, it took 10 to 12 developers a year to create a customer-facing site, at a cost of \$5–\$7 million. Since implementing the new architecture, Corbis has found that three developers can create a new site in four months, saving the company hundreds of thousands of dollars per site and shortening its time to market. Equally important, from a strategic vantage point, Corbis is now able to pursue micro markets that previously were prohibitively expensive to target. It is making aspects of its Web site (*e.g.*, search, catalog, lightbox, shopping cart) available to business partners that target niche audiences that Corbis is not reaching directly.

Publicly shared services

Because they run over Internet protocols, Web services may be deployed publicly. Today the directories for discovering such public services are not yet well established, but that hasn't kept enterprising businesses from testing the model.

Blurring boundaries. One example of shared services that blur boundaries among publishers and their suppliers comes from the journal publishing community. In the past three years, more than 1,000 journals have exposed their manuscript-submission and manuscript-tracking processes to authors over the Web. In the large majority of cases, the publisher (often a nonprofit society) does not run this service itself. The service is either hosted for the publisher by a vendor, such as ScholarOne or Aries Systems, or in some cases, by a large commercial publisher (such as Wolters Kluwer) that runs the software internally and will host the service on behalf of a publishing society. If the branding is done well, authors should feel as though they are interacting with the publisher, not with the host system provider.

That same transparency is now being extended to related processes that may be hosted by different vendors. A good example is the integration of the Digital-Expert image-preflighting tool from the Sheridan Group into ScholarOne's manuscript-submission and tracking service, Manuscript Central. Rather than write its own image preflighting tool, ScholarOne used Web services to seamlessly integrate Sheridan's Digital-Expert, which the printer developed as a way to attract journal publishers to its editorial and production services. Through this integration, the society publisher is able to expose an internal process (in this case, checking to see if accepted art meets quality standards) hosted by a printer and, at the same time, to tie that process into a larger one (submitting and tracking manuscripts through the peer-review process), which is exposed and managed by a separate third-party vendor. From the outside, the entire process appears to be hosted by the publisher. From the inside, the integration of software and processes are made much simpler through the use of quasi-public Web services.

Growing ecosystems. A third example of public services are the APIs that have been made public by Google, Amazon and eBay. By exposing their internal systems, these leaders of the Net economy are creating new ecosystems of small businesses that rely on them.

In Amazon's case, for example, it viewed Web services as a way to enable associates to build better and tighter links to Amazon's online store—enhancing the partner's site and sales, which in turn drives revenue for Amazon. For third-party sellers, Web services provides a way to plug into Amazon's product information. Dealers of used goods are now using the Amazon and eBay APIs to value items they buy and sell. Developers such as Anacubis are writing specialized search applications that query both Amazon and Google (http://www.anacubis.com/amazondemo/), showing, for example, what news or Web sites are saying about an item that you've found in an Amazon store.

Business process management

Above the layer of software services are business processes, governed by business process management (BPM) systems. As Howard Smith and Peter Fingar define it in their book, *Business Process Management, The Third Wave*, the BPM system "enables companies to model, deploy and manage mission-critical business processes, that span multiple enterprise applications, departments, and business partners—behind the firewall and over the Internet." Smith and Fingar predict that these systems will arrive either from vendors specializing in BPM, like Intalio, or as part of the IT infrastructure that companies already have in place.

Business process management has been tried before, but this time it will be different, because instead of codifying existing processes, the next generation of BPM systems will facilitate process change. The drive toward customization of products, and the multiplicity of channels over which companies will have to communicate, will force companies to become adept at changing their communication processes. As Howard and Fingar aptly put it: "Creating a new process is not enough—companies must be able to efficiently manufacture many process variants." A new breed of systems, complemented by emerging standards—will make it possible for companies to change the way they communicate, even as they continually change the products and services that they sell.

LexisNexis. In the media industry, business process management systems are catching on first in the non-creative areas of the business. We see their adoption in the back-office of media companies as paralleling that of other industries and as a precursor of what will eventually come: BPM for creative processes.

For LexisNexis, one of its challenges to growing sales has been the inefficiency of corporate licensing. Although the company has standard pricing, salespeople have had considerable wiggle room in negotiating major deals, often without a clear sense of the cost implications of making special provisions. Every time special terms were suggested, other people at LexisNexis had to become involved in order to determine these costs and to write a custom contract; the established fulfillment system wasn't set up for dynamic pricing models.

For large deals, the attention of so many people was worthwhile. But as LexisNexis went after mid-tier buyers, the inefficiency of the process became a stumbling block. There were opportunities, but the company needed a way to roll out flexible pricing on a wider scale without getting burned by salespeople making overly aggressive deals.

With help from BPM specialist Intalio, LexisNexis abstracted its pricing rules into a separate layer of software. Now, salespeople can price a package within a range of limits and see what kind of discount is preapproved and what will require special approval. LexisNexis cut costs and time associated with human review of sales contracts, and the finance department can now monitor the orders and better manage the discount levels. Thus, the BPM system is not a fixed system that covers a static product and process, but rather a dynamic system that helps the organization adapt its processes (and pricing) over time.

Redefining the Value Chain

When we take a step back from the details of what and how we will publish and look at who will be involved and what tasks they will perform, we see that in many cases there not only are new players, but also that the roles of established players change in this next stage of publishing.

A value chain is a connected series of organizations, resources and knowledge streams involved in the creation and delivery of value to end customers *(See Fig. 10)*. This concept is not unique to publishing. Value systems integrate supply-chain activities in all industries, from determination of customer needs

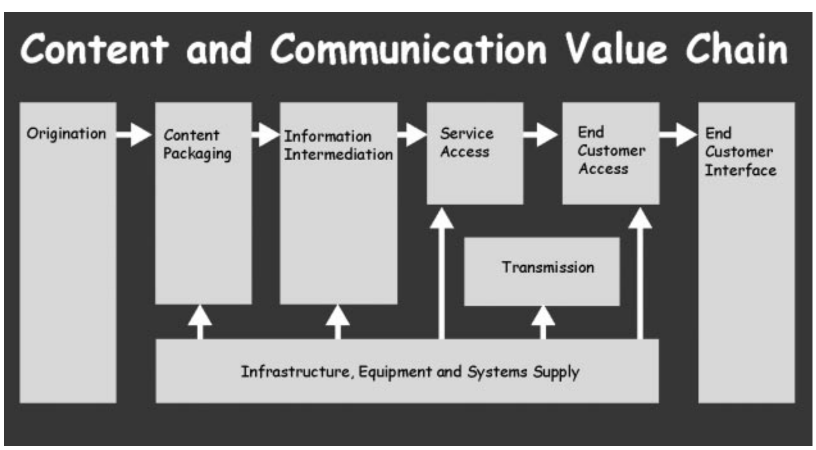


Fig. 10: Content and communication value chain. The illustration depicts a basic sequence of activities for next-wave publishing, which encompasses Internet, mobile and conventional media. Value-adding activities are shown from left to right in the boxes labeled origination, content packaging, information intermediation, service access, transmission, end-customer access and end-customer interface. The lower box, labeled infrastructure, equipment and systems supply, is the enabling support for this chain.

through product/service development, production operations and distribution, including (as appropriate) first-, second- and third-tier suppliers.

Next-wave publishing technology and process will redefine value chains in multiple ways. Most companies will experience changes in:

- Content sources they publish;
- Audience needs they address;
- Media channels they distribute;
- Role(s) they play in the value chain;
- Roles that others play in the value chain;
- Functions they outsource and insource;
- Providers of outsourced capabilities they use;
- Internal deployment of functions;
- External deployment of function;
- Level of collaboration (and of process integration).

The changes are outlined in Fig. 10. To expand upon the terms in the diagram, we note that:

- Content originators include authors, publishers, libraries and enterprises. Content can be either pre-existing or created. Published content is pre-existing, hence acquired, for example via syndication. Originated content is created and then acquired via the information network.
- Content packaging puts content into a form that can be delivered in one or more media formats. Editing, design and production are typical functions. Packaging is specific to the constraints of chosen media channels.
- Information intermediation stores and forwards content. Also, it may transform or customize content and packaging to the preferences of the end customer.
- Service access is the process of connecting customer demand with content supply via a delivery channel. In conventional publishing, retailers may play this role. Over the Internet, it might be an ISP, cable operator or wireless provider.
- Transmission is the way content is delivered to the distribution point. There are multiple players and choices for transport, both physical and digital. Transport is usually invisible to the end-customer unless it overly constricts an end-customer's content experience.
- End-customer access is how content comes to the point of end use. In a store, access is about shelf space and format of the media product. For mail order, e-store or home delivery, service access is virtual (*e.g.*, ordering via a catalog) followed by delivery of the product. In a 4G world (next-generation, high-speed converged mobile and terrestrial digital communications), where anything and everything travels over Internet protocols, access may be via a home entertainment system, a mobile device or a computer, or may be embedded in any number of gadgets or products and appliances we use daily.
- End-customer interface refers to the media format and interactivity that determine a customer's experience of the content. Readers have well-formed expectations of how to interact with printed products. Similarly, the conventions for graphical user interfaces for personal computers are broadly understood. Game stations, cell phones, digital gadgets and consumer electronics all have their own user interface conventions, all of which are still evolving.
- The infrastructure, equipment and systems portion of the value chain includes organizations that provide the hardware and enabling services that power operation and management of the next-wave publishing environment.

Consumer games: bandwidth redefines a value chain

A key technology driver for value chain realignments is the bandwidth of communications. To glimpse the future of high-bandwidth consumer telecommunications, just look to Asia—and specifically at what's happening today in South Korea, where more than half the population has had broadband access for more than three years. During the three-year period of 1999-2001, when broadband was widely adopted in South Korea, online gaming grew by 650 percent and its share of the overall gaming market rocketed from 2 percent to 21 percent. As North America and Europe follow South Korea's lead in adopting broadband, the changes in gaming that South Korea has witnessed could signal how consumer entertainment will change throughout the globe. Transitions in gaming may presage value-chain shifts that will become common across other areas of publishing. As a media format, games have a close relationship between advancement of the medium and advancement of its underlying technology and production processes.

From boxes to services. Traditionally, the value chain for software video games has been similar to book publishing, with publishers and developers (such as

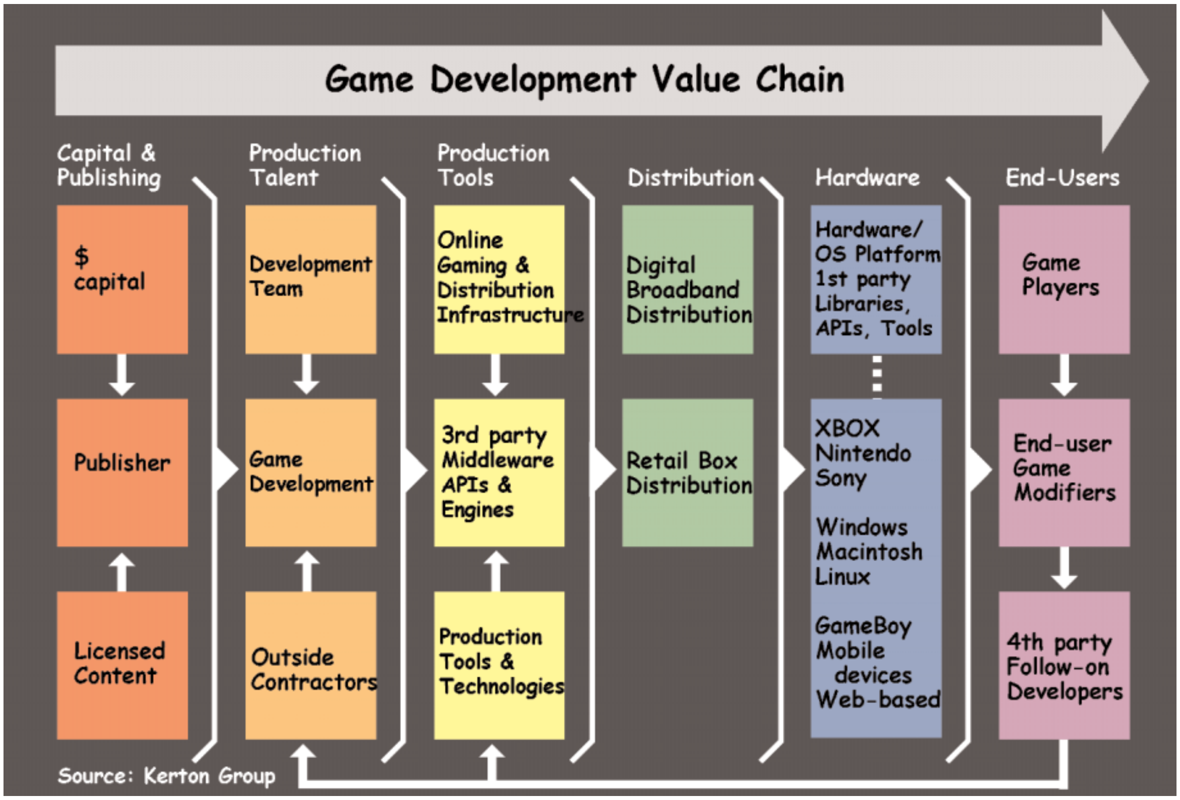


Fig. 11: Game development value chain. The illustration depicts the value chain for game development. Starting from the left:

- Capital and publishing refers to all sources of investment, publishing entities and third-party intellectual property that might be licensed for a particular game title.
- Production talent includes the developers, designers and artists. Here the pendulum swings between teams that are under contract (rather like the movie industry), to inhouse developers, to outsourced specialists (like other publishing endeavors).
- Production platform refers to the suppliers of content production tools, game-development middleware, customizable game engines and production-management tools. It also includes physics libraries, asset management, 3D rendering engines and interactive audio, as well as interfaces to online gaming infrastructures. The trend in production platforms is toward commercial tools rather than custom-built ones; separation of "game content" from "game-engine" (which echoes the rise of databases and business-process engines elsewhere); and integration of production tools with tools for end-user modification and usage, as well as online distribution.
- Distribution includes providers of broadband distribution and online gaming delivery, as well as retail box distribution.
- Hardware refers to the suppliers of arcade equipment, gaming consoles, PCs and mobile gaming devices.
- Lastly, end-users refers to roles that consumers are playing in the value chain as players, modifiers of games and, more recently, "fourth-party" developers.

Electronic Arts) and retailers (such as Electronic Boutique) capturing about 70 percent of the money spent for a boxed product. However, with the emergence of broadband Internet infrastructure (and, later this decade, 4G converged high-speed wired and wireless digital communications), the value chain is shifting significantly. Internet downloading will become the dominant distribution mechanism. Broadband interactive gaming will emerge as a major entertainment format. And games themselves will become platforms for continuing development by game players as well as game development teams *(See Fig. 11)*.

Several trends are impacting next-wave consumer game value chains, which could have significance for other areas of content and publishing.

Games began as custom hard-coded, closed-end experiences for a single end-user machine. The trend today is toward multiplayer games built as a layer of content and logic on top of fundamental engines that can be repackaged for simultaneous distribution across multiple player platforms and global markets. This approach separates representation of core content, character behaviors, levels of complexity, digital effects, graphics, sound and product-delivery infrastructure, and it standardizes the toolsets and building blocks.

Many new games are increasingly becoming platforms unto themselves—"open shell" game-styled operating systems upon which entirely new productions can be professionally developed by third-party developers. There is even a trend towards generic (open or public source) engines for developing staple game types, such as real-time strategy, turn-based strategy, racing and the ever-popular shooter games such as CounterStrike, Diablo, StarCraft and Warcraft

One value-chain implication is that game development will become more horizontally differentiated:

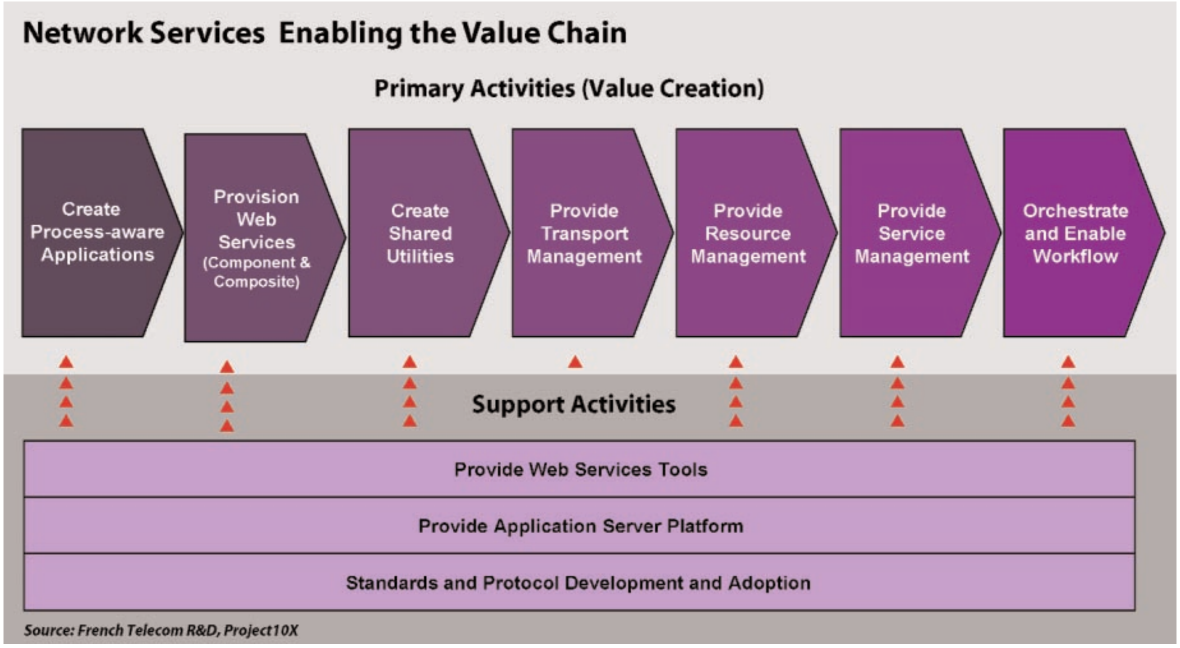


Fig. 12: Network services enabling the value chain. The drawing depicts ten categories of services that network service providers can provide in order to integrate themselves into new content value chains.

that is, more B2B collaboration, more outsourced specialists, and more transient development teams, rather than inhouse employees. In this sense, consumer games will become more like movies. Also, as game development and distribution platforms become more integrated and standardized, the studios or development houses will become more like "editing shops," where the elements of content they develop become separately packaged in a way that better exposes them to the overall value chain.

This enables selling game features individually (*e.g.*, difficulty levels), reusing character content between game versions, and developing new games that exploit content, game platforms and distribution solutions already developed. In fact, the trend will be towards continuous development over a life cycle consisting of a number of versions, add-on products, extensions and user-developed content.

The user-developed component is especially important in massive multiplayer online role-playing games (MMORPGs), where gamers invest time and energy in creating their characters and imaginary worlds and then get emotionally invested[1] in picking up where they left off. Because of this emotional investment in keeping the game-state persistent, MMORPGs, such as EverQuest or UltimaOnline, are typically sold by subscription, with recurring royalties for the publisher and distributor.

Seeking ways to monetize product derivatives is not new. What is different about consumer games is the entrance of new members into the value chain. One of those is "fourth-party" developers, the 1–2 percent of gamers who develop modifications or extensions for a released game and who will participate in the revenue stream, especially as off-the-shelf toolsets and product platforms become integrated and standardized.

Global growth of pervasive gaming experiences is another key trend. Increasingly, products will span platforms and utilize mobile and wireless technologies to reach people wherever they are, whenever they want. As broadband becomes the preferred method of game distribution, new value-chain opportunities arise. In addition to the retail store, the game developer may distribute directly to the end user via e-commerce sale or monthly subscription. In South Korea, developers distribute to owners of more than 25,000 PC cafés, where gamers play via a pay-per-use model similar to an arcade. A third channel will be broadband service providers eager to increase revenues through content-based services. They will offer games and take a slice of the revenue, whether the model is an initiation fee, a subscription or pay-as-you-play.

Two considerations in these business models are the availability of broadband services and the level of infrastructure investment needed to support online gaming for multiple players. In South Korea, Korea Telecom and Hanaro Telecom successfully targeted their DSL services at online gamers, offering them a pay-as-you-use model where gaming charges are aggregated onto their monthly DSL service bills. In North America, companies like K2 Network are introducing another new member of the value chain: "game aggregation services" for broadband providers. These aggregators provide central and regional centers that host game operations and ensure seamless integration between the points where games are created and where games are played. For a \$10/month subscription, the revenue split might be \$3 to the game developer, \$3 to the network service provider and \$4 to the aggregator.

Change in other sectors

In virtually every sector of the global economy, the value chains of graphic communications are shifting. Here are but a few examples:

[1]In late January, a lawsuit was reportedly filed in China seeking damages for the loss of game character development due to a server crash.

Retailing. In retailing, the WorldWide Retail Exchange (www.wwre.com)—a nonprofit cooperative founded four years ago by 17 retailers—has saved its members more than $1 billion by eliminating inefficiencies in the retailing supply chain. WWRE sits as a digital hub between manufacturers and retailers, reducing the complexity of the many-to-many relationships that retailers and suppliers establish with one another. As part of its services, WWRE operates a globally accessible digital asset management system, where concept designs can be shared and product shots can be dropped off and picked up by retailers for use in their ads, circulars and in-store displays. The system is saving not only money but also time, shortening product-development and publishing cycles, which gives retailers a later deadline for preparing their ads.

Telecommunications. In telecommunications, voice telephony has become a commodity. Companies are looking to high-bandwidth services to redefine customer expectations and open the way for them to participate in the revenue streams of new value-add services. These new services will include a variety of utilities and activities, as illustrated in Fig. 12.

Printing. Demand for offset printing is shrinking. To counteract that trend, smart printers are redefining their businesses. They are expanding their commercial printing or document-manufacturing businesses to include management of clients' printed inventories and even management of marketing campaigns. They are tying real-time, on-demand digital printing to Web-based transaction and tracking systems that link directly to corporate customers' desktops. The automation of a digital supply chain saves the customer time and money and enables the preferred partner to increase its share of its customers' printing. The strategy is paying off for large companies, such as Moore-Wallace and Quebecor, as well as for smaller printers, such as LaVigne in New England.

Conclusion

Global 5000 corporations buy more than 50 percent of all printing and media purchases. Faced with shorter product and information cycles yet ever-increasing numbers of media channels, they are struggling to reduce the cost of media spending and to speed time to market. At the same time, they want to improve the effectiveness of their communications—fewer errors, more consistent branding and messaging, and better returns for marketing campaigns.

Historically, the systems large corporations used to create content, and the processes surrounding the management, procurement, production and distribution of content, were not well integrated with other systems in the enterprise. Today, in an effort to make their core business processes more efficient, many of these corporations are linking their business systems to their content systems and media-communications processes. Thus, cost-accounting systems are being tied to print procurement, and print-on-demand media delivery is becoming a component of a larger move toward personalized, relevant communication with the customer. The network-services architecture provided by next-wave publishing technologies is facilitating these integrations and will create many opportunities to redefine the content value chain in the process.

Commercial media companies are changing as well. They are taking a multi-channel approach to their products and services, building new content-product platforms and forging new partnerships in support of their strategies to penetrate new media. In the years ahead, they, like corporate publishers, will extend their ability to dynamically make new media products by developing an ability to dynamically change their internal processes in response to customer demand.

The next wave of publishing technologies—just like previous ones—will have a dramatic impact on both types of publishers and on all of their suppliers and partners as well. As we've tried to illustrate here, the network-services stage of this next wave is hitting all sectors of industry, washing across all types of visual communications, from consumer-oriented ads, games and media to business publications, Web sites and information services. Its reach is global; its impact will be both liberating and devastating. The examples we cited above show that the revolution has begun. It will take time to play out. But in the years ahead, next-wave publishing technologies will transform publishing processes, redefining our businesses and how we go about them. And in running its course, this wave will set the stage for the one that follows—the revolution in content. **TSR**

About the authors

Mills Davis is managing director of Project10X. A researcher, visionary and industry consultant for more than two decades, Davis focuses on improving profitability through process innovation, technology investment and strategic partnerships. Previously, he directed the Digital Roadmaps Project and the HiFi Color Project. He can be reached by e-mail at mdavis@project10x.com or by phone at 202-667-6400.

Mark Walter is senior editor of *The Seybold Report* and the lead Seybold analyst on the Next-Wave Publishing Technologies project. Contact him at mwalter@seyboldreports.com.

Look for our conclusion to the series, "Part 3: Revolution in Content," to appear in an upcoming issue.

The Seybold Report℠

Vol. 3, No. 23 • March 15, 2004

Publishing Strategies

Next-Wave Publishing, Part 3: Revolutions in Content 3

OUR SERIES ON THE NEXT WAVE OF PUBLISHING TECHNOLOGY concludes in this issue with a look into the future, when semantic computing revolutionalizes digital content. The arrival of semantic computing heralds the dawn of the knowledge age, in which new vistas will open for publishing, information technology and manufacturing.

Semantics and smart data. What is semantics all about? We open the article with an introduction that chronicles the evolution of data from the dawn of computing through the current decade.

Semantics for content. How will semantics improve our content? How do we move from simple searches to knowledge-based retrieval? We detail the progression and characteristics of knowledge representations, from glossaries to taxonomies to ontologies, and outline the stages of constructing an ontology.

Semantics in computing. The W3C posits that we need semantics to make the World Wide Web work for machines as well as for people. But the implications are even wider: Knowledge will trump programming, leading to new forms of computing.

Semantics in publishing. What role will semantics play in product strategy? We predict that publishers stand to win big in the next wave—if they can adapt their skill set to semantic content.

Semantics in government and manufacturing. As was the case in other areas of technology, government agencies and large manufacturers are a hotbed for testing the practicality of cutting-edge theories, in this case semantic-content applications. The stakes are large—but so are the potential paybacks.

ALSO IN THIS ISSUE:

For news between issues, visit:

Seybold365.com

Seybold
SEMINARS PUBLICATIONS • CONSULTING

Business & Editorial/Production Office
428 E. Baltimore Ave., PO Box 644
Media, Pennsylvania 19063, USA
phone: (610) 565-2480
fax: (610) 565-4659, 565-3261
web: www.seyboldreports.com

editor: Peter E. Dyson (phone x122)
e-mail: pdyson@seyboldreports.com
managing editor: Patricia S. Evans (x125)
senior editor: Mark Walter (x126)
executive editor: George Alexander (x142)
contributing editor: Steve Edwards
correspondents: Kurt Wolf
e-mail: kwolf@uitikon.ch
contributors: Laurel Brunner, James Felici, Hans Hartman, Matt McKenzie, Bill Rosenblatt, Ron Roszkiewicz, Andy Tribute, Bernd Zipper, David Zwang
production manager: Edmund H. Rozecki (x118)

Customer Service
phone: (800) 325–3830 or (610) 565–6864
fax: (610) 565-1858
e-mail: pubsvcs@seyboldreports.com
accountant: Shirley Gay (x 112)

San Francisco (Seminars) Office
795 Folsom St., 6th floor, San Francisco, CA 94107-1243
phone: (415) 905-2300
fax: (415) 905-2FAX
publisher: James Smith
editorial advisor: Craig E. Cline

The Seybold Consulting Group
The Seybold Consulting Group advises both users and vendors on an individual project basis. Please direct inquiries to Mark Walter at (610) 565-2480, x126, or mwalter@seyboldreports.com.

BCR takes over Seybold Publications subscription services

Note new address, phone number for customer service

As we announced in our last issue, Seybold Publications' fulfillment and customer-service operations will now be handled by our sister company, Business Communications Review (BCR).

Effective March 11, 2004, customer service, subscription inquiries and orders, back-issue requests, billing questions and subscription payments should be directed to:

Seybold Publications
C/O BCR
999 Oakmont Plaza Drive, Suite 100
Westmont, Il 60559
(630) 986-0385 Phone
(800) 325-3830 Toll-free
(630) 323-6324 Fax
pubsvcs@seyboldreports.com

For editorial inquiries and article reprints, please contact:

Cynthia Wood
Director of Content
Seybold Seminars & Publications
C/O MediaLive International
795 Folsom Street
San Francisco, CA 94107
415/905-2518 office
cynthia.wood@mlii.com

For Conference Sales and general inquiries, contact:

(888) 886-8895 Toll-free

People

Personnel changes around the industry

Phil Rugile has joined **Managing Editor Inc.** (MEI), a Philadelphia-area vendor of newspaper and magazine production software. In past lives, he was director of sales for NewsStand, Inc., COO of Reuters Personal Finance, CIO for The Baltimore *Sun*, director of technology for *Newsday*, and product marketing manager for Mediabridge Technologies (Cascade Systems).

Among his other credits, Rugile managed the "perfect page" pagination project with Atex Systems (and later SII) and Time Inc. He also managed the installation of the largest ad-automation project in South America at *Jornal O Globo* (a Cybergraphic site) in the early '90s. At Cascade Systems, Rugile introduced early natural-language search technologies and brought to market its first open-platform database for advertising workflow for the Mac.

Xaar announced three new appointments: **Bryan Palphreyman** and **Peter Addington** have been appointed as business development managers for the packaging sector, and **John Attard** has taken on the role of business development for organic semiconductors and other "functional fluid" applications.

Palphreyman will focus on new business in the label market. He previously held senior sales positions within LogoPack, Weber Marking Systems, Willett and Sessions of York Ltd.

Addington will chase new business in the carton market. He previously ran his own consultancy in cost-efficient packaging solutions, primarily for clients in the food and drink industries, and has held development roles at Schweppes International and Express Foods Group.

Attard will focus on the organic semiconductor and printed electronics markets (think RFID). He previously worked at Cambridge Display Technology, Rescitech Ltd., Philips Electronics in The Netherlands and SINTEF Unimed in Norway. **TSR**

The Seybold Report (ISSN: 1533-9211; USPS 020-953). Subscriptions are available in the U.S./Canada for $595 per year and Internationally for $640 per year (24 issues mailed semimonthly).

POSTMASTER: Send address changes for ***The Seybold Report*** to: Seybold Publications, 428 E. Baltimore Ave., Media, PA 19063
Periodicals Postage Paid at Media, PA and at additional mailing offices.

Next-Wave Publishing, Part 3: Revolutions in Content

BY MILLS DAVIS

The conclusion to our research into next-generation publishing looks ahead to the future, when semantics redefines the nature and value of content. We examine the potential impact semantics will have on data, computing, people, publishing, government and manufacturing.

In Part 1 of this series *(see Vol. 3, No. 15)*, we saw that next-wave technology is the result of several long cycles of innovation: computing, distributed intelligence, nanotechnology and biotechnology. We pointed out that the dominant driver of next-wave publishing technology is *distributed intelligence* rather than computerization. We outlined a framework for interpreting technology advances that distinguished five technology waves. The first three were past and present waves; the last two defined present and future waves, which we set out to explore in Part 2 and Part 3.

In Part 2 *(see Vol. 3, No. 20)*, the focus was on revolutions in process. We examined the impact that networked services and new media technologies will have on corporations and commercial publishing businesses. We explored repercussions in four areas—content strategies, media platforms, product platforms and infrastructure, and value chains—with industry examples illustrating how the broader trends play out in specific industry applications.

In Part 3, the conclusion to the series, we will turn our focus to revolutions in content. The unfolding story of semantic webs at this stage will lead to the emergence of a new kind of content—semantic-form declarative knowledge. In this stage, semantics (the meaning of something) gets encoded separately from content. Ultimately, semantics gets encoded separately from process.

In previous technology stages, we saw that digitization of content and the separation of data and process representations from software applications created new economic value in the form of new tools and product categories, new categories of output and new markets, resulting in major breakthroughs in process economics.

Semantics, it turns out, can directly encode ideas and patterns of thought—all theory, all knowledge; in fact, anything that has or can ever be thought by anyone. The arrival of semantic computing heralds the dawn of the knowledge age, in which truly new vistas open for publishing, information technology and manufacturing. The direction is toward systems that know, learn and can reason the way humans do.

In the following pages, we explore what these revolutions in content are and what they'll mean for us. We'll start by asking:

- What is semantics all about?
- What is the significance of semantics technology for:
  - Content?
  - Computing?
  - People?
  - Publishing?
  - Government?
  - Manufacturing?

Semantics and Smart Data

What is semantics all about? Semantics is defined as the meaning of something. In a computer, what exists is what can be represented. So, digital semantics is a kind of content. But more properly, semantics is a kind of knowledge, because the meaning of anything is something that we know about that thing and represent separately from it in the computer.

When we talk about the semantics of content, we're referring to something we know about the content. Similarly, if we talk about the semantics of processes, we're referring to something that we know about a process. The trend, as we know more about something, is to represent this knowledge digitally as data, rather than as program code. Of course, it didn't start this way *(see Figure 1)*.

Evolution of data. The evolution of digital semantics started with the "age of programs." At that stage, data was simply embedded in applications and used locally. Data was less important than the application code.

The next two stages were the ages of proprietary and open (based on HTML) data exchange. Here data became separately managed and widely shared. Data became just as important as the programming code.

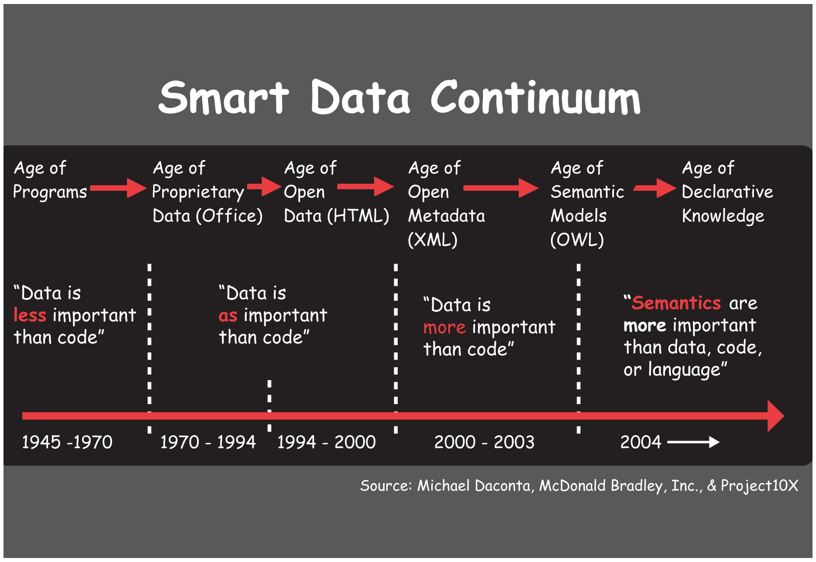


Figure 1: Smart Data Continuum. The illustration depicts the evolution of data. The basic direction is that data gets smarter, but the evolution happens in stages.

With adoption of XML for metadata, followed by RDF for the semantic web, we entered the age of open metadata exchange between systems and across networks. Metadata structures embed information and abstract ideas within recognized linguistic, graphical or symbolic forms (such as schemas). Metadata structures convey accepted patterns of meaning and well-understood relationships, which provide a needed component of "theory" that helps integrate data, content and processes.

Now, with the release of the Web Ontology Language (OWL), we are entering the age of semantic models. The key advance in this era will be that semantic models employ formal "named or coded relationships" to spell out and make explicit the "theories" left implicit within a metadata structure. Plus, behind the very plastic definition of a relationship, there now exists some underlying axiom or theory that explains and constrains how one related concept affects the other. This formality enables mechanization of reasoning, which is why semantic models improve information search, simplify integration, enable advisory services, automate custom communications and facilitate interoperability across systems and repositories. Semantically modeled metadata is now more important than the program code.

The progression does not stop with deployment of language-based ontologies. Next, a wholly new species of content emerges, and we will enter the age of declarative knowledge. At this stage, it becomes widely recognized that the only way to gain the benefits of precise meanings is to move away from natural language toward pure semantic codes and relationships. Semantics moves to center stage and becomes more important than either data or program code or, for that matter, natural language. This shift leads to new categories of products and services that open multi-billion dollar markets in publishing, IT and manufacturing industries.

Semantics for Content: Information Management, Libraries and Research

Information managers, librarians and researchers have a long history of experience with managing digital and physical collections. Library science goes back thousands of years and continues to evolve. While the environment and economics of content management are changing, the key concerns of librarians and information managers have remained pretty constant.

Key questions that these professionals are asking include:

- How can I automate the process of managing the repository?
- How do I classify, index and organize my collection?
- How can I improve search with digital technology—make it faster, more productive and less time-consuming for researchers?
- How can researchers find what is in the repository (and relevant to their questions) without having to look at it all, read it all or listen to it all?
- Can I have self-organizing repositories?
- How can I navigate across multiple repositories that have been indexed by different communities?

Information managers and librarians have been trying a broad range of knowledge representation techniques in order to resolve language ambiguities and improve the quality of content searches. To illustrate, we recently came across a "Dictionary of Search Terminology" (www.topquadrant.com) that discusses 70 categories of digital search technology. All of these are in use.

An example, from the NSF Digital Library Initiative, is the bio-science, medical and health-care repository being developed by the CANIS center at the University of Illinois. A sample question might be: For a patient with rheumatoid arthritis, what is a drug that reduces the pain but does not cause stomach bleeding? According to Dr. Bruce Schatz, finding the answer requires navigating across hundreds of millions (currently, 250 million) of concepts within millions of documents stored in multiple repositories, each of which has been indexed by a different community, but nevertheless can be accessed using the searcher's own special vocabulary.

From search to knowing

Semantics is strategic for information management, libraries and research for two important reasons. First, semantics is the key to better information search. Better search is worth millions of dollars. Second, semantics enables applications that know. This is what happens when the entire library is seen as one extended book, with all of its concepts, theories and factual matters present, unambiguously encoded and organized in a way that enables exploration of every question as well as every relevant path of reasoning. A researcher could not only navigate concept across repositories to locate sources, she or he would be able to reason and simulate directly across all knowledge contained in these repositories.

Semantics trumps linguistics

Semantic-stage applications always start with some sort of knowledge representation. An application reasons across this knowledge. The question is, what sort of knowledge and representation are we talking about?

Figure 2, "Types of Knowledge Representation," identifies 15 forms of knowledge representation arranged in five categories along an axis of increasing knowledge value. As we review the capabilities of different categories of knowledge representation, we'll also identify some limitations as well. A key theme is that language-based approaches suffer from ambiguities inherent in natural language. We contrast language-based knowledge representation with semantic-form declarative knowledge and conclude that *semantics trumps linguistics*.

Glossaries and data dictionaries. The simplest forms of knowledge representation are glossaries and data dictionaries. Lexicons collect terms used in information systems. Dictionaries and glossaries define terms in natural language, and can be used to define keyword searches.

Thesauri and taxonomies. A *thesaurus* identifies synonyms, related and contrasting words, and antonyms. *Taxonomies* categorize, abstract and classify terms. Metamodels of knowledge frequently employ taxonomies as a hierarchical arrangement (tree structure) for all abstract concepts. These trees may focus at any given level to show how concepts are exclusively different. Models at the base of taxonomies are very detailed and specific. Progressing up the tree, models become progressively more abstract. To abstract is to ignore certain distinctions in order to focus selectively on certain other commonalities. So, moving up the tree, each category subsumes or includes the broadest, most general characteristics of those below.

Typically, no two groups will make the same abstract choices as they move up or down. Different users have different social roles and needs. So, they produce multiple taxonomies (or mappings of concepts). For example, if there is one catalog (such as a collection in a library), then there are just as likely a dozen or more taxonomies depending on the usages and purposes of the taxonomies' users. Still, taxonomies allow concept-based search and guided navigation as an alternative to search.

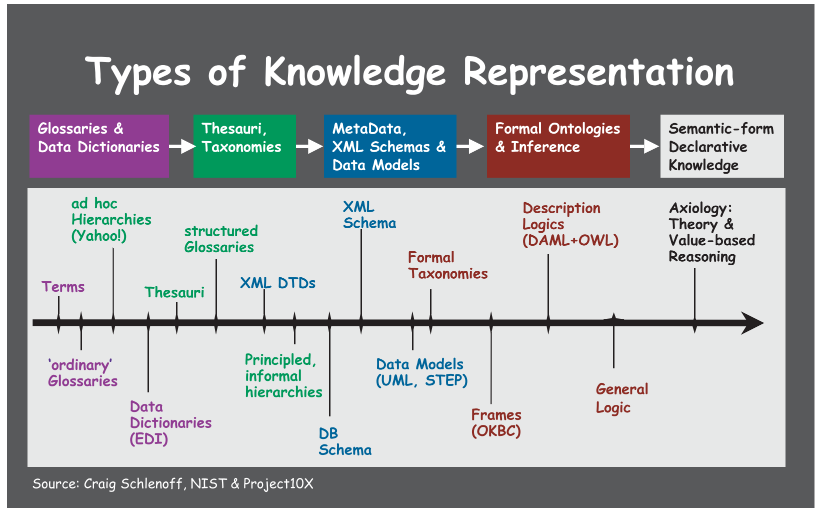


Figure 2: Types of Knowledge Representation. This diagram arranges 15 forms of knowledge representation in five categories along an axis of increasing knowledge value. Semantic-stage applications always start with some sort of knowledge representation. An application reasons across this knowledge.

Taxonomies have proven ROI in a number of areas. For example:

- Knowledge management—Within the organization, capturing and sharing institutional knowledge, such as best practices and corporate intelligence.
- Customer service—Reducing the time it takes customers to find answers in technical documentation, troubleshooting guides, how-to's, FAQs and best practices, organized product by product.
- E-commerce—Customers can't buy what they can't find, so businesses must support different modes of shopping (*e.g.*, classic "men's vs. women's" methods); organize materials according to marketing needs; offer technical details of different items; and merchandize related items.

Metadata, XML schemas and data models. Metadata, XML schemas and data models make up the next group. Knowledge is represented in the form of an XML DTD, XML schema, RDF, database schema or data model (*à la* UML or STEP).

Metadata here are still definitional data that provide information about or documentation of other data managed within an application or formal computer language environment.

A model describes how concepts and phenomena are similar and how they differ.

- An *object* model is a networked data-set tightly bound to the procedures that directly access or update it. Each describes its own modular world inside a system. Base classes behave like abstractions high in a taxonomy, expressing those properties that are shared by the many more-specialized systems derived from it.

- A *data* model describes a world outside the system. Several applications can share a database. In a relational DBMS, each table in the schema models an abstract concept dictating every distinction its collection of concept-instance records are prepared to save—some common, others different. Another table sets forth its one common idea and fixed list of shared or different property values. Typically, relationships are defined only by the records they connect, so each instance has no unique identity, property or justification in theory.

The problem we still have to solve is the ambiguity of language used in creating DTDs or schemas.

XML is one formal language used for defining collections of metadata to be exchanged between applications over the Internet. XML has given us syntax interoperability.

Given that the XML grammar is fixed, the problem we still have to solve is the ambiguity of language used in creating the DTD or schema. XML Namespaces allow groups to register vocabularies used in tags. This is a step toward interoperability, but names alone don't supply the knowledge needed to resolve many ambiguities. Does the label (or tag) infer one or many possible semantic concepts? No one knows or can know. Another party can plausibly take your words and then describe a very different semantic definition than the one you chose. That is the fundamental and inescapable burden of using language to define semantics.

Resource Description Framework (RDF) and RDF Schema provides a model based on XML syntax to represent and transport metadata. RDF Schema is an extension of RDF that provides mechanisms for describing groups of related resources and the relationships between them. RDF integrates a variety of applications: library catalogs and worldwide directories; syndication and aggregation of news, software and content; personal collections of music, photos and events, and so forth.

RDF gives us an Internet with two-way named hyperlinks and thus a way to expose metadata. However, in schema architecture, relationships are where you put them. What you call them has few constraints. In a "two-way, named hyperlink," then, the only relationship naming concern is that the name previews what you are likely to see first, given that you take that link. If so, intersite negotiations will favor a minimal set of abstract relationship types, because one-time users are not going to spend much time learning some new subtlety roaming through yet another site. They only need enough to take or not take the branch offered. Abstract schemas present minimal constraints—*i.e.,* offer little knowledge value.

Formal ontologies and inference. Formal ontologies and inference make up the next level of knowledge representation. These include formal taxonomies, topic maps, frames, description logics and general logic.

As the semantic model becomes richer, it more completely specifies not only the formal class-subclass relationships, but also relationships between concepts, and the descriptive logic and conditional assertions that are used to perform inference.

Topic maps are a method of using XML to represent networks of concepts to be superimposed on content resources such as documents of various types, providing a means to represent, navigate and query the topic-map network itself, rather than the full text of a document collection. Topic maps support both hierarchies of concepts (topics) and relationships between them (associations).

An *ontology* is an explicit formal specification of how to represent the objects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among them. In computer systems, what "exists" is that which can be represented. Ontologies organize concepts and their interrelationships in ways that facilitate machine reasoning and inference.

The Web Ontology Language (OWL) is a semantic markup language for Web resources and ontology construction that has just been developed and approved by W3C. It is based on RDF and extends the work of DAML and OIL languages for defining ontologies. The OWL language has three levels, with progressively more expressiveness and inferencing power. OWL will be used as a semantic markup for Web pages, data sharing and Web services. The goal is to have resources, repositories and processes semantically related and interoperable through ontologies.

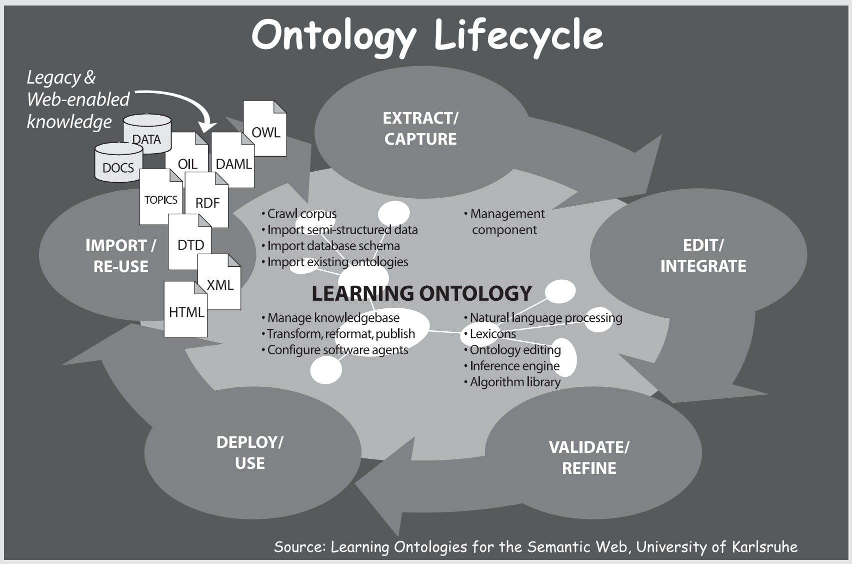
Ontology-based solutions are already having economic impact. For example, global publisher Bertelsmann owns Empolis, a semantic-technology supplier. Empolis has developed an ontology-based, self-help and customer-service application for a division of Siemens, which produces a wide array of control and monitoring devices that are sold all around the world. This application, distributed to Siemens service personnel (who are not necessarily Siemens employees) on-site for use with 65,000 users worldwide, is helping Siemens save $3 million a year.

Figure 3, "Ontology Life Cycle," presents an overview of a computer-aided process for constructing ontologies. Companies providing software for building ontologies include: CognIT, Empolis, Intelligent Views, Language and Computing, Lockheed Martin, Modulant, Network Inference, Ontoprise, Plugged In

Figure 3: Ontology Life Cycle

This diagram outlines the process for constructing ontologies. The stages include:

(1) *Import and reuse* legacy and Web-enabled knowledge sources. It is possible to "crawl" the corpus to identify and download sources in a variety of formats as well as to mine text to extract terminology from documents. Tools should allow import of legacy forms such as database schemas, product catalogs, yellow pages listings, semi-structured data sources and Web-enabled data. Also, it is possible to reuse ontologies, in whole or in part, that have already been developed.



(2) *Extract and capture.* This entails format conversion, analysis of source material, and creation of class structures and subject indices.

(3) *Edit and integrate.* Conceptual analysis is done to determine what an information object is "about" and to establish a knowledge-organization scheme that includes subject index, thesauri, classifications, etc. This can be approached manually, semiautomatically or automatically. Automated indexing ranges from simple natural-language processing (recognize parts of speech and words in a document) to sophisticated analyses that identify key names, words and phrases. Automated classification assigns documents to categories or classes.

(4) *Validate and refine* the ontology. Detect and correct errors and omissions; respond to changes in the environment. To validate the ontology, critique the formal semantics of what the class structure means. Description logics provide frameworks that support this sort of reasoning. Some tools employ automated description-logic engines to determine if an ontology has contradictions or gaps in the knowledge representation, or to detect when a particular concept can be classified differently, according to its description and that of other concepts. These critiques can be used to modify the ontology automatically, consolidating information contained within it.

When maintenance requires merging ontologies from diverse provenance, some tools provide human-centered capabilities for searching different ontologies for similar concepts (usually by name) and merging the concepts. Other tools perform more elaborate matching, based on common instances or patterns of related concepts.

(5) *Deploy and use* the ontology in any of several ways. An ontology provides a natural index of the concept instances described in it, and thus serves as a navigational aid for browsing contents. An ontology can drive similarity measures for case-based retrieval and reasoning. Web-ontology languages (such as DAML and OWL) have capabilities for expressing axioms and constraints on concepts that enable powerful reasoning engines to draw conclusions about instances in the ontology. Probably the fastest-growing area of development for ontology-based systems is semantic integration across various applications and content repositories. **TSR**

Software, Sandpiper Software, Semagix and Unicorn Solutions.

To date, most ontology work is driven by the assumption that knowledge search is the goal, and so the primary interest is in amassing and searching the set of all independent concepts (subject indicators) in the ontology.

A key goal of language-based knowledge representation is to eliminate the ambiguity of describing things with labels and natural language, leading to improved search and easier integration of content and processes. But, as we have seen, this goal is difficult to achieve when we use language to describe what we mean. Natural language use is inherently ambiguous. Many words have multiple meanings. There is no way to guarantee that two occurrences of the same word have the same meaning.

A standard ontology with just one noun-phrase selected for every category is an improvement. But, what else do you (must you) offer in evidence to guarantee the unambiguous meaning for this particular noun-phrase? Ask everyone to give their form, defining each concept. Do they have text definitions? Do they offer a series of sentences that connect different word concepts to one another? Can they use those sentence structures to define relationships? How do they define the relationships so that they are rare and effective discriminants between meanings? How many other concepts use these same relationships?

Ultimately, the only way to ensure precise meanings is to move away from natural language toward pure semantic codes and relationships; that is, use unique identifiers to identify concepts. (We may draw an analogy here with the UPC [universal product code] identifier that has no significance other than that it is unique.) Do not use labels or names of things. Rather, determine meaning by the sum of all the relationships the concept has.

Semantic-form declarative knowledge. Semantic-form declarative knowledge is the next form of knowledge representation. It is based on theories of knowledge and computation that define an emerging science of representation.[1] Here, conceptualism and semantics replace nominalism and language to solve seemingly unsolvable problems of great complexity and ambiguity. For example, model and instance codes are the same for all languages. A concept (model-instance) appears only once in any semantic web; the codes identifying it also locate it instantly without search. Meaning is defined only by the unique web of relationships tying every concept to those connecting to it. Properly implemented, semantic-form declarative knowledge webs approach absolute limits on physical size, speed, and efficiency.

A central goal of declarative knowledge is the development of a standard scientific representation for "all knowledge"—one that can be shared jointly by humans and machines—a common, evolving, most-fundamental representation seeking to encompass all science and all learning. A unique, scientifically precise and minimal representation for encoding every idea known and imaginable then provides the basis for a language-independent, standard, global ontology (initially hundreds of millions of "concept models").

Machine theories and computation will model human theories and reasoning. All things can be represented in semantic form. Semantics can directly encode ideas and patterns of thought—all theory, all knowledge; in fact, anything that has been or can ever be thought by anyone.

According to Dr. Richard Ballard (rlballard@earthlink.net), the scientific foundations for semantic-form declarative knowledge are made quantitative by the following axioms:

- Knowledge is defined by an unanswered question or that set of questions defining some field.

- Knowledge becomes measurable when one can define one or more "expected" or "acceptable" answer forms.

- The number of acceptable forms defines the initial answer uncertainty, or problem size, describable in bits.

- Knowledge is anything—influential theory or observable fact—that decreases answer uncertainty.

- Knowledge is measured by the amount that its possession (theory) or receipt (information) reduces uncertainty.

Knowledge is then defined as a learned and stored system of constraints that tell us why particular answers or classes of answers should be discarded. These constraints do not assure us initially that there is just one or any particular "right answer." They only narrow or eliminate possibilities. We deal with the remainder as we choose. Knowledge does not solve problems. It only helps us to predict, anticipate and manage their consequences.

Semantic-form declarative knowledge is composed of theory and information. Theory is a "metaphysical" constraint that asserts something thought absolute (this always/never happens) or conditional (this happens when/if/while/after some condition exists). Information is a "physical reality" constraint (observably—this situation is happening or has happened).

Theory and information constraints are absolutely and fundamentally different. To have influence, theory has to be learned and known before (often long before) an event. By contrast, information can only be known during and after an event. A "new" theory is likely to be 30–50 years old; most go back 3,000 to 40,000 years, and some are more than 2 million years old. Information looses value continually as situation awareness fades in minutes, days or years, becoming past history—recorded or not.

The expectation of semantic-form knowledge representation is that ontologies will be massive, rather than small and handcrafted. They will top 100–500 million concepts within five to ten years. Ontologies will be stable, slow evolving and, subject independent, and will encompass all knowledge codes.

Professions, publishers and governments will lead ontology codification. Knowledge content is a capital asset. Knowledge ownership and practice will define an organization's value.

Semantics for Computing

The information technology (IT) community has a 60-year history with computing that has evolved from vacuum tubes to networked services. Whether we approach the matter from the standpoint of consumer electronics companies, telecommunications providers, hardware manufacturers, software publishers or enter-

[1] The semantic-form representation of knowledge described here is based on "Physical Theory of Knowledge and Computing," by Richard Ballard, which expands the Mathematical Theory of Communication developed by Claude Shannon to provide limit-case bit-measures for the constraints of theory and information and a quantitative definition and measure of knowledge.

According to Ballard, semantic representational form should combine: (a) the concept-relationship formalism of Porphyry (ca. third century AD) with (b) structuralist semantic definitions (Ferdinand de Saussure) and (c) mediating structures (John F Sowa/Charles Sanders Peirce Synthesis) expanded to (d) higher logical order empirically by use of n-ary information-limit relationships.

Relevant reading includes: Claude E. Shannon and Warren Weaver, The Mathematical Theory of Communications (University of Illinois Press, 1949); John Sowa, "Ontologies: Lattice of Categories," in Sowa and Dietz, Knowledge Representation: Logical, Philosophical, and Computational Foundations (Brooks/Cole, 2000); and Richard Ballard and Robert Smith, "On the Evolution of a Commercial Ontology and Coding System" (Knowledge Foundations, 2001).

prise IT departments, semantics are now essential to make IT work.

The fundamental issues for IT today are managing complexity and uncertainty. Key challenges, such as security, pervasive services, stack complexity, autonomic systems and legacy conversion, demand solutions designed for the era of distributed intelligence, not for the desktop or the client-server world. On a global scale, market dominance will be worth trillions of dollars to the group of companies that can develop and deploy the winning architecture.

We're not making this up. IBM, Sun Microsystems, Microsoft, Oracle and other major vendors in the IT space have all focused on network services and semantic technologies as the way to fundamentally improve the economics of information technology, for both solution providers and their customers. The core argument has been that the cost to integrate, manage and maintain the expanding array of (component) systems and processes has become unsustainable. They believe the "grand challenge" for the industry to be developing systems and processes that are self-declaring, self-integrating, self-optimizing, self-protecting and self-healing. They give this vision different names, such as "autonomic computing" or "the net-effect," but they concur that you must have semantics to implement this vision. It's not an option.

We need semantics to make the World Wide Web work for computers as well as people. We need shared meanings to make Web services work between applications and organizations. It's the only effective way to reduce the time and cost required to integrate our processes. We need semantics to make grid computing work for large numbers of dynamic projects and distributed resources. Nothing scales without it. Likewise, we need semantics to cope with the swarms of devices, services and resources that will be orchestrated through the pervasive service grid. There's no other way to pull it off. And finally, we need semantics if we're ever going to develop systems that can self-configure, self-integrate, self-deploy, self-optimize and self-repair. The world of autonomic computing is unreachable without semantics.

Knowledge trumps programming

When it comes to semantics for computing, what kind of semantics do we need? The semantics of processes are represented separately from the application code and the content. But is this going to be enough? In some cases, the answer is yes. However, the argument we are making is that language-based approaches to semantics may max out at some point. They may not scale to handle the diversity of communities, the size and weight of the "standards" stack, the security requirements, etc. The IT community may find that it needs to examine assumptions and take a fresh look. In the following discussion, we point out two new directions that appear promising: context computing and declarative knowledge computing. Our take is this: *Knowledge trumps programming*.

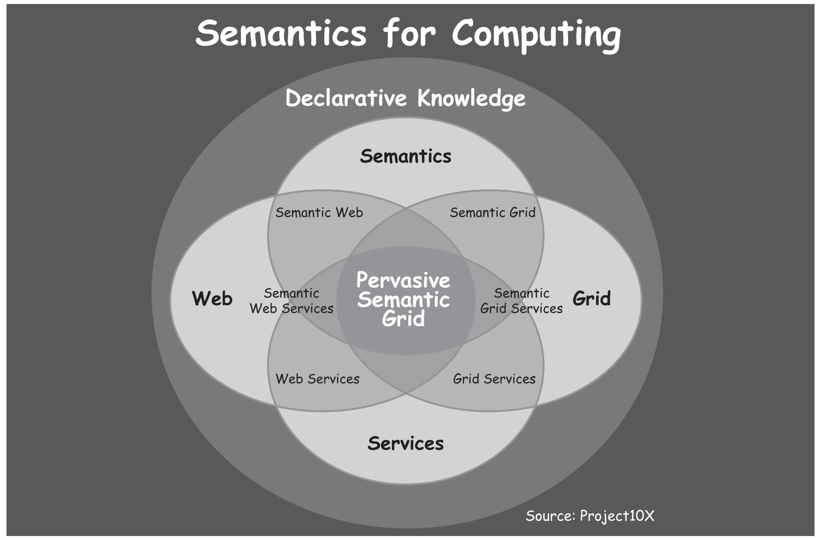


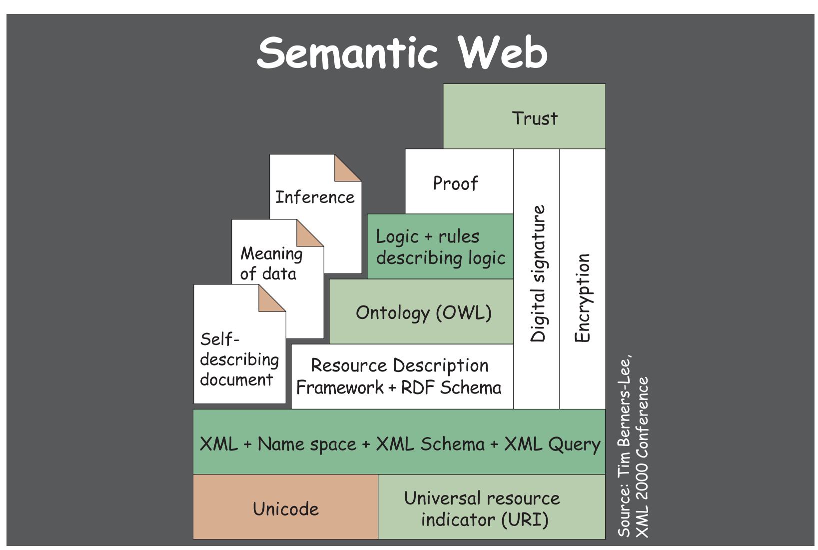
Figure 4: Semantics for Computing. This diagram depicts major lines of development for semantic technology in computing. Semantics for information processing is converging on the pervasive service grid as its new paradigm. This is the intersection of four major technology themes: semantics, the Web, grid computing and services. Surrounding and subsuming this is another emerging knowledge-technology paradigm, which we call declarative-knowledge computing.

Semantic tsunami

As shown in Figure 4, the major lines of semantic technology development for computing can be visualized as the intersection of four technology themes: the worldwide Web, grid computing, services and semantics, leading to the pervasive semantic grid as the emerging inner paradigm of information computing. Surrounding and subsuming this, we are suggesting, is another emerging technology paradigm: declarative-knowledge computing.

Semantic web. The *semantic web* (*see Figure 5*) envisions a transition from simple HTML linkages to machine-interpretable tagged relationships among

Figure 5: Semantic Web. This diagram depicts the arrangement of specifications (called a stack) that define the semantic web.



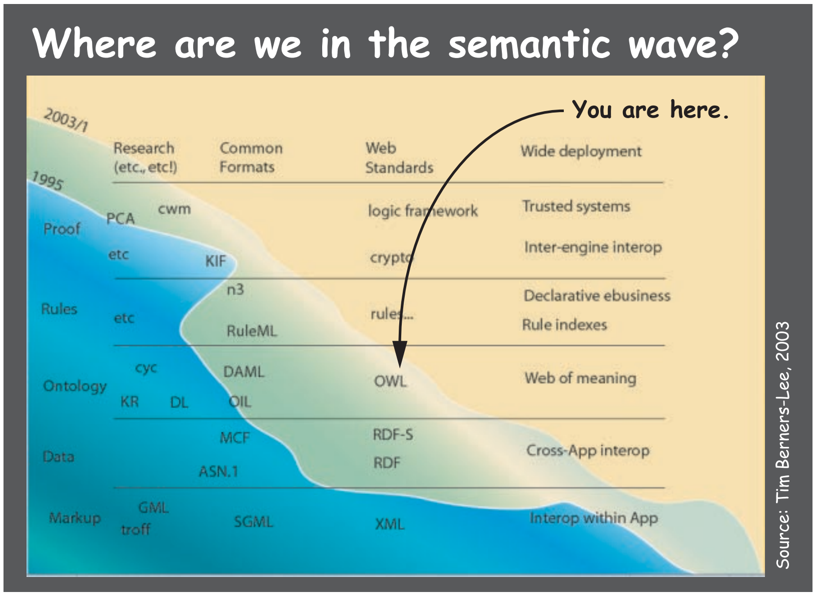


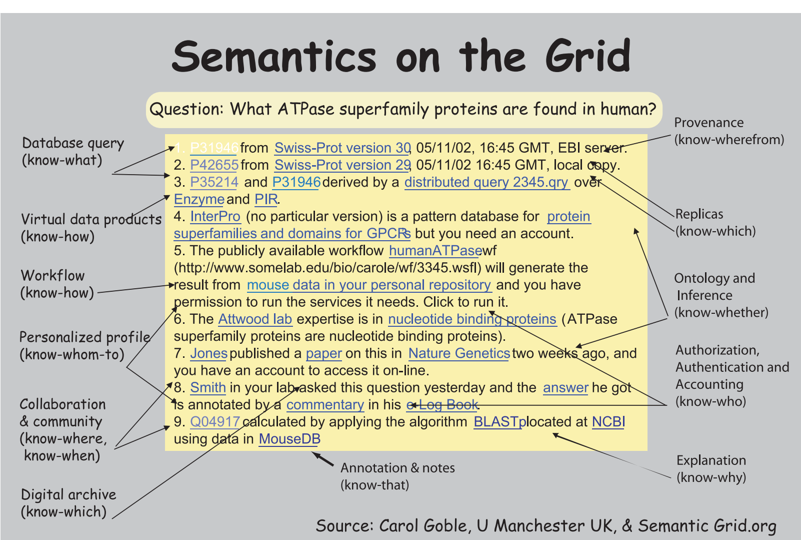
Figure 6: Where Are We in the Semantic Wave? This chart uses the metaphor of high watermark of an advancing wave along a beach front to describe the progression of semantic-web technologies from research to common formats, to web standards and to wide deployment.

resources. The goal: a web of globally linked, semantically related and distributed resources. A semantic web, as envisioned by Tim Berners-Lee, is a web of machine-processable data.

The semantic web is "crossing the chasm" now (*see Figure 6*). We'll see the tipping point within three years. Businesses will see it in portals. Consumers will see it in the integration of e-mail, calendar and contact lists with personal knowledge-bases (music, video, vacation, etc.).

Semantic-web services. Semantic-web services, employing shared semantics, enable disparate systems to discover a Web service and understand what it does, how it works and how to access it. Without semantics, who knows what the service provider meant?

Figure 7: Semantics on the Grid. This diagram provides an illustration of the types of application knowledge that are required to support e-science projects across the semantic grid.



Semantic-grid services. Semantic-grid services (see www.semanticgrid.org) envision the use of shared semantics to facilitate multi-participant dynamic specification, allocation and persistent management of distributed computational resources. The original motivation for grid computing was the orchestration of distributed computing resources.

The motivation for semantics in the grid *(see Figure 7)* comes from recognition of the need to support collaborative projects and virtual organizations from diverse provenance, as well as the need to access knowledge developed and indexed by diverse groups, using different languages and methods in different regions of the world.

Pervasive computing. Pervasive, or ubiquitous, computing envisions an environment where devices that compute and communicate are everywhere, for example, the environment, clothing, everyday artifacts, sensor arrays, etc. Pervasive computing and semantic-grid services face similar issues. They are large-scale, massively peer-to-peer distributed systems. They must provide for service description, discovery and composition. They must solve issues of availability and mobility.

Pervasive service grid. The center of the diagram in Figure 4 shows that these lines of Internet development are converging on a new information-computing paradigm called the pervasive service grid. It is the manifestation of the semantic grid in the physical world through pervasive computing. This is also called ambient intelligence.

Context computing

Context computing is what will enable the pervasive service grid to function. Context might be thought of as the next level of semantic-web stack. It treats everything as nodes, spaces and relationships. These are multidimensional. The sum of these dimensions is called a context. In the semantic web, a model of a context is an ontology that integrates web, grid and mobility services, and unifies both content and procedural language forms across all layers of abstraction, right down to the bits, in a way that scales across swarms of devices and resources in a network, enabling them to combine and interact securely and efficiently.

According to Sandy Klausner, (klausner@coretalk.net), a leading proponent of context computing, the central aim is to integrate all data and processing languages and usage into a common ontology-based representation that spans all levels of the stack, all memory and all nodes of the network, and processes across all (peer) nodes as unified (symbols to bits) binary. In short, context computing attempts a grand synthesis that integrates and unifies all of the interior layers of the IT process paradigm. The advantages of context computing include scalability (which is needed to handle pervasive computing), immunity to viruses and security attacks once within context, built-in digital rights management providing a business model even for open-source components, and extreme performance across diverse network nodes.

One area where context computing is already receiving attention is national defense. According to a recent white paper by Michael Daconta (McDonald Bradley, Reston, VA) entitled "Semantic Web Foundations of Net-Centric Warfare," a key aim of context computing is the ability to model the context of content, services and transaction message spaces from multiple user perspectives. This enables synchronization of these contexts with each other, which is a key enabler for the new kinds of high-mobility, edge-powered, rapid-deployment combat operations where force superiority depends on better knowledge, high coordination and faster sense-response.

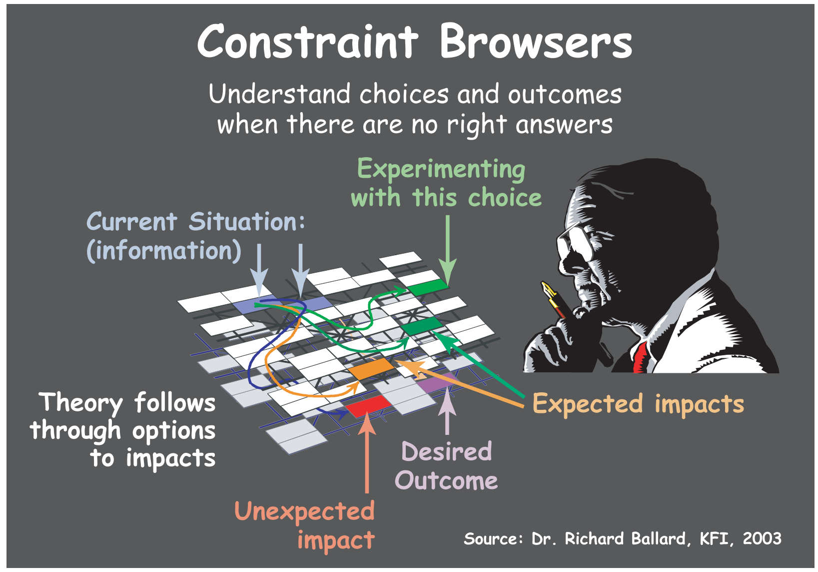


Figure 8: Constraint Browsers. This diagram illustrates a kind of tool that a decision-maker might use to reason with semantic-form declarative knowledge.

Knowledge computing

Referring back to Figure 4, the surrounding circle shows that semantic technologies are simultaneously evolving in another, profoundly different direction, which we call declarative-knowledge computing. The goal of knowledge computing is systems that know, learn and reason the way humans do—not just with logical consistency and determinism, but also with all value systems and patterns of thinking.

In declarative knowledge, all theory and information is present and expressed as pure, extra-linguistic semantic crystals. There are no labels. Meanings are structural, the sum of relationships with other concept instances. Declarative forms require no "execution." All paths, from any and all inputs to any and all outputs, are always already present. Reasoning, then, is not restricted to executing hierarchies of deterministic procedures. There is no search. Answers to questions are paths that are defined and constrained by theory and information (facts, situation awareness). It's ironic to think that the most practical solution to problems of process complexity and language ambiguity may be to abandon both in favor of declarative knowledge.

One way to think about declarative computing is to imagine a computer program (a coded algorithm), plus all of its possible inputs and all of its possible outputs, arranged so that all of its possible execution paths are already taken and recorded in the form of a knowledge web.

A constraint browser is a type of tool that a decision-maker might use to reason across a semantic-form declarative knowledge base *(see Figure 8)*. For example, there might arise a question about Sarbanes-Oxley compliance, and if a decision-maker had established that for one division the company, if one knew the capabilities of people—what they were doing, the flow of work, the functions of systems used, the separation of controls and the monitoring and testing being applied—then an answer about the extent of compliance, including material weakness of controls, could be given. Further, if the decision-maker had established that when these informational constraints were combined with legal and regulatory requirements, standards, industry benchmarks and current case law, then one could determine the nature and extent of risks, possible consequences and trade-offs that the company was facing.

In this example, input about the current situation is simply an informational constraint. If you know it, then theory follows through the options to the various impacts, and you know the output(s). If you want to know "how" you got to that output, just follow the path in the forward direction. Conversely, if you want to know "why" you ended up with that result, then you can follow the path in the reverse direction.

Also, you may have a bundle of different "trade-off" paths that end up at a particular expected or unexpected result. Then you can explore "what ifs" by experimenting with different choices and examining the effect of changing different information and constraints.

Using declarative semantic-decision tools, users walk down all branches and examine the consequences of making a decision one way and not another. Only after all of the potential decision impacts are known do users actually pick a particular branch. That is the classic "what if" methodology. As a rule of thumb, the criteria or theory that they say influenced them most in making the trade-off will end up being heuristic (in the user's head, but not formally in the knowledge asset). They have figured out some plausible advantage in choosing one thing.

Declarative-knowledge tools can have self-awareness of limitations of knowledge (*e.g*,. by identifying gaps in otherwise-valid reasoning patterns), and can take steps to learn. For example, having established an acceptable path of reasoning about Sarbanes-Oxley compliance, this "template" might be applied to another division. Doing so might reveal gaps in information. The decision maker might learn that personnel profiles were incomplete, or that the separation of controls was unknown. The tool can reason to determine the seriousness and consequences of incomplete information, and indicate what determination could be made if missing knowledge or facts were provided. Also, it can take directed action to seek and fill in missing information.

So we can see that declarative-knowledge computing is really quite different from simply executing a program, because an algorithm reasons in only one direction to a logically consistent output. Language-based, programmed solutions tend to be looking for some way to divide and conquer a problem. They assume that they need to model the sponsor's tasks and organization, and they assume that they need to model the sponsor's desired answer form. They assume that logic and algorithms are the only possible rational knowledge. And, they assume that the same inputs must always return the same answers and never learn or change.

Humans don't reason this way. They face complex decisions leading to trade-offs with varied consequences. They reason, not only with logical consistency, but using other value systems. Judging guilt or innocence, for example, is a much deeper decision to make than determining logical truth or falsity. Semantic-form knowledge bases make trade-offs in lives, dollars, morality, job loss, education, base closings, laws broken—that's complexity. It requires reasoning in many directions, using multiple value systems, the way humans do.

Perhaps, back when the price of memory was high and Von Neumann computers were all we had, then algorithms seemed like the smart way to go because that approach required less investment in memory. The price of procedural computing, however, is that the program only executes in one direction; it doesn't remember and it never learns.

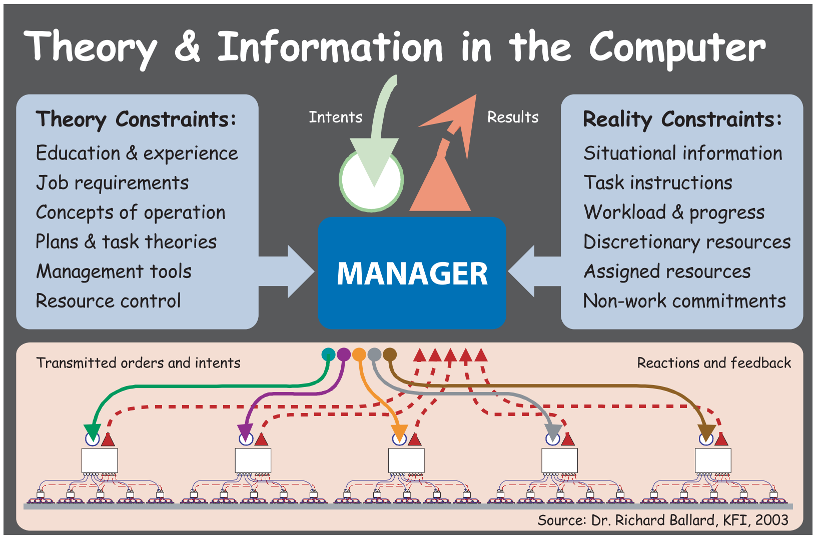


Figure 9: Theory and Information in a Computer. This diagram depicts a scenario in which management systems provide both theory and information for conducting the work of the enterprise. More than 80 percent of management decision-making relies on theory, with the rest being facts and situation awareness.

Semantics for People

Semantics is key to changing the economics of labor, including the cost of education, personnel acquisition, productivity and labor rates. There are several current strategies for improving workforce productivity and managing labor costs:

- *Mechanization* seeks to substitute capital investment in machines for labor.
- *Outsourcing* seeks to exploit differentials in labor rates and other costs among different geographies and business entities.
- *Labor transitions* (*e.g.*, from professional to paraprofessionals in law and medicine) seek to substitute less-skilled workers for higher-cost workers in certain tasks.
- *Service automation* seeks to displace labor or maximize productivity.
- *Self-service* seeks to offload labor costs to the customer or supplier.
- *Information technology* seeks to improve labor productivity through digitization, automation, integration and optimization of information-based tasks and activities.
- *Education, training and distance learning* seek to transfer knowledge efficiently from sources to empower new generations (of labor).

Knowledge technologies (*i.e.*, semantics embedded in tools, processes and infrastructure) will accelerate and dramatically intensify the impact of all of these approaches for dealing with labor costs. Knowledge technologies will promote an unprecedented degree of career mobility and enhanced productivity at all levels of the job market. Given a professionally adept machine backup, early-career specialty training will be substantially shorter, but adaptive mid-career training will be constant. This is good news. However, labor transitions will impact professions, management and technical ranks—categories that previously have been less impacted than agriculture, manufacturing and service industries. Sustainable careers for highly educated, specialized professions will shift toward new knowledge discovery and marketable knowledge-asset creation.

Knowledge tools

In any job, declarative knowledge becomes the basis for new categories of research, analysis, planning, design, diagnosis and decision-management tools. Today's information systems focus on bringing information to the job, that is, situation awareness. Next-wave knowledge systems will deliver all of the theory and information needed to perform the job or task *(see Figure 9)*.

The realm of theory has never been fully present in information technology. For example:

- Rules engines in software applications and databases are trivial.

- Semantics (meanings of things) has been hard-coded into IT technology, so it can never learn.

- AI, to date, has remained machine-like in reasoning, not delivering a foundation for engineering practical knowledge solutions in adaptive human settings.

Rather, the knowledge required to do a job is something an employee has to bring with him or her (via previous education and experience) or learn (on the job or by formal training). This education is expensive to acquire. Also, when people leave, the knowledge is rapidly lost to the organization. Similarly, to automate work on the farm, in the factory or in the office, the knowledge required to accomplish the task is laboriously hard-coded into mechanical parts, circuitry and software algorithms. Improvements in capability require repeated investments in next-generation solutions.

Visualize the role of declarative knowledge in discovery, simulation, diagnosis and decision-making. Having theory-in-the-computer enables a legal researcher to both retrieve case law that is relevant to the brief *and* see its reasoning applied to the case at hand. Also, imagine a paraprofessional with knowledge-based tools that enable less educated personnel to perform diagnoses and other key functions of professionals, in legally defensible ways.

In engineering, theory-and-information-in-a-computer leads to a new kind of design-build process in manufacturing, architecture and engineering. Here, semantic-form declarative-knowledge tools accelerate the design cycle, especially for complex engineered products such as cars and airplanes.

The question might be: What are all designs that have specified properties of performance, noise and safety characteristics? The declarative-knowledge web includes all science, engineering, manufacturing, standards and regulations, as well as history. Design tools embody knowledge and theory for all possible designs in a solution space, enabling "what if" simulations that reason from desired results and attributes backwards.

From a semantic-form model, design flows directly to a manufacturing process that proceeds from virtual to actual. Designs can be automatically rendered as drawings, described as specifications, presented as briefings, planned and scheduled as a work breakdown structure and bill of materials, outsourced and subcontracted through a multi-tier supplier network, submitted for regulatory approval, and so on.

Simulation is another low-hanging fruit of declarative knowledge. Theory-in-a-computer immediately calls for some way to test it. Simulation is the preeminent way to test. Semantics and the abundance of theory concerning every physical, rational and social process will make knowledge-based simulation a central subject of every argument on plans, policies, strategies, new law, economics, social values, etc. Proponents and skeptics alike will test macro and micro models against past history and their suitability to predict the future will become the basis for debate and vivid interactive demonstrations.

Demonstrations of knowledge-based simulation will extend to historical, professional and archetypical personalities. We'll probably see this first as entertainment, then as models of great teachers in action, of enlightened prophets and practitioners—or, alas, of individuals trapped in narrow and ignorant worldviews. Ultimately, the most attractive of these may become images of the kind of person or expert or teacher or parent others might become if they could apprentice themselves to the training and use of particular knowledge assets.

To summarize, knowledge tools have broad applications. There are as many domains for knowledge-enabled labor tools and systems as there are:

- Industry sectors and segments—government, manufacturing, services, energy, publishing, etc.

- Job categories—by role and responsibilities within an organization.

- Functions—such as decision-making, research, design, planning, analysis, marketing, sales, support.

- Disciplines—including management, projects, engineering, accounting, finance, software development, medicine, law, scholarship, etc.

- Hobbies and interests—gardening, home improvement, entertainment, games.

Semantics for Publishers

In the knowledge age, the concept of publishing needs redefinition. Part 2 of this series discussed the content cycle through which publishers create, acquire, manage, package, deliver and make content public for use.

FIGURE 14: CHARACTERISTICS OF NEXT-WAVE INFORMATION AND KNOWLEDGE TECHNOLOGIES

Summary of some of the general characteristics of next-wave information and knowledge technologies.

	Information	Knowledge
New Capabilities	Solve or better manage existing IT problems, such as search and integration of content and processes.	New product categories, different capabilities tap new value sources; know, learn, communicate; do things we couldn't do or afford to do before.
Operating System	Context computing creates pervasive service grid.	Knowledge operating system maintains and reasons across massive semantic webs of theory and information.
Data management	Ontology-based smart data with massive associations exceeds capabilities of RDBMS.	Semantic-form knowledge stacks avoid exponential complexity growth associated with relational DBMS technology.
Software	Thinner applications as more "smarts" goes into the data.	Thin applications for declarative knowledge computing; semantic form tools; knowledge engines embedded in intelligent systems.
Hardware	New virtual machines optimized for context processing in a massively peer-to-peer mobile world.	Long-lived, non-Von Neumann computing architectures, optimized for *n*-ary memory traversal and mobility.
Systems	Autonomic systems that can self-declare, self-configure, self-integrate with other systems, self-optimize, self-protect and self heal.	Autonomous systems that know, reason like people and can learn.
Ecology	Improved IT life-cycle economics benefits existing participants: reduced effort, cost and time to develop, deploy, operate, service and maintain, and upgrade solutions.	Software value chain with new players including content providers, government, and third- and fourth-party developers. Self-evolving products transform life-cycle economics of IT, publishing, manufacturing and other industries.

To this we must now add the life cycle for declarative knowledge, which is about knowing, learning and communicating. Together, these open huge new market opportunities for the publishing industry.

Business-information services and professional publishers have a long history of working with information and reference sources in digital form. They've needed to solve problems of corpus building, maintenance, classification and indexing (including multiple indices), print and digital delivery, currency and relevance to customer need, ease of use and integration with their customers' processes and usage context.

Having experience with different approaches, publishers and business-information services recognize that they need semantics for their content and processes. They've built taxonomies to facilitate access. They recognize the need for markup and metadata to enable better machine processing and searching, as well as content multi-use and multi-channel delivery. Also, some have recognized that ontologies can extend the effectiveness of user interfaces. As they gain experience with Web services for internal-process integration, as well as for customer-facing services, they recognize that process semantics play an important role.

Historically, publishers with knowledge-rich content assets enjoyed the greatest success in domains where content was well structured, or organized so as to be reasonably well understood by the using community. Here, the limitations of language-based approaches to semantics were not overly burdensome, because the target audience could supply the knowledge needed to use the service effectively. But this limited the opportunity to add value, since the customer was supplying the smarts. Semantic-form declarative knowledge provides a way for publishers to escape that limit and move from an information service to a knowledge platform for their products. This leads to product families that change the rules of the game by taking significant time and cost out of their customer's or client's process through knowledge tools.

For publishers, declarative knowledge creates a new class of business opportunity that applies across many categories of business-information service as well as professional and scholarly publishing. For that matter, it applies to many categories of consumer, hobby and entertainment publishing just as well.

Moving from searching to knowing adds a new level of value. The direction is from computer-aided access to information, to semantics-enabled navigation of concepts, to declarative knowledge enabled reasoning across knowledge assets. At each stage, semantics increases asset value.

Using semantics in product strategy

The basic strategy is to convert legacy assets from a publishing division or business-information service to create a new type of product that combines all relevant theory and fact into a reasoning tool. For example, when asked a question, the new tool answers the question and shows tradeoffs and reasoning. It doesn't just return a list of sources. It isn't a book, but it might be a DVD with a book about it. Complex theory and details (probably organized as tables) and other reference knowledge, which would just not be practical to publish in a 2D format, become practical and valuable as a semantic-form encoded knowledge tool.

Putting both theory and information into the computer creates a wholly new experience for the customer. It's like the difference between reading a book about playing a game of chess and having an expert advisor to help you strategize and play the game better than ever before. That's why the future of reference services will be embedded software rather than static reading material.

Combining relevant theory with information changes the rules of the content marketplace. It opens a new competitive vector. The new category of asset is a tool rather than just a publication. The tool commands critical reference knowledge. The new product becomes an active (not passive) asset that is self-evolving and self-learning, and that increases in value as it is used.

For publishers with vision, next-wave competition will be based not just on the completeness and timeliness of information, the quality of its organization and the ease of access to sources, but also on the performance of the knowledge asset at conducting specific tasks and functions performed by those using the service. Success will depend on the quality of the results that customers achieve from applying the knowledge-based tool to activities such as research, analysis, planning, simulation and testing and evaluating alternatives, consequences and trade-offs.

Declarative knowledge becomes the cornerstone of knowledge-age publishing strategies and the basis for sustainable brand dominance. This will be true in areas of popular culture, fashion and entertainment; in news and information segments; and in professional, scholarly and business segments. Knowledge dominance will be the key determinant of who "owns" which media space.

The content providers that are in the best position to win are likely to be publishers that already have a strong base of (knowledge-rich) content as well as strong, established relationships with the specific (consumer or other) micro market. But this cannot be taken for granted, since competing interests could tap a broad range of sources to develop competing knowledge assets. The prize for the publisher is to "own" the meeting place for those that want to learn and do (whether they be hobbyists or business professionals), with those that want to market to that interest (advertisers, etc.), and those that have something to say or communicate or teach to this audience.

Owning the forum, and with it strong, highly valued life-cycle relationships with end customers, is what enables the publisher to collect (multi-channel) subscription revenues, advertiser-based revenues (including co-marketing and co-selling revenues, when the publisher's business model includes e-commerce), and ancillary service revenues.

Moving first to establish a knowledge-age market is important because it creates a barrier to entry to other publishers. The first into a new market spends the least and gains the greatest share. The second to arrive must spend twice as much to gain half the share of market. And so on.

To summarize, the opportunity for commercial publishers is to amass and organize a dominating reference source. The marshaling and authentication of any body of theory is a capital expense conveying ownership and creating substantial barriers to competition. Dominating the theory positions in economically significant markets creates the frameworks for structuring all tasks and information use. In declarative semantic-form, theories will remain relevant for tens to hundreds (potentially thousands) of years, independent of facts and language changes that make them appear different. New market opportunities build firmly on existing customer relationships, content assets and subject-matter expertise.

Publishing models facilitate knowledge capture

A key gating factor for knowledge-age publishing markets is the cost of converting legacy content to declarative semantic-form knowledge. The economics of knowledge acquisition have matured over the past 30 years. In that time, important lessons have been learned. They include:

- Hand-building ontologies and rational architectures is too expensive.
- Experts validate system particulars, but these are self-assembled on a productive industrial scale requiring investment and ongoing production and maintenance support.
- The sub-language hypothesis for ontology construction (*i.e.*, that a semantically well-formed context of language exists and is widely understood in some domain) doesn't work out in practice. Rapid change intermixes specialized terminology in short order.
- Language-based ontologies have inherent ambiguities.
- Capturing the declarative knowledge from legacy reference-content sources is economically feasible.
- The cost of converting from semantics to language forms is tractable and has bounded economics.
- The cost of attempting to capture semantics using language forms is unbounded.
- Developing language-based content and then converting it to semantic form adds about 5–7 percent to the cost of the original content product.

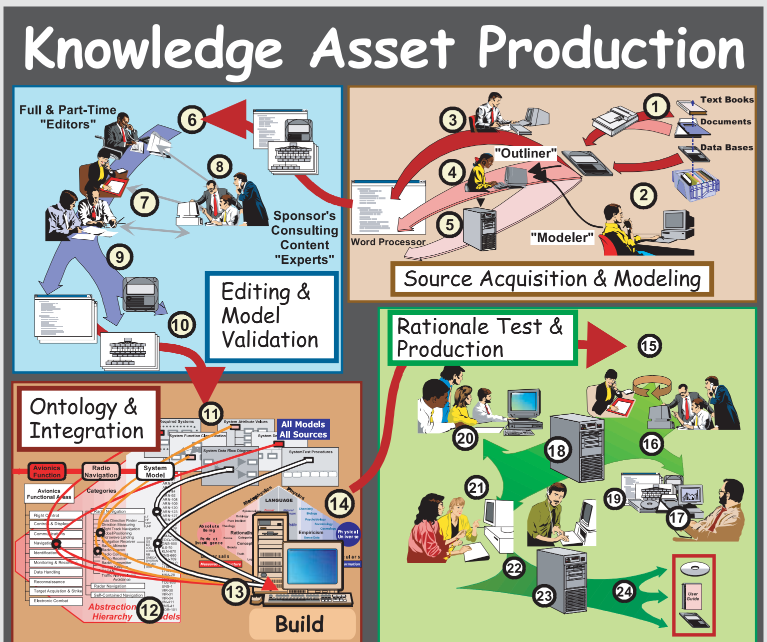
Figure 10: Producing Knowledge Assets

This diagram depicts 24 steps in the process of building knowledge assets from legacy sources. This process is divided into four phases:

- Source acquisition and modeling.
- Editing and model validation.
- Ontology and integration.
- Rationale test and production.

The key roles in this process include the following:

- Project manager, product designer.
- Outliners, transformers, production personnel.
- Knowledge editors.
- Subject specialists.



Source: Dr. Richard Ballard, KFI, 2003

The process of knowledge-asset production introduces several categories of new tools, for example:

- Learning agents extract meaning from any and all forms of content, and encode it in semantic form.
- Acquisition tools convert content to semantics.
- Creation tools express and manipulate ideas in semantic form.
- Building and editing tools amass and integrate knowledge.
- Knowledge-worker tools enable individuals and teams to work with ideas in semantic form and to integrate knowledge assets together.
- Semantic browsers (or "knowing" tools) view knowledge and related content, following reasoning paths or answering questions.
- Knowledge engines form part of knowledge-based computing and application processes.
- Communicating and teaching tools translate from semantic form to language, picture, simulation and other content forms.

- Developing content directly in semantic form and then expressing it in multiple language forms across multiple media can save 50–60 percent of content life-cycle costs.[2]

For the next few years, the process of creating semantic-form knowledge assets depends on human modelers and editors to carry knowledge across the semantic gap from linguistic ambiguity into precise and validated semantics. *(See Figure 10 for an overview of this process.)* The good news is that publishers already have the business and staff models to transform existing content assets into products. Further, commercially marketed knowledge assets from different sources can be integrated into small to massive layered stacks *(see Figure 11)*. However, at the point where science and commercial work products are created originally and delivered preferentially in the more valuable and sharable semantic codes, this need for some human agent in the loop disappears.

What about knowledge computing across the semantic web? For humans, knowledge computing probably does not take place across the Internet directly, but rather through massive semantic webs of knowledge that are local to them, but regularly updated via networks. For humans, the speed of thought doesn't wait on speed-of-light delays or interminable transmission jumps, but it does depend on navigating an *n*-ary reasoning path. Machines, by contrast, operate over longer times without organic short-term loss of attention.

Lost in translation

In the 1980s, the Interagency Language Roundtable (ILR), made up of 18 agencies of the U.S. Federal Government, developed a framework for testing and evaluating human proficiency with language. The framework for language proficiency addresses speaking, listening, reading and writing at five levels of complexity, *i.e.*, none, elementary, limited working proficiency, general professional proficiency and advanced professional proficiency. The framework was then expanded to measure performance at translation. The ILR framework treats translation as a composite of skills that includes reading in the source language, writing in the destination language and mak-

[2] In more than 50 knowledge-engineering projects for publishers and government agencies conducted by Knowledge Foundations, Inc. (KFI), the researchers found that:
(1) The cost for capturing a source completely into semantic form averaged \$5,000–\$7,000 in direct labor, compared with an initial investment in the \$80,000–\$100,000 range to research, edit and produce the source document in digital form.
(2) Reuse of knowledge assets in a field accounted for 70–80 percent of the knowledge base of subsequent projects, once the first "definitional" project had been completed.
(3) Savings of 50–60 percent over the content-media life cycle are attainable for projects involving natural-language generation of dialogs, documents, graphics and instructional materials from semantic-form knowledge.

ing congruity judgments. The framework distinguishes among professional, transitional and pre-professional levels of performance and introduces four terms associated with bringing texts across languages: translation, rendition, code-matching and glossing.

What is significant in the ILR framework is the importance it assigns to semantics in attaining progressive levels of proficiency with a language, as well as levels of performance with translation. To some extent, this echoes common sense. If you want something translated well, then choose someone who knows the subject matter, not just the language. And ideally, find someone who knows the culture of the target audience as well and how to communicate with it.

The realm of semantics contains hundreds of millions of unique ideas and concept instances, while language consists of only a few hundred thousand words, at most. As shown in Figure 12, the path from language to language is always ambiguous. The path from linguistics to semantics is always approximate, partial, unbounded and economically open-ended. You cannot get there from language using language as the encoding for meanings. On the other hand, the path from semantics to linguistics (from semantic-form meaning to text, pictures or sound) is straightforward and achievable.

Language to semantics is an "inverse problem." If you start from semantics, where every idea has a unique coded identifier in an ontology of millions of ideas, you can look backward from that ontology with far more certainty. For example, you might discover that there were 43 common uses for that sound and 2,489 rare uses by groups numbering 100 or fewer. Then, if all possible associations are known, moving in the forward direction (from semantics to linguistics) is achievable.

In the coming era, it is not unlikely that the machines accompanying us (such as cell phones and PDAs) may play a major role in mediating our conversations, asking and answering questions, noting significant agreements and differences in our planned objectives, and raising awareness of the outstanding issues we may seek to resolve.

Ideas become products, and prospects become business relationships, through a process that entails a cycle of communications. Semantic-to-linguistic translation promises extraordinary improvements in disambiguation. In the not too distant future, we should accept that the machine's semantic web-based language skills may be better than our own. Who then should write our technical literature? If we do not want the costs of continual factual edits, then probably the machine should.

Semantics-based natural-language generation will play a major role in all stages of the life cycle of customer relationships, product design and manufacturing, supply-chain relationships, legal and regulatory matters, and health care.

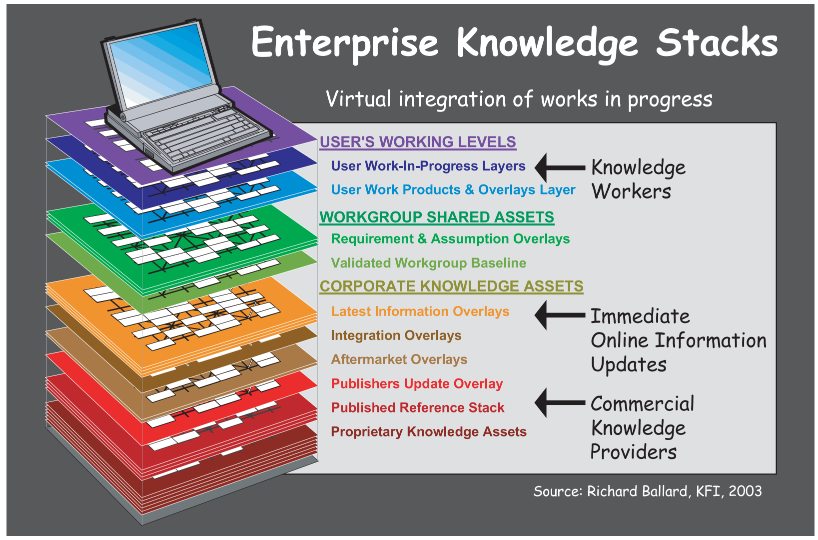


Figure 11: Enterprise Knowledge Stacks. This drawing shows a work environment wherein a collection of layers has been gathered, organized, integrated and worked upon by any number of contributing users. A session overlay creates this stack by dynamically linking together the working layers. Linkages are virtual. In this example, user work-in-progress layers may be updated, while proprietary product layers are treated as "read only" within such stacks. Combined with encryption and digital rights management, knowledge stacks provide an enabling infrastructure for knowledge commerce.

Semantics for Government

A "Semantic Technologies for E-Government" conference was held at the White House Conference Center in September 2003. Among the many agencies represented by more than 130 attendees were the Army, Census Bureau, CIA, DIA, DOE, EPA, GSA, IRS, Navy, NARA, NASA, NSA, NSF, SSA, USDA and the U.S. Patent Office. A number of attendees were from nonprofit organizations such as Aerospace.org and Mitre. Major government contractors were also well represented, including BBN, CSC, Lockheed Martin and SAIC.

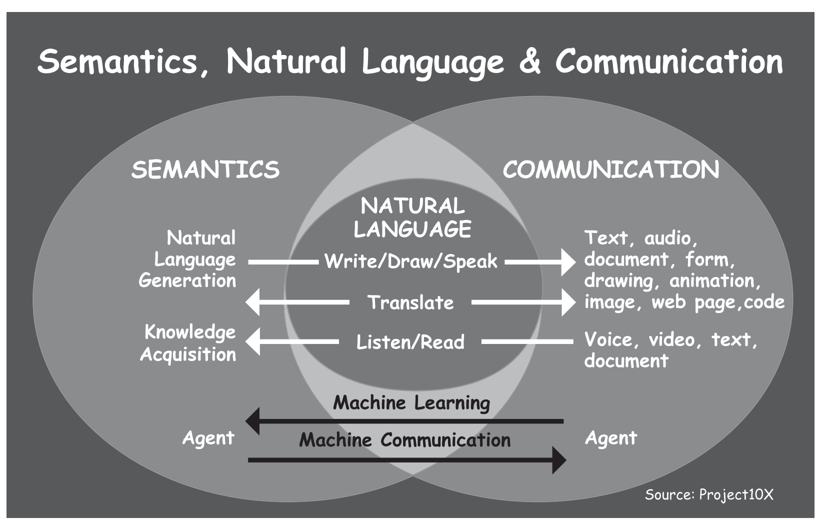


Figure 12: Semantics, Natural Language and Communication. This diagram depicts the relative sizes of the intersecting realms of semantics, natural language and communications. It summarizes pathways of knowledge acquisition, natural-language generation, translation, machine learning and machine-to-machine communications.

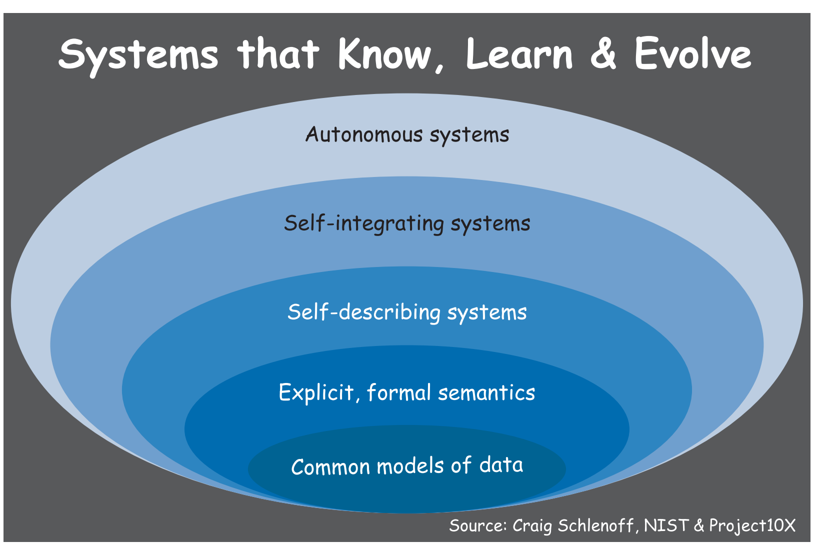
Q&A (question and answer) systems were a common concern of many of these agencies. That is, how could they put together systems that a policy researcher, program manager, intelligence analyst, executive, congressional staffer or constituent could use to integrate knowledge with public and classified information resources to rapidly explore and answer complex questions?

Semantics are the only practical way to build a Q&A system. The capabilities can be surprising. Here, for example, we envision capabilities that could become available to an enterprise or federal agency transitioning from current information-systems technology to a Q&A system based on semantic-form knowledge technology.

Stage 1. Initial capital investment builds the ontology, converts and bulk loads the key reference knowledge (hundreds or thousands of documents), validates and links the core knowledge assets, and deploys basic knowledge tools such as a constraint browser. Knowledge assets and tools enable professionals to research questions, alternatives and trade offs. The replacement for searching is the creation of ontological hierarchies filled with abstract models defined semantically by the relationships and associations that are explicit within the ontology. The expectation is that critical issues and questions, which used to take months to answer, can be researched and evaluated in hours to reach the point of decisive recommendation.

Stage 2. The next stage of development builds new tools for knowledge and information acquisition and machine learning. This includes self-learning capabilities for (1) updating situation awareness, (2) incorporating advances in theory, and (3) expanding the range of policy research and decision-making that the knowledge stack is capable of addressing. Knowledge is an active asset. The value that this system provides continues to amplify as the knowledge base grows and the system gains more experience—which it does as people use it. The system has features that enable it to learn and improve its reasoning and communications skills.

Figure 13: Systems That Know, Learn and Evolve. This diagram depicts stages in the evolution of intelligent products and processes.



Stage 3. The next wave of development would focus on communications capabilities. These include semantics-based natural-language generation. One development is the capability to produce good-quality briefing books and presentations and other communications directly from the system. Another is the system's capability to teach what it knows. The system can compose lesson plans, conduct sessions, answer questions and customize materials to the needs and preferences of the learner.

Stage 4. A further round of development gives the system the capability to speak, listen and write so that the system is capable of communicating effectively in multiple languages. Language proficiency can reach level 3 (requiring both subject matter and cultural semantics) in the framework developed by the Interagency Language Roundtable. Translating from one language into another can reach a practitioner's level of performance—transitional to professional, and much more than a rendition. With the acquisition of language skills, any of the system's knowledge tools can carry on a conversation with humans in any language of their choosing. Similarly, the system is capable of assimilating information written in different source languages.

Semantics for Manufacturing

Manufacturing paradigms have changed. Manufacturers used to focus on regional market dominance, vertical process integration and strategies to contain local labor costs. Now, the model is to design where the knowledge is, manufacture where labor and other factors are most economical and compete in global markets. Studies and research over the past decade have focused on how to build advanced integrated manufacturing technologies and processes. A key theme is the role of knowledge-based technologies in "smart" products and processes.

According to Craig Schlenoff of NIST (National Institute of Standards and Technology), the evolution toward smart products and processes starts with common models of data, then advances to explicit, formal semantics (dealing with the relationships rather than just the terminology), to self-describing systems, and eventually to self-integrating systems. As shown in Figure 13, the journey doesn't stop there. The goal is to create autonomic and autonomous systems that know, learn and can reason as people do and can self-evolve.

Currently, the aims of advanced manufacturing studies are to develop methodologies and approaches to machine learning and rational theory construction in every area well practiced by humans. Originally, this goal was targeted to 2010–2015, dates set by DoD and NASA for large-scale introduction of autonomous aircraft and intelligent robotic planetary and giant moon explorers. There, the distances and time delays require systems to both explore and solve their own problems during the mission.

Intelligence in aerospace. Some of our best illustrations of what intelligent machines can do are taken from aircraft designs. These applications are well aware that humans can behave most ignorantly as operators, so the machine itself has to trade off what it is commanded to do against the other imperatives it has for accomplishing its longer-term mission. This was precisely Hal's dilemma in *2001: A Space Odyssey.* (As explained in the sequel, Hal had balanced his secret instructions from a higher authority against the mission threat posed by Dave and the rest of the crew in conspiring to turn him off.)

Better examples are found in the way flight and mission computers in airplanes take a pilot's steering commands as "suggestions" rather than overrides. They have to keep such commands from tearing the wings off or steering a perfectly fine airplane into the ground, so they make trade offs between safety and radically unwise control or emergency actions. For example, if an aircraft is inverted and close to the deck, then a pilot-ejection command would kill the pilot by blasting him into the ground. The plane may automatically do a snap-role-sacrificing its wings to point the ejection skyward. These things sound far-fetched, but they are indeed in the avionics instructions of fighter jets.

Similarly, aircraft know their own condition far better than any pilot, and they can report it directly to maintenance crews even before landing. Though the pilot may too tired or stressed to go on to sortie again after landing, the plane may be quite capable of continuing operations. This is why carrier aircraft are switching from large flight sorties to smaller groups operating in "pit stop" fashion, cycling through fueling, arming and reconfiguring for the next mission. The limit to carrier productivity is the surge-sortie rate, and the limits are defined by exhaustion of the deckboard launch and handling crews. Pit-stop sequences conserve launch crews, because maintenance crews know what each plane needs long before it returns from its last mission, and because most of the planes can be serviced on deck. Currently, aircraft life cycles and refits are tied strictly to flight hours, but airplanes that had only three computers before now have close to 50. So the subsystems can monitor their own operating condition and expected lifetimes, and can carry this information and history from one platform installation to the next.

Strategies for managing aircraft life cycles vary considerably based upon availability of platform and subsystem replacements. Sometimes, management favors phasing out whole model lines at about the same time to guarantee comparable efficiency for all operating units and to limit maintenance training to just one generation of machines. Where replacements are uncertain, the strategy is to keep a few always working by scavenging parts from the ever-present "hangar queens" that always have some undiagnosable malady.

As we enter the knowledge age, one expectation is that industry will begin moving away from unique, rapidly changing hardware aggregations toward longer-lived platforms enduring for tens to hundreds of years. New hardware "limit machines" will be engineered to be flexible hosts for virtually any function within the everyday environmental limits associated with a wide range of locations and situations of regular and extraordinary use. There are many examples of such platforms in aviation (DC-3, B-52, C-130, etc.), but the knowledge age will see this strategy applied across a range of industries and product categories. These long-lived hosts will tend to remain in continuous production for extended periods so as to continually supply and evolve capabilities operating near the physical limits to material and ultimate system performance for type. These will minimize scarce resource usage and fully close the materials-recycling loops, providing an equalizing foundation for all of the global civilization.

Dawn of the Knowledge Age

In this article, we have focused on the semantic wave as a revolution in content with major economic consequences.

We started by explaining that digital semantics are all about representing more and more of the things we can know about something as a new kind of data that can be processed with computers. The trend is toward representing this knowledge as data.

We examined the role of semantics in content management, libraries and research We saw that knowledge representation plays a vital role in the present and future of content search, and it is worth millions of dollars in productivity. We reviewed the capabilities and limitations of forms knowledge representation. We learned that it is possible to overcome limitations of language-based approaches to knowledge representation through semantic-form knowledge.

Next, we investigated semantics for computing. We found that process semantics are more than strategic; they are absolutely essential for the success of all major lines of information-technology development, starting with the semantic web. All of the major consumer electronics companies, telecommunications companies and IT companies are weighing in, and the market stakes, which are already vast, are rising. Yet here, too, we learned that language-based semantics and object-oriented procedural models of computing may not be enough to win the day. We identified context computing as a key focus for a new, unified processing paradigm. And we talked about declarative-knowledge computing as, potentially, an ultimate solution to the problem of process complexity and language ambiguity.

We discussed semantic-form knowledge as a driver for labor productivity. We explored new categories of knowledge tools for research, design, planning, analysis, simulation, and decision-making that have the potential to revolutionize professions, management and most knowledge-worker job categories. Taking a longer view, education is destined to be transformed as well.

We saw that publishers and professional groups stand to win big in the knowledge age. They own the reference assets it will take to jump-start new markets. They have the organizational and editorial disciplines needed to build knowledge assets. And they have the established customer relationships needed to dominate new markets.

We briefly touched on semantics across government to sketch the type of knowledge-based capabilities that might be brought to bear on policy-making, defense, intelligence, program management, regulation and public information.

Lastly, we examined semantics for manufacturing and discovered the mainstream role that knowledge technologies will play in advanced manufacturing processes and fundamentally new categories of intelligent products and services. For manufacturers, the market stakes are high, because the economic impact of competitiveness and new markets will be played on a global scale.

The bottom line is that we are at the dawn of the knowledge age. In this report, we've only been able to sketch broad outlines of a major transition coming for the world economy that will occupy several decades *(see Figure 14)*. Knowledge technologies based on science and engineering will power economic expansions measured in the trillions of dollars world-wide. The economic driving force, as we pointed out in Part 1, is a hundred-fold shift in the economics of knowledge (as contrasted with information). The impacts cut deeper and have a much wider scope than previous waves. First, knowledge technologies impact the life-cycle costs of labor and education. Second, knowledge technologies directly affect the global competitiveness of entire industries, especially IT and manufacturing. Third, knowledge technologies open major new markets for "smart" products, services and processes that tap new sources of value. And fourth, knowledge technologies establish the horizons and means for a level of global planning and coordination that is unprecedented in human history. **TSR**

About the authors

Mills Davis is managing director of Project10X. A researcher, visionary and industry consultant for more than two decades, Davis focuses on improving profitability through process innovation, technology investment and strategic partnerships. Previously, he directed the Digital Roadmaps Project and the HiFi Color Project. He can be reached by e-mail at mdavis@project10x.com or by phone at 202-667-6400.

Quark, Canon, Adobe jostle for Japanese market share

XPress 6.1J was a yawner, but Canon's Edicolor sparked interest. Adobe's Creative Suite debuted, too, catching our eye with new typographic functions in Illustrator CS.

The recent JAGAT Page show, a annual early February prepress rite, was a good opportunity to catch up on recent developments in the Japanese market. There was not much in the way of new hardware: Anything worth showing had been shown already at IGAS or is still under wraps for a Drupa rollout. On the software front, though, there was plenty to see. The biggest announcement by far was the unveiling, finally, of Quark XPress 6.1J. Unfortunately, Quark pitched the announcement so low-key that it sounded more like a whimper than a roar.

Orient XPress. The official rollout took place at a free seminar hosted by Quark Japan president Yoko Hayakawa, Quark media house specialist Ajit Singh from India and product manager Hiroyuki Nishimura. The first slide of the first presentation said it all: "Productivity *vs.* Creativity"—guess which concept is Quark's message to the Japanese market. If this were a political campaign, the title would read: "Don't change horses in the middle of the stream."

This is not a bad strategy if the aim is to hang on to an established market share. But it doesn't make sense for the Japanese market, where over 50 percent of prepress production is still done on proprietary systems. That market segment has a more basic choice to make: What is the most cost-effective replacement for our proprietary system that still gives us most of the functionality we currently have? The answer is not Quark XPress 6.1J.

Quark's product demonstration made it abundantly clear that 6.1J is a simple localization of the current English product. There is not a single new Japanese feature (over the currently shipping product, Quark Xpress 4 J). This means there is still no support of advanced Japanese OpenType features such as extended character sets; no advanced typography features (features that you get for free with Apple's MacOS X Hiragino font and that are accessible via the TextEdit application in MacOS 10.3); and none of the sophisticated vertical layout you get with Adobe's InDesign J.

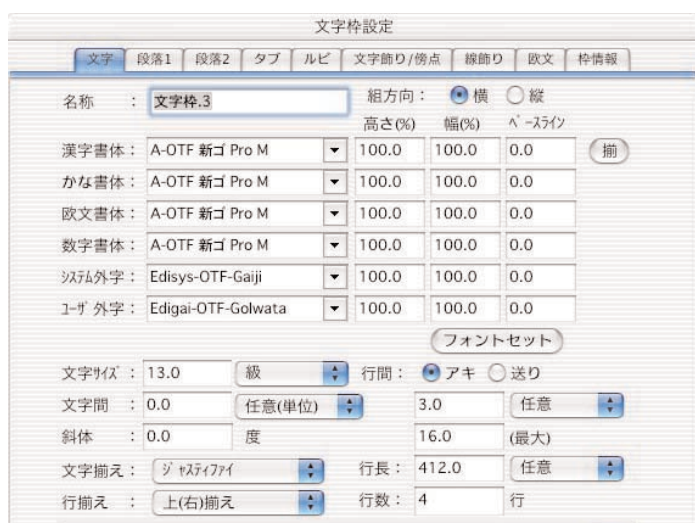
The out-of-touch atmosphere was illustrated in Singh's presentation and comments: "Publishing on paper will continue for the next 20 years." "Quark is becoming an open company with an open heart and open mind." And finally, "XPress is still easy to use…The design is exceptionally good. The output is unmatched." This last comment may be true, but after looking at the competition and Quark Japan's marketing message, we think the result may be one long, slow decline for XPress in the Japanese market.

Canon updates Edicolor. Edicolor is a Japanese layout application with a long pedigree going back to 1993. At Page, the development team from Canon was on hand to show off the newest version. Edicolor 7.0 adds native MacOS X support, Unicode support and a new feature called Virtual Font for printing legacy Japanese PostScript fonts that reside on the printer but not necessarily on the computer. The table features are almost on par with InDesign J, and there are some original touches such as live linking with Excel data.

One interesting new feature, coming in a minor upgrade this month, is the ability for users to add custom-made *gaiji* characters (special characters that fall outside of standard Japanese encoding sets) and edit the font-ID and name data directly. Edicolor has never had the big marketing muscle of Adobe InDesign or the market share of Quark XPress, but Canon has taken good care of its customer base. This has paid off, as big publication houses such as Recuit have chosen Edicolor for production.

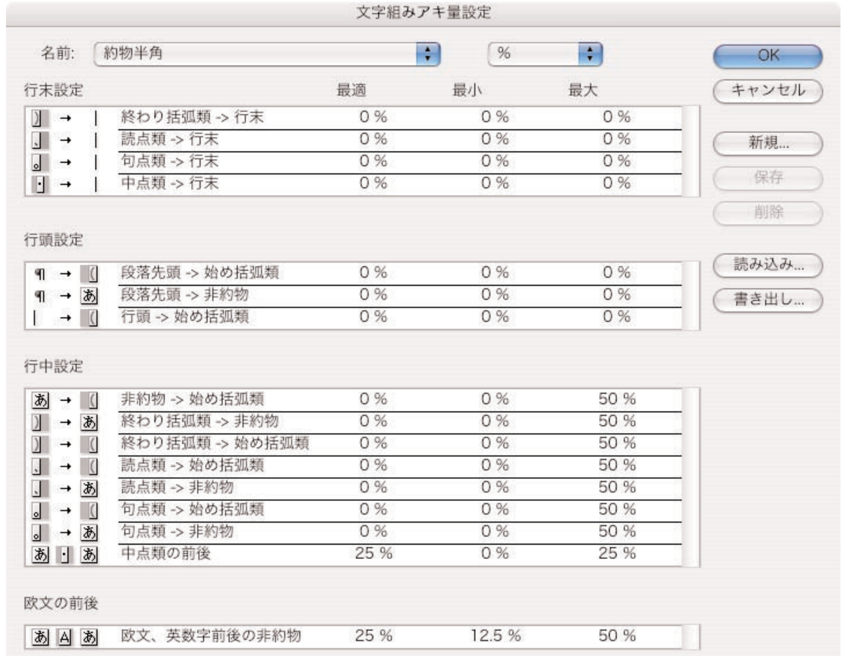
The TNG co-marketing venture of Apple, Adobe, Screen and Morisawa was out in force to help unveil Adobe's new Creative Suite applications. The most important of these was Illustrator CS, which has gained major new Japanese features. In addition to full OpenType support, advanced typography and sophisticated settings for Japanese punctuation and line breaks, there is the abili-

At last. QuarkXPress 6.1J is finally to be shipped in the late spring. Will it make a difference in light of the rich Japanese features of InDesign and Illustrator CS?



Composite controls. Edicolor had composite-font settings long before Illustrator added them.

Framing the data. One nice feature of Edicolor, not found in InDesign J, is the ability to specify a frame around a text grip.



High-end type. Among other improvements, Illustrator CS adds sophisticated punctuation and line-break control.

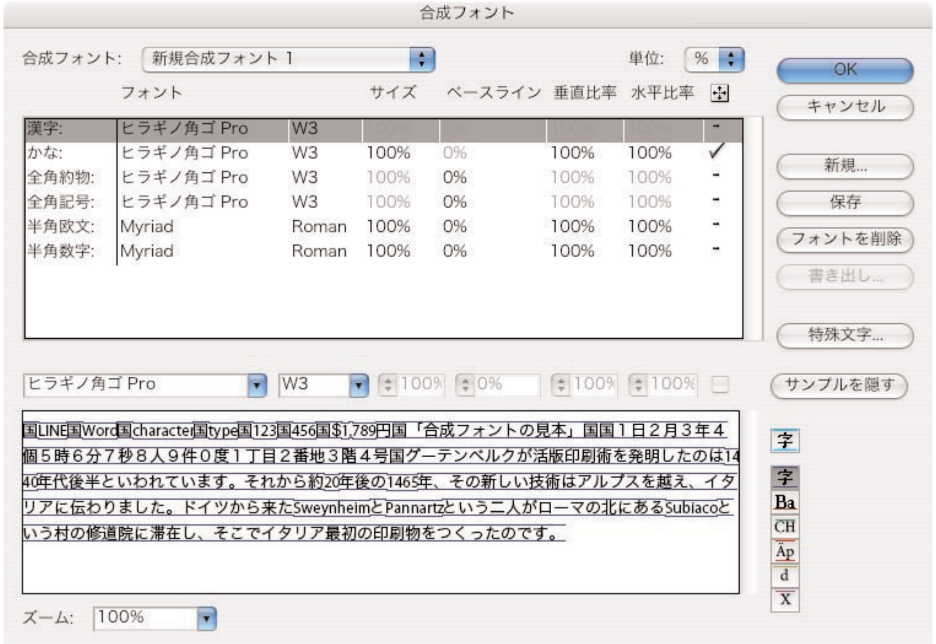
Advantage of OpenType. A side-by-side comparison of OpenType in the Adobe Japan 1-4 character set *(above)* with the Adobe Japan 1-5 character set in Mac OS X's Hiragino font *(below)*. There are many more glyph variations available, which can be a real timesaver when creating flyers.

ty to create composite fonts. Japanese designers like to mix-and-match fonts in a single style: one font for Japanese characters, a western font for Roman characters and yet a different Japanese font for Japanese Hiragana and Katakana scripts. Illustrator CS handles this easily. More than InDesign, Illustrator CS will probably deliver the promise and power of Japanese OpenType to the average designer and drive the OpenType font upgrade cycle.

Another year, another font-encoding specification

The last day of Page had a conference session on Japanese font-encoding with a distinguished panel of font experts from Apple, Microsoft, NTT and Kyodo News. There was a lively discussion of the forthcoming Unicode 4.0 specification and a few new tweaks to the Japanese Industrial Standard (JIS) X 0213:2004. Microsoft, predictably, plans to ignore Adobe's Asian encoding efforts (AJ 1-4 and AJ 1-5) and to focus on Unicode for Longhorn, while Apple will be on friendly terms with both standards.

Apple's Yasuo Kida made the interesting comment that Unicode is sometimes weak in dealing with context. In Japanese, as in a few other languages, the difference between a glyph and a character is not always clear: Sometimes a character is a character; but, in a certain context, it is



Better composites. The Illustrator CS Composite Font Preference dialog is sleeker than Edicolor's.

also a glyph variation. Kida proposed having a ideograph variation selector to help define the difference between glyphs and characters.

It was clear that Japan's font experts are ready to throw in the towel. JIS X 0213 will probably be the end of the line for homegrown font-encoding standards. There will be tweaks (2001, 2002, 2004, etc.) but no major new effort. One Japanese font designer said, "Many Japanese font people fought Unicode for a long time, but they can't anymore." It was a matter of pride, but from now on, Unicode will lead the font-encoding parade.

Adobe does a favor. Behind the scenes, there was talk of yet another Adobe Japan character set, AJ 1-6, which will probably incorporate Unicode 4.0 and the U-Press character set for newspapers. To our eye, Adobe is doing the Japanese font industry, and the users, a big favor by publishing a font-ID specification rather than letting everybody stumble over each other with private ID mappings.

Adobe is also working on a character-creation InDesign plug-in for the Asian market that should do well in Japan. The high-end market always needs better gaiji-creation tools, but only Adobe has the deep pockets and talent to deliver one. Making it part of InDesign CS will only increase its power in the Japanese market. The tool should appear with the next CS upgrade, which is due in the summer of 2005. The file format of InDesign CS is also due for a makeover, and it is said the next one will become XML-based.

Joel Breckinridge

Creo to debut third-generation VLF imager, spot-color workflow, gravure system

Magnus was designed for automation from the start; exposes plates for the largest available offset presses.

At Drupa, Creo will show a VLF product, Magnus. It has many similarities to the Trendsetter VLF, but it has a smaller footprint. (But it still takes more space than the competitive machines from Agfa and Screen.) Automated plate loading has been designed in from the start.

Magnus will be available in four models: the 4570, 5183, 5570 and 6383. Each model number indicates the drum size in inches. The Magnus 6383, at 63×83 inches, is the largest VLF machine available. Magnus will be offered at three speed grades, the fastest of which can expose 15 full-size (up to 80.7×59.4 inch) plates per hour. Alternatively, it can load two B1 plates at a time and thus crank out 31 plates/hour. Creo claims that this is the fastest fully automated CTP device on the market, although we believe that both the Agfa Xcalibur VLF XXT and the Screen PlateRite Ultima 32000Z may be faster.

Spotless. Spotless is Creo's approach to eliminating the use of spot-color inks. There will be two versions: Spotless 4 for CMYK and Spotless X for five or more colors. The process is built around two of Creo's core technologies, SquareSpot process control and Staccato FM screening. The Spotless software accurately predicts on-press color within the prepress workflow and automatically converts spot to process colors, taking printing conditions (press, paper, ink) into account.

Spotless will have plenty of competition from vendors in packaging and from Agfa, which has been offering Pantone's Hexachrome. It will be interesting to compare the results of all these products.

Gravure. Creo has a joint development partnership with Italian supplier Acigraf to produce the Exactus thermal gravure system. Exactus uses Creo's SquareSpot imaging at 3,200 dpi to expose a mask on the cylinder, which is then chemically etched.

In the Exactus process, a blank copper cylinder is coated with a thermal resist called Graviti. This compound is similar to the thermal mask that Creo developed when it attempted to break into the printed-circuit-board market. The resist is non-ablative and white-light tolerant. After exposure, the excess resist is dissolved in a water-based solution and the cylinder is conventionally etched.

The system is claimed to reduce the density-management and depth-control problems that gravure is heir to, and to reduce the cost of cylinder preparation compared with engraving systems. It is also said to be faster than engraving systems—and advantage that will be more apparent for large cylinders, because the etching process is the same for large and small cylinders.

Andy Tribute

In *The Bulletin* Since Last Issue

Volume 9, Number 21
February 18, 2004

Adobe takes aim at corporate DRM. Sensing an opportunity to leverage its experience in digital rights management (DRM) for commercial media by testing DRM in the corporate sector, Adobe has introduced Policy Server, a new server for setting and enforcing document-security policies. The Policy Server provides tools for applying security controls to PDF documents, but it is new technology developed by Adobe specifically for the corporate market.

HP to update Indigos, Designjets; touts CMYK Plus. HP announced that it will debut the first HP press to be developed entirely under HP auspices at Drupa. Called the HP Indigo 5000, HP's engineering contribution is extensive monitoring of internal conditions that supports a failure-prediction system. The press runs at 4,000 four-color A4 pages per hour, or 16,000 single-color pages per hour. It will be offered with four- to seven-ink printing. HP also plans to announce updates to its Designjet line of large-format printers and debut CMYK Plus color-management software.

Chinese developer tackles Microsoft. Evermore Software LLC, China's leading developer of office software, is launching a challenge to the Microsoft Office monopoly. It announced an English-language version of Evermore Integrated Office (EIOffice). EIOffice consolidates the components (word processor, spreadsheet, slide-maker) of a conventional Office suite into a single application. EIOffice stores all text, worksheets, graphics, audio, video and slides in one file format. The software is written in Java and runs on several platforms, including Windows and Linux. (Mac and Solaris are said to be in the works.) It imports and exports Microsoft Office xls, doc and ppt files, and Evermore says that files are compatible if users transfer data back and forth between programs. Users may also save documents in PDF, rtf and txt formats. Starting in May, Evermore will lease, rather than sell, EIOffice 2004 for $99 annually or $249 for three years.

Newsstand. Agfa, Thieme to develop ink-jet printer; Arbortext, DMSi explore XML in Word; Apago releases PDF Enhancer 2; Dot-Photo offers talking pictures; Xeikon adds Finnish distributor; Altona Test Suite kit now available; Océ debuts new engine; Quark releases XPress update; Enfocus PitStop Server 3.1 is PDF 1.5-compatible; IXOS, Open Text release ECM platform; Ixiasoft intros version 3.0 of TextML; Creo acquires second plate-manufacturing business.

Insider Perspective: Peter Luit explains why innovation is needed in PDF workflows.

Volume 9, Number 22
February 25, 2004

Agfa, Creo, Océ lift Drupa product veil. Agfa, Creo, Océ and Xeikon preannounced the products they will be showing at Drupa in May. New CTP machines and plates, updated workflow systems and new digital printers were among the more significant announcements. Details will follow in an upcoming issue of *The Seybold Report*.

Microsoft battles security concerns. Two years into its "trusted computing initiative," Microsoft reaffirmed its commitment to fight malicious hackers by improving the security controls of its operating systems. In a keynote speech delivered at the RSA Security Conference, Microsoft chairman Bill Gates announced Microsoft's partnership with RSA Security to develop better log-in security and demonstrated an improved firewall that will be part of the upcoming Service Pack 2 for Windows XP due later this year.

Ascender font-development company debuts. Some well-known font-industry veterans have joined forces to start a new font-development and licensing company. Called Ascender Corp., the Chicago-based firm's founders include the founder and former president of Monotype Typography, Ira Mirochnick; former VP of marketing for Agfa Monotype, Bill Davis; and font designers Steve Matteson and Thomas Rickner.

Newsstand. Time deals with HP for archive; Visual Data joins DAM market; Adobe PDF JobReady 1.6 now available; Easypress releases XML white paper; Hummingbird unveils Enterprise 2004; Markzware releases version 3.1 of FlightCheck Workflow.

Insider Perspective: Hans Hartman looks at the relationship between PDF and XML. Do they compete or are they complementary?

Periodicals Postage Paid
at Media, PA and
additional mailing offices