
8 Knowledge Management

Kimiz Dalkir

CONTENTS

Introduction.....	111
Definition of Knowledge Management.....	111
The Multidisciplinary Nature of KM.....	112
The Intellectual Capital Perspective.....	114
Core KM Concepts	115
The Knowledge Processing Cycle	117
History and Evolution of KM	118
Emerging Roles for Information Professionals.....	120
Conclusion	122
References.....	122

INTRODUCTION

This entry provides an overview of knowledge management (KM), both as a scholarly discipline and a professional field of practice. Key terms such as “knowledge” and “knowledge management” are defined. Core concepts such as intellectual capital, lessons learned, best practices, and value-added knowledge reuse are introduced and defined. Key stages in the knowledge processing cycle are defined and described. A brief historical chronology of KM is outlined to show the multidisciplinary roots and the evolution that has taken place to date. The emerging roles for information professionals in KM are presented and the links to information studies are emphasized. The primary goal of this entry is to illustrate how KM that rests on a solid foundation of information management emerges as a stronger and more rigorous field of study and practice.

DEFINITION OF KNOWLEDGE MANAGEMENT

There is no universally accepted definition of KM. Part of this stems from the fact that there is a lack of consensus in defining knowledge. A sample dictionary definition of knowledge is

The fact or condition of knowing something with familiarity gained through experience or association; acquaintance with or understanding of a science, art, or technique; the fact or condition of being aware of something; the range of one’s information or understanding (e.g. answered to the best of my *knowledge*); the circumstance or condition of apprehending truth or fact through reasoning; the sum of what is known: the body of truth, information, and principles acquired by humankind; cognizance, awareness, learning.^[1]

It is useful to view the knowledge in KM as being composed primarily of experiential knowledge that is highly subjective in nature. This can be contrasted with information that, although subject to different interpretations, is typically thought of as having a more neutral and verifiable nature. The same dictionary defines information as

Knowledge obtained from investigation, study, or instruction; intelligence, news data or facts; the attribute inherent in and communicated by one of two or more alternative sequences or arrangements of

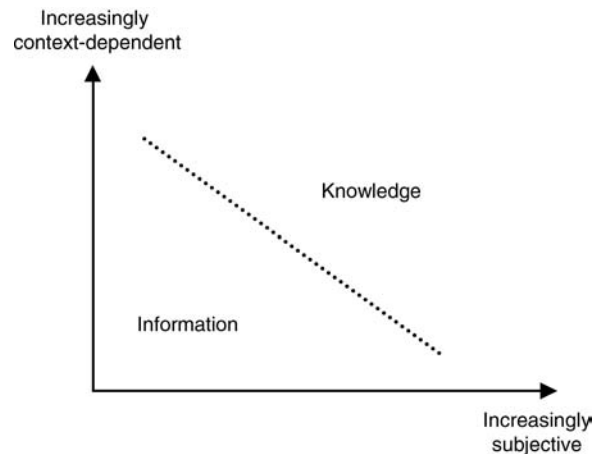


FIGURE 8.1 Key attributes of knowledge assets.

something (as nucleotides in DNA or binary digits in a computer program) that produce specific effects
c (1): a signal or character (as in a communication system or computer) representing data (2): something (as a message, experimental data, or a picture) which justifies change in a construct (as a plan or theory) that represents physical or mental experience or another construct.^[1]

Other definitions from the KM literature include

Knowledge is a subset of information: it is subjective; it is linked to meaningful behaviour; and it has tacit elements born of experience (p. 113).^[2]

Knowledge is always recreated in the present moment. Most of us cannot articulate what we know. It is largely invisible, and often comes to mind when we need it to answer a question or solve a problem.^[3]

Knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of those who know. In organizations, it often becomes embedded not only in documents or repositories but also in organizational routines, processes, practices and norms (p. 5).^[4]

A typical example of information would be to read the listings of movies being shown on a given week, reading the reviews, and consulting the times and places the film will be showing at. An example of knowledge would be calling on a friend who has, in the past, demonstrated an uncanny inverse relationship with respect to likes and dislikes: if this person liked a movie, experience has shown that one would hate it and vice versa. Knowledge is also brought into play in making a decision to drive or take public transport, given the time of year, possibility of inclement weather, and the rapid trend analysis that concludes that it has never been possible to find parking in time to see a movie being shown downtown.

A key point is that not all information is of value to a given individual, group, or organization. A knowledge asset is defined as knowledge that is of value, which is more context dependent and more subjective in nature, as illustrated in Figure 8.1.

THE MULTIDISCIPLINARY NATURE OF KM

There is a great diversity to be found in KM definitions. This is due to the multidisciplinary nature of the field. In fact, most definitions appear to provide only one piece of the KM definition puzzle, as they are heavily influenced by their particular discipline. Table 8.1 lists some of the different disciplines and specialties that have contributed to and that make use of KM.

TABLE 8.1
Diverse Disciplines that Contribute to Knowledge Management

1. Organizational development, organizational science, and organizational learning
2. Cognitive science, individual learning, and mental models
3. Linguistics and computational linguistics
4. Design of information technologies such as knowledge-based systems, document and information management, electronic performance support systems, and database technologies
5. Information and library science, classification, and taxonomy
6. Technical writing and journalism
7. Anthropology and sociology, sociograms, and social network analysis
8. Education and training
9. Storytelling and communication studies
10. Collaborative technologies such as Computer Supported Cooperative Work (CSCW) and groupware as well as intranets, extranets, portals, and other Web technologies

Here are a few sample definitions from the business perspective:

Knowledge management is a business activity with two primary aspects: treating the knowledge component of business activities as an explicit concern of business reflected in strategy, policy, and practice at all levels of the organization; and, making a direct connection between an organization's intellectual assets – both explicit (recorded) and tacit (personal know-how) – and positive business results.^[5]

Knowledge management is a collaborative and integrated approach to the creation, capture, organization, access and use of an enterprise's intellectual assets.^[6]

There is much more to knowledge management than technology alone. Management is a business process (p. 95).^[7]

Some definitions from the cognitive science or knowledge science perspective:

Knowledge—the insights, understandings, and practical know-how that we all possess—is the fundamental resource that allows us to function intelligently.

Over time, considerable knowledge is also transformed to other manifestations – such as books, technology, practices, and traditions – within organizations of all kinds and in society in general. These transformations result in cumulated expertise and, when used appropriately, increased effectiveness. Knowledge is one, if not THE, principal factor that makes personal, organizational, and societal intelligent behavior possible.^[8]

Some sample definitions from the process/technology perspective:

Knowledge Management (KM) was initially defined as the process of applying a systematic approach to the capture, structuring, management and dissemination of knowledge throughout an organization to work faster, reuse best practices, and reduce costly rework from project to project.^[9–12]

American Productivity and Quality Center (APQC)^[13] defines knowledge management as

. . . the managing of knowledge through systematic sharing. Even in highly sophisticated modern knowledge organizations, the most valuable knowledge—the know-how in terms of what really gets results and what mistakes to avoid—often resides mainly in people's minds. Knowledge Management works towards migrating that knowledge from one person to a wide-range of individuals within an organization.

Some “people-oriented” definitions:

Knowledge management is not seen as a matter of building up a large electronic library, but by connecting people so they can think together (p. 104).^[3]

Most executives seem to understand that knowledge is highly people-based, but they are stuck with an investment model that is geared primarily toward technology implementations (p. 86).^[12]

Knowledge management and its varied definitions need to achieve a balance: there cannot be an overemphasis on any one of the key dimensions, such as technology.^[14] Effective KM (and comprehensive KM definitions) should include people, process, technology, culture and measurable organizational objectives.^[15,16]

Wiig^[8] also emphasizes that given the importance of knowledge in virtually all areas of daily and commercial life, two knowledge-related aspects are vital for viability and success at any level. These are knowledge *assets* that must be applied, nurtured, preserved, and used to the largest extent possible by both individuals and organizations; and knowledge-related *processes* to create, build, compile, organize, transform, transfer, pool, apply, and safeguard knowledge. They must be carefully and explicitly managed in all affected areas. According to Wiig^[6]

Historically, knowledge has always been managed, at least implicitly. However, effective and active knowledge management requires new perspectives and techniques and touches on almost all facets of an organization. We need to develop a new discipline and prepare a cadre of knowledge professionals with a blend of expertise that we have not previously seen. This is our challenge!

THE INTELLECTUAL CAPITAL PERSPECTIVE

The focus of intellectual capital management (ICM), on the other hand, is on those pieces of knowledge that are of *business value* to the organization—referred to as intellectual capital or assets.^[17] While some of these are more visible (e.g., patents, intellectual property), the majority consists of know-how, know-why, experience, and expertise that tend to reside within the head of one or a few employees.^[18,19]

Knowledge management is often characterized by a “pack rat” approach to content: “save it, it may prove useful some time in the future.” Many documents tend to be warehoused, sophisticated search engines are then used to try to retrieve some of this content, and fairly large-scale and costly KM systems are built. Knowledge management solutions have proven to be most successful in the capture, storage, and subsequent dissemination of knowledge that has been rendered explicit—particularly lessons learned and best practices.

Intellectual capital management is characterized by less content—because content is filtered, judged, and only the best are inventoried (the “top ten” for example). Intellectual capital management content tends to be more representative of peoples’ real thinking (contextual information, opinions, stories) due to its focus on actionable knowledge and know-how, with the result that less costly endeavors and a focus on learning (at the individual, community, and organizational level) results, rather than on the building of systems.

It is essential to identify that knowledge which is of value and is also at risk of being lost to the organization, through retirement, turnover, and competition. The best way to retain valuable knowledge is to identify intellectual assets and then ensure legacy materials are produced, and subsequently stored in such a way as to make their future retrieval and reuse as easy as possible.^[20] These tangible by-products need to flow from individual to individual, between community of practice (CoP) members and, of course, back to the organization itself, in the form of lessons learned, best practices, and corporate memory.

The knowledge capture and transfer approaches described here help to

- Facilitate a smooth transition from those retiring to their successors who are recruited to fill their positions.
- Minimize loss of corporate memory due to attrition and retirement.

- Identify critical resources and critical areas of knowledge so that the corporation “knows what it knows and does well—and why.”
- Build up a toolkit of methods that can be used with individuals, with groups, and with the organization to stem the potential loss of intellectual capital.

Knowledge management is often mistakenly perceived as a process of documenting knowledge held in people’s minds, then storing or archiving this knowledge.^[21] In fact, it is better to view KM in a broader context, one that encompasses all the processes used to generate value from knowledge-based assets. Knowledge management is facilitated by technology and by the culture of an organization. A good way of defining KM is

... the deliberate and systematic coordination of an organization’s people, technology, processes and organizational structure in order to add value through reuse and innovation. This is achieved through the promotion of creating, sharing and applying knowledge as well as through the feeding of valuable lessons learned and best practices into corporate memory in order to foster continued organizational learning (p. 3).^[22]

CORE KM CONCEPTS

Krandsdorff^[23] coined the term “corporate amnesia” to refer to the loss of accumulated expertise and know-how due to employee turnover as people take what they know with them when they leave. The costs of employee turnover to the organization have been well documented (e.g., separation costs, recruitment and selection costs, training of replacements, initial lack of productivity of new hire, and the loss of productivity of coworkers during the transition). Far less research attention has been paid to the cost to the firm of losing know-how that resides within the minds of individual employees who depart. In an era of knowledge workers, learning organizations and service economies individuals are increasingly responsible for value creation. Although many organizations have succession plans in place, the process usually involves transferring know-how from the departing employee to their successor but the whole process has to be repeated again for the next departure. Organizations need to “capture” this know-how and transfer it to a stable, easily accessible, cumulative knowledge base—an organizational memory—to retain and make accessible valuable knowledge gained through the experiences of all knowledge in a continuous and uninterrupted manner.

Organizations need to effectively manage their organizational memory in order to prevent the loss of essential knowledge, particularly knowledge that resides predominantly in the heads of their knowledge workers (“tacit knowledge”) and less in documents, procedures, and other tangible forms (“explicit knowledge”).^[24] More often than not, it is this difficult-to-articulate “know-how” that is of greatest value in organizational competitiveness and viability and which represents the vast majority of experiential know-how. Figure 8.2 shows how the metaphor of an iceberg is often used to depict tacit and explicit knowledge forms.

The National Aeronautics and Space Administration (NASA), for example, has publicly admitted that the knowledge of how to put a man on the moon has been lost.^[25] The lessons that were learned and the innovations that were sparked cannot be found in the collective organizational memory of NASA. This means that NASA’s organizational memory cannot be used as a resource to plan a more effective mission to send another manned flight to the moon or to Mars. A well-designed and well-managed organizational memory does not only combat corporate amnesia, but it ensures knowledge continuity—the effective transfer of know-how amongst peers and to future generations of knowledge workers. A better understanding of the nature of organizational memory, what it should include (content), how it can best be retained (technological containers), and how the accumulated lessons learned and best practices can be used by newcomers (connections), will help mitigate the cost of lost, forgotten, or un-transferred knowledge and know-how.

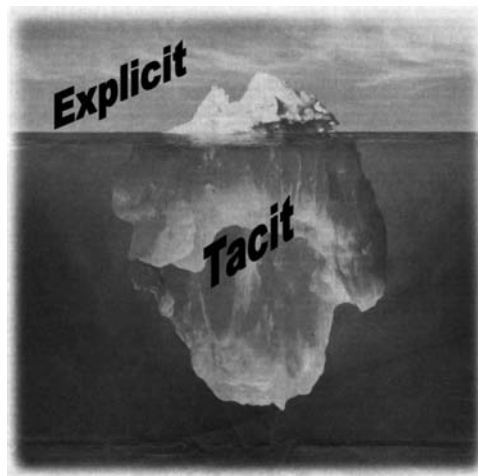


FIGURE 8.2 The iceberg metaphor used to show the respective proportions of tacit and explicit knowledge.

Organizations today face escalating risks of losing strategic knowledge and know-how and face incredible difficulties in recruiting and retaining skilled employees. Many industries face astounding rates of high and constant turnover. In addition, we are experiencing a demographic pressure as baby boomers rapidly approach retirement age, which means that a “critical mass” of knowledge will literally walk out the door over the next five to ten years. Approximately 11,000 baby boomers are turning 50 years old every day.^[26] Over 80% of Canadian government federal executives will be eligible to retire by 2010 and 53% of U.S. federal civil servants will be eligible to retire within the next two years.^[27] The number of workers age 55 and over is expected to increase by 47% the year 2010.^[28] Other vulnerable organizations include high technology companies with very high turnover rates and the military or banking, where constant turnover is the norm given rotational postings.

Most successful organizations will state that their two greatest assets are the people who work for them and the knowledge they possess. The imminent turnover signals a potential for the loss of valuable accumulated knowledge and know-how in the form of the competence and expertise possessed by the departing individuals. This valuable knowledge and know-how exists in both formal or tangible forms, such as documents, but also in less visible forms—often referred to as tacit or difficult-to-articulate knowledge. Particular emphasis must be placed on the tacit form as this often resides within a given individual or group and is therefore more easily and completely lost when the people leave the organization.^[29]

The typical technological tools used in KM are some form of centralized database system an intranet or Web portal to access the database content as well as other organizational knowledge, some type of messaging system (typically e-mail and discussion groups), some content management tools to organize the knowledge resources (e.g., content management software, taxonomy software) and some form of collaborative software (e.g., virtual meeting tools, groupware).

The essential elements that go into organizational memory include best practices, which are innovations or improved practices, processes, and so forth; together with the flip side of the coin: lessons learned, outcomes that were not desirable, or errors that should not be repeated. By carefully capturing, documenting, organizing, and making available accumulated experiential knowledge in the form of best practices and lessons learned, KM can provide a framework within which we can process knowledge throughout its useful life cycle.

Knowledge management typically addresses three levels within an organization: the individual, the group, and the organization. Organizational memory resides at the level of the organization and this is the preservation layer of KM. Innovation and knowledge creation occurs at the individual

level while knowledge dissemination and sharing occur at the group level. In KM, groups of a particular nature, called “communities of practice” (CoPs) are often key elements in ensuring that valuable knowledge flows or moves around appropriately. These CoPs are essentially thematic networks of like-minded individuals who share a common purpose, a common professional zeal, and agree to mutually help out one another to achieve some strategic goals for the organization.^[30,31]

Szulanski^[32] introduced the notion of knowledge “stickiness” to refer to knowledge that was difficult to move around, and thus could provide little in the way of value to individuals, groups, or the organization as a whole. The knowledge processing cycle is used to describe the various stages a particular knowledge resource goes through during the course of its “lifespan” within an organization.

THE KNOWLEDGE PROCESSING CYCLE

Nonaka and Takeuchi described how knowledge was disseminated and transformed as it made its way around an organization in their Knowledge Spiral Model.^[9] Essentially, knowledge is created by an individual, (an innovative idea for example), but remains initially tacit. This innovation or idea is typically shared only with trusted friends and colleagues at first (“socialization”). With time, others may be aware of the innovation and the knowledge may be transformed from its tacit state to an explicit one (“externalization”), by documenting it in the form of text, audio, or other formats. Once explicit, it is easier to disseminate more widely and to preserve in an organizational memory system. At this stage (“combination”), the content may be organized and reworked (e.g., addition of an abstract or executive summary). In the final stage (“internalization”), explicit knowledge once again becomes tacit as an individual learns, understands and “buys into” a knowledge resource (e.g., sees its value and decides to apply this new knowledge). The knowledge resource then continues cycling through the four quadrants in a “knowledge spiral.” Figure 8.3 outlines the major stages in the Nonaka and Takeuchi KM model, as knowledge is transformed, disseminated and applied within an organization.

As knowledge resources circulate throughout an organization, value can be added at each step. A more detailed view of what happens to knowledge as it undergoes transformation from tacit to explicit and back again is provided in the form of a knowledge life cycle. The first stage is to either capture existing knowledge (e.g., already existing in tangible form such as project reports, conference reports, or other documents), or to create knowledge that does not yet exist in documented form (e.g., by interviewing subject matter experts to document a particular procedure). This knowledge resource can then be contributed to or made available for reuse by others. Knowledge is then codified by the addition of sufficient descriptive information about both the content and the format it is in (e.g., metadata describing theme, whether it is a text or also available as a short video clip). At this point, knowledge is further refined by sanitization (removing all sensitive information that may

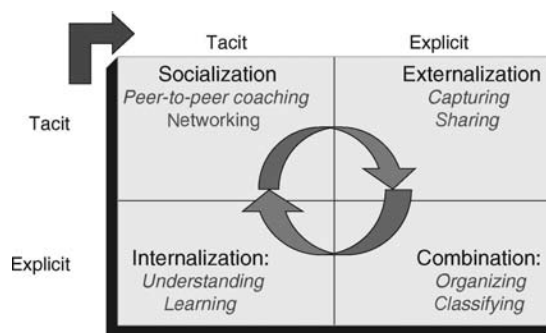


FIGURE 8.3 The Nonaka and Takeuchi knowledge spiral model of knowledge management.

TABLE 8.2
Key Steps in Knowledge Processing Cycle

Step Number	Description
1.	Knowledge capture
2.	Knowledge creation
3.	Knowledge contribution
4.	Knowledge codification
5.	Knowledge refinement (including sanitize) and reconstruction (e.g., synthesis)
6.	Knowledge selection—filter contributions
7.	Knowledge sharing and pooling
8.	Knowledge organization and access
9.	Knowledge learning and application
10.	Knowledge evaluation and reuse or divest (e.g., destroy, archive)

identify a person, place, or event, or other confidential content) and reconstruction (e.g., recombining or synthesizing in a new version). Contributions that have been refined and reconstructed are then subjected to a selective filtering process and those that are deemed to be valuable are made available for knowledge sharing.

Knowledge sharing involves the pooling and distribution of valuable content, either globally (e.g., a “broadcast” mode) or selectively (e.g., targeted to specific users, communities, or based on user profiles, subscriptions, etc.). The knowledge is better organized (classified using a taxonomic scheme) and access is typically provided through an intranet or portal interface. These valuable knowledge resources are then learned and applied by knowledge workers. Each knowledge resource is evaluated periodically in order to ensure continued relevance and to decide whether to discard or to archive at the appropriate time. Table 8.2 lists the key phases in knowledge processing that occurs during the useful life cycle of a knowledge resource, such as a best practice or lesson learned.

HISTORY AND EVOLUTION OF KM

Knowledge and management are concepts that have obviously been around for some time—the combination, “knowledge management,” however, is fairly recent. The emergence of communication technologies that create real-time networks such as the Internet, intranets, e-mail, and so forth have made KM easier to implement.^[33]

A number of management theories have contributed to definition of KM. Two KM gurus were responsible for establishing the field as a serious target for scholarly study: Peter Drucker,^[34] who coined the term “knowledge worker” and who wrote about KM from a management perspective in the 1970s and Nonaka and Takeuchi, who wrote the seminal book *The Knowledge Creating Company* in the 1990s.^[9] Others such as Peter Senge^[35] and Chris Argyris^[36,37] have focused on the “learning organization” and cultural dimensions of managing knowledge. The notion of knowledge as an asset became more widely adopted by the mid 1980s. Early pioneers in information technologies from artificial intelligence and expert systems at Carnegie Mellon University (CMU) and DEC (Digital Equipment Corporation) developed hypertext and groupware tools that allow information to be better shared and organized. Knowledge management-related articles began appearing in prestigious journals such as the *Harvard Business Review* and *Organizational Science* and a number of foundational books were published such as *The Fifth Discipline* and *The Knowledge-Creating Company*. Table 8.3 illustrates the major milestones in KM history.

In 1989, the International Knowledge Management Network (IKMN) was founded in Europe and was joined by the U.S.-based Knowledge Management Forum in 1994.

TABLE 8.3
Major Milestones in KM History

Year	Source	Event
1980	DEC, CMU	XCON Expert System
1986	Dr. K. Wiig	Coined KM concept at UN
1989	Consulting firms	Start internal KM projects
1991	HBR article	Nonaka & Takeuchi
1993	Dr. K. Wiig	First KM book published
1994	KM network	First KM conference
Mid-1990s	Consulting firms	Start offering KM services
Late 1990s	Key vertical industries	Implement KM and start seeing benefits
2000–2003	Academia	KM courses/programs in universities with KM texts

Businesses began implementing KM solutions in earnest in the 1990s. At first, projects tended to be technological implementations, mostly portals and knowledge repositories. This was followed by a wave of “people not technology” emphasis that led to the popularity of thematic networks or CoPs. A third wave emerged as knowledge workers, much like information workers, found themselves faced with “content overload” which in turn led to a number of content management, knowledge organization, or taxonomic classification projects to make the knowledge more readily accessible to users.

With the new millennium, KM faded a bit from public view, likely due to disappointment, the result of expectations having been raised too much by “KM hype” and focusing too much on technologies that were largely left untouched. Practitioners and researchers came to the realization that converting theory into practice was more challenging than previously thought. Among the missing elements were ensuring that KM was not being done for KM’s sake but that there was a clear link to organizational objectives and that all participants saw value in KM (incentives for knowledge sharing, for example). By 2004, the field had finally shaken its burden of being yet another management fad. Organizations appreciated the holistic nature of KM and paid attention not only to tools but also to organizational culture requirements for successful KM solutions.

In parallel, the nature of work evolved, with collaboration becoming the default rather than the exception, not only for our professional lives but social interactions and learning objectives. Knowledge sharing with a group of trusted peers has become the de facto way of working, learning and living. Knowledge management found itself in tune with these changes that greatly facilitated its adoption as a philosophy, a mindset, a strategic perspective, and a way of interacting to create, share, and use valuable content.

Knowledge management today is being taught in universities around the world, typically in business, computer science, education, and information studies departments. Doctoral students are completing Ph.D. theses on KM topics. International conferences are being held around the world, with both practitioner and researcher tracks. A number of consortia and KM professional organizations have been established. The number and type of KM journals have also changed. While at first publications were primarily authored by KM tool vendors and practitioners, and in more general management, computer science and organizational science journalists, there has been a significant shift to more scholarly venues and publications that represent KM research rather than KM project mandates. Along with the proliferation in the number of journals, there has also been increasing specialization (e.g., *International Journal of Nuclear Knowledge Management*), more interdisciplinary journals (e.g., *Interdisciplinary Journal of Information, Knowledge and Management*), and more emphasis on international KM (e.g., *International Journal of Knowledge Management*). Up until about the year 2000, only non-peer reviewed journals were available (e.g., *KM World*, *KM*

TABLE 8.4
Major Peer-Reviewed KM Journals

Journal Title	Year Founded
<i>IBM Systems Journal</i>	1962
<i>Journal of Managerial and Organizational Learning</i>	1970
<i>Journal of Management Information Systems</i>	1984
<i>Journal of Organizational Change Management</i>	1989
<i>Knowledge and Process Management</i>	1996
<i>Journal of Knowledge Management</i>	1997
<i>Journal of Management Studies</i>	1997
<i>Journal of Knowledge Management Practice</i>	1998
<i>Journal of Intellectual Capital</i>	2000
<i>E-Journal of Organizational Learning and Leadership</i>	2002
<i>Journal of Information and Knowledge Management (JIKM)</i>	2002
<i>Electronic Journal of Knowledge Management</i>	2003
<i>International Journal of Knowledge, Culture and Change Management</i>	2003
<i>Knowledge Management Research and Practice</i>	2003
<i>Interdisciplinary Journal of Storytelling Studies</i>	2004
<i>International Journal of Nuclear Knowledge Management (IJNKM)</i>	2004
<i>International Journal of Knowledge Management</i>	2005
<i>Knowledge Management for Development Journal</i>	2005
<i>International Journal of Knowledge Management Studies</i>	2006
<i>Interdisciplinary Journal of Information, Knowledge and Management (IJKM)</i>	2007

Review). Some of the major peer-reviewed KM journals available today are shown in Table 8.4, ordered by first year of publication.

Knowledge management is found in all vertical industry sectors (e.g., banking, pharmaceutical); in all sizes of organizations (including volunteer-run associations with less than 10 full-time employees); profit and nonprofit organizations and organizations with a strong hierarchical structure (for example, the U.S. military was one of the early pioneers of KM). Organizations today are defining and staffing KM positions. Knowledge management is increasingly perceived as an element that cuts across organizational structures, projects, and silos—it enjoys a status similar to that of “quality assurance” or “ethics” and it has become a critical element of organizational viability. Knowledge management is no longer something interesting to explore—it has become subsumed in the way in which individuals, groups, and organizations work, learn, and remember.

EMERGING ROLES FOR INFORMATION PROFESSIONALS

The KM field has transformed from one led primarily by consultants and other KM practitioners to a bona fide discipline, with a distinct body of knowledge. This has been paralleled by the growing number of academic programs that offer KM as compared to the predominately private sector training that had been the only way to learn about KM up until now.^[38] TFPL^[39] is a specialist recruitment, advisory, training, and research services company with offices in London focusing on KM, library and information management, records management, and Web and content management. Since 1987, TFPL has worked with organizations in both public and private sectors to help them develop and implement knowledge and information strategies and to recruit and train information and knowledge leaders and their teams. TFPL has drafted a KM skills and competencies guide to provide a clear and practical overview of KM skills and competencies that draws on

the practical experience of organizations in a wide range of sectors and with varying approaches to KM.

In general, these KM skills include

- Time management to use their time and energy effectively for acquiring knowledge (spending all day surfing the Net is probably counterproductive).
- Use of different learning techniques to absorb key knowledge and learning quickly.
- Effective skills of advocacy and inquiry to present knowledge to, and gather knowledge from, others.
- Informal networking skills to build influence to gain access to people with knowledge.
- Resource investigation skills.
- Effective IT skills for recording and disseminating information.
- Skills of cooperative problem solving.
- Open dialogue skills.
- Flexibility and willingness to try new things and take educated risks.
- Active review of learning from mistakes, risks, opportunities, and successes.

The TFPL KM skills inventory^[40] is based on an extensive international research. The project team contacted over 500 organizations involved in implementing KM, and identified the roles that they had created, the skills that were needed in those roles, and the additional skills that were required across the organization. These key skills included an understanding of the KM concept—the philosophy and theory—and an awareness of the experience of other organizations in developing KM solutions and approaches; an understanding of, and the ability to, identify the business value of KM activities to the organization; an appreciation of the range of activities, initiatives, and labels which are employed to create an environment in which knowledge is effectively created, shared, and used to increase competitive advantage and customer satisfaction.

Knowledge management roles are quite diverse and include such categories as senior and middle management roles—Chief Knowledge Officer (CKO) who ensures that KM goals are in line with organizational strategies and objectives; Chief Learning Officer (CLO) who ensures that the organization acts like a learning organization, improving over time with the help of accumulated best practices and lessons learned; Knowledge Managers who are typically responsible for the acquisition and management of internal and external knowledge. Other roles include

- Knowledge navigators, responsible for knowing where knowledge can be located, also called knowledge brokers.
- Knowledge synthesizers, responsible for facilitating the recording of significant knowledge to organizational memory, also called Knowledge Stewards.
- Content editors, responsible for codifying and structuring content, also called content managers; roles involving capturing and documenting knowledge—researchers, writers, editors.
- Web developers, electronic publishers, intranet managers, content managers.
- Learning-oriented roles such as trainers, facilitators, mentors, coaches—including those with responsibility for developing information and knowledge skills.
- Human resources roles with specific responsibility for developing programs and processes that encourage knowledge-oriented cultures and behaviors.
- Knowledge publishers, responsible for internal publishing functions, usually on an intranet, also called Webmasters, knowledge architects, Knowledge editors.
- Coaches and mentors, responsible for assisting individuals throughout the business unit or practice to develop and learn KM activities and disciplines.
- Help desk activities, including the delivery of KM and information related to training, also called Knowledge Support Office (KSO).

Knowledge management professionals require a multidisciplinary skill-set that consist of such competencies as finding, appraising, and using knowledge, beginning to be able to reformulate questions, navigating through content, evaluating the relevance of content, filtering out what is not needed, and synthesizing from diverse sources in order to apply the knowledge (e.g., to make a decision). Last but not least, they must contribute to the recording of such valuable experiences to organizational memory systems.^[41]

CONCLUSION

Knowledge management has assumed a greater priority due to the demographic pressures created by retiring baby boomers. Retirement en masse means a loss of valuable knowledge and know-how that has been accumulated by experience over the years. All organizations, regardless of type of industry, size, country, and so forth will have to face the issue of knowledge continuity. Knowledge continuity is analogous to business continuity: while the latter targets data loss prevention due to disasters, the former targets knowledge loss due to turnover in personnel. Valuable knowledge must not only be preserved and transferred to future workers, but also transferred across geographical and temporal boundaries due to the fact that work today is spread out over networks that may extend beyond organizational boundaries. Knowledge management is the discipline and field of practice that will equip us with the necessary tools, processes, and skill-sets to ensure that knowledge reaches all knowledge workers so that they can carry out their work more efficiently and more effectively.

REFERENCES

1. Merriam-Webster Dictionary Online. <http://www.m-w.com/dictionary/knowledge> (accessed November 20, 2008).
2. Leonard, D.; Sensiper, S. The role of tacit knowledge in group innovation. *Calif. Manage. Rev.* **1998**, *40* (3), 112–132.
3. McDermott, R. Why information technology inspired but cannot deliver knowledge management. *Calif. Manage. Rev.* **1999**, *41* (4), 103–117.
4. Davenport, T.; Prusak, L. *Working Knowledge*; Harvard Business School Press: Boston, 1998.
5. Barclay, R.; Murray, P. What is knowledge management? Knowledge Praxis. <http://www.mediaccess.com/whatis.html>, 1997 (accessed November 20, 2008).
6. Grey, D. What is knowledge management? *Knowl. Manage. Forum* March **1996**.
7. Sarvary, M. Knowledge management and competition in the consulting industry. *Calif. Manage. Rev.* **1999**, *41* (2), 95–107.
8. Wiig, K. *Knowledge Management Foundations*; Schema Press: Arlington, TX, 1993.
9. Nonaka, I.; Takeuchi, H. *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*; Oxford University Press: New York, 1995.
10. Pasternack, B.; Viscio, A. *The Centerless Corporation*; Simon and Schuster: New York, 1998.
11. Pfeffer, J.; Sutton, R. *The Knowing-Doing Gap: How Smart Companies Turn Knowledge into Action*; Harvard Business School Press: New York, 1999.
12. Ruggles, R.; Holtshouse, D. *The Knowledge Advantage*; Capstone Publishers: Dover, NH, 1999.
13. APQC (The American Productivity and Quality Centre). What is KM? From the APQC knowledge base: http://www.apqc.org/portal/apqc/ksn?paf_gear_id=contentgear-home&paf_dm=full&pageselect=detail&doid=101261 (accessed November 20, 2008).
14. Hibbard, J. Knowledge tools debate. *Information Week* March 16, 1998. <http://www.informationweek.com/673/73iudeb.htm> (accessed November 20, 2008).
15. Frappaolo, C. Knowledge management: A 2001 perspective. http://www.kwork.org/Stars/KM_perspective.pdf (accessed November 20, 2008).
16. Levinson, M. The ABCs of knowledge management. *CIO Magazine* http://www.cio.com/article/40343/ABC_An_Introduction_to_KM/40343/40343/1#1 (accessed November 20, 2008).
17. Bontis, N.; Nikitopoulos, D. Thought leadership on intellectual capital. *J. Intel. Capital* **2001**, *2* (3), 183–191.

18. Klein, D. *The Strategic Management of Intellectual Capital*; Butterworth-Heinemann: Oxford, U.K., 1998.
19. Stewart, T. *Intellectual Capital*; Doubleday: New York, 1997.
20. Stewart, T.A. Software preserves knowledge, people pass it on. *Fortune* **2000**, 142 (5), 390–392.
21. Rothberg, H.; Erickson, G. *From Knowledge to Intelligence. Creating Competitive Advantage in the New Economy*; Butterworth-Heinemann: Boston, 2005.
22. Dalkir, K. *Knowledge Management Theory and Practice*; Butterworth-Heinemann: Boston, 2005.
23. Kransdorff, A. *Corporate Amnesia: Keeping Know-How in the Company*; Butterworth-Heinemann: Oxford, U.K., 1998.
24. Polanyi, M. *The Tacit Dimension*; Peter Smith: Gloucester, MA, 1966.
25. Glesinger, J. The not-so-hidden cost of lost knowledge. *Energy Press J.* September 2, **2008**. <http://energy.pressandjournal.co.uk/Article.aspx/800114/?UserKey=> (accessed November 20, 2008).
26. Hutchison, L. Phased retirement saves intellectual capital. *St. Louis Commerce Magazine* **2001**. <http://www.stlcomm ercemagazine> (accessed November 20, 2008).
27. Liu, S.; Fidel, R. Managing aging workforce: Filling the gap between what we know and what is in the system. In *Proceedings of the 1st International Conference on Theory and Practice of Electronic Governance*, Macao, China, December 10–13, 2007; (ICEGOV '07) ACM: New York, 2007; Vol. 232, 121–128. <http://doi.acm.org/10.1145/1328057.1328084>.
28. Liebowitz, J. Bridging the knowledge and skills gap: tapping federal retirees. *Public Personnel Manage.* **2004**, 33 (4), 421–447.
29. LaBarre, P. People go, knowledge stays. *Fast Company* **2001**, 17 (4). <http://www.fastcompany.com/magazine/17/wyp17.html> (accessed November 20, 2008).
30. Wenger, E.; Snyder, W. Communities of practice: The organizational frontier. *Harvard Bus. Rev.* January–February **2000**, 139–145.
31. Wenger, E.; McDermott, R.; Snyder, W. *Cultivating Communities of Practice: A Guide to Managing Knowledge*; Harvard Business School: Cambridge, MA, 2002.
32. Szulanski, G. *Sticky Knowledge—Barriers to Knowing in a Firm*; Sage: Thousand Oaks, CA, 2003.
33. Alvesson, M.; Kärreman, D. Odd couple: Making sense of the curious concept of knowledge management. *J. Manage. Stud.* **2001**, 38 (7), 995–1018.
34. Drucker, P. *Management: Tasks, Responsibilities, Practices*; Harper & Row: New York, 1973.
35. Senge, P. *The Fifth Discipline: The Art and Practice of the Learning Organization*; Doubleday: New York, 1990.
36. Argyris, M.; Schön, D. *Theory in Practice. Increasing Professional Effectiveness*; Jossey-Bass: San Francisco, 1974.
37. Argyris, C. *Knowledge for Action. A Guide to Overcoming Barriers to Organizational Change*; Jossey-Bass: San Francisco, 1993.
38. Al-Hawamdeh, S. *Knowledge Management: Cultivating Knowledge Professionals*; Chandos Publishing: Rollinsford, NH, 2003.
39. TFPL website: <http://www.tfpl.com>.
40. TFPL Skills for knowledge management, a briefing paper by TFPL Ltd: Based on research undertaken by TFPL on behalf of the Library and Information Commission; TFPL Ltd.: London, 1999.
41. Henzcel, S. Supporting the KM environment: The roles, responsibilities, and rights of information professionals; Constructive thoughts on knowledge management. *Information Outlook* **2004**, (1), 14–19.

This page intentionally left blank

9 Information Management

Brian Detlor

CONTENTS

Introduction.....	125
A Process View of Information Management.....	126
The Organizational Perspective	127
Information as a Resource.....	127
Managing Information Processes.....	128
Information Technology's Role.....	129
Organizational Information Processing.....	129
The Library Perspective	130
The Personal Perspective	131
Conclusion	132
References.....	132

INTRODUCTION

Information management (IM) is a broad conceptual term that has various meanings and interpretations among different constituencies. Often the term is used interchangeably with others. For instance, information management is often equated with the management of information resources, the management of information technology (IT), or the management of information policies or standards.^[1] Some suggest that information management draws upon ideas from both librarianship and information science.^[2] The purpose of this entry is to clarify the meaning of the term “information management,” to showcase its major perspectives, and to illustrate how it relates to associated terms.

What is information management (IM)? Information management is the management of the processes and systems that create, acquire, organize, store, distribute, and use information. The goal of information management is to help people and organizations access, process, and use information efficiently and effectively. Doing so helps organizations operate more competitively and strategically, and helps people better accomplish their tasks and become better informed.

This entry adopts a process orientation toward information management where IM is viewed as the control over the information life cycle. Various models of information processes exist; some of the major processes involved in information life cycle concern those that create, acquire, organize, store, distribute, and use information.

Importantly, this entry recognizes three major perspectives of information management: the organizational, library, and personal perspectives. By far, the organizational perspective is the most predominant and popular. In a nutshell, this perspective deals with the management of all information processes involved in the information life cycle with the goal of helping an organization reach its competitive and strategic objectives. A variety of information resources are managed by organizations. These include transactional information stored in databases, summarized information found in data warehouses, and unstructured information content found in documents and reports. From the organizational perspective, the management of IT is a major component of any IM plan. It is with this perspective where associated terms like information systems management,

IT management, data management, business intelligence, competitive intelligence, content management, and records management have relevance.

The library perspective of information management recognizes the unique role of information provision organizations, such as libraries, whose central mandate is to provide their clientele with access to information resources and services. As such, these types of organizations view information management primarily as the management of information collections, such as books and journals. With respect to information collections, libraries are neither the creators nor the users of this information. As such, this perspective deals with the management of a subset of information processes involved in the information life cycle with the goal of helping library patrons access and borrow information items held in the collection. It is with the library perspective where associated terms like knowledge organization, classification, cataloging, digital libraries, indexes, and information retrieval systems have relevance.

The personal perspective of information management, though not strong, is similar to the organizational perspective in that it involves the management of all information processes in the information life cycle. The major difference, however, is that the organizational perspective concerns the management of information of interest to the success and well-being of an organization, while the personal perspective deals with the management of information of relevance and concern to the individual.

In terms of this entry's organization, the remainder is structured to further explain the major points just described. First, description is given on the information life cycle and the process view of information management. This is followed by a detailed discussion on each of the three major perspectives of information management: the organizational, library, and personal perspectives. Finally, concluding remarks are made.

A PROCESS VIEW OF INFORMATION MANAGEMENT

Leading information management scholars and organizations purport a process-driven view of IM. For example, Wilson defines information management as "the application of management principles to the acquisition, organization, control, dissemination and use of information relevant to the effective operation of organizations of all kinds."^[3] Similarly, Choo defines IM as the management of processes that acquire, create, organize, distribute, and use information.^[1] Likewise, the U.S. Government Accountability Office defines information management as "the planning, budgeting, manipulating, and controlling of information throughout its life cycle."^[4]

The conceptualization of information management from a process perspective emerged in the early 1990s.^[5,6] Advocates of this approach stress that a process model of information management should encompass all or some parts of the information value chain or life cycle. Choo proposes six discrete information-related processes or activities that need to be managed: 1) identification of information needs; 2) acquisition of information to address those needs; 3) organization and storage of information; 4) design and development of information products; 5) distribution of information; and 6) information use.^[1] Wilson positions the creation and use of information outside the information management rubric and purports the following six information-related processes pertaining to IM: 1) acquisition; 2) organization; 3) storage; 4) retrieval; 5) access/lending; and 6) dissemination.^[7] Not all information management frameworks include information needs identification and information use as processes to be managed. However, many see the most critical issue facing information managers is eliciting information requirements and matching those information needs in the design and delivery of information systems to promote effective and efficient information use.^[1,8,9]

With respect to this entry, the following are considered to be the predominant information processes to be managed in IM: information creation, acquisition, organization, storage, distribution, and use. Information creation is the process where individuals and organizations generate and produce new information artifacts and items. Information acquisition is the process where information items are obtained from external sources. Information organization is the process of indexing or

classifying information in ways that support easy retrieval at later points in time. Information storage is the process of physically housing information content in structures such as databases or file systems. Information distribution is the process of disseminating, transporting, or sharing information. Information use is the process where individuals and organizations utilize and apply information made available to them. Effectively managing these information processes helps get the right information to the right people in the right forms at the right times and at reasonable costs.^[1,10]

THE ORGANIZATIONAL PERSPECTIVE

The organizational perspective is the most predominant perspective of information management. Under the organizational perspective, IM concerns the management and control over the full life cycle of information processes ranging from creation to use for the betterment of the organization itself. In this sense, the management of information processes is seen as a strategic advantage that affords four kinds of benefits to an organization: 1) a reduction of costs; 2) a reduction of uncertainty or risks; 3) added value to existing products and services; and 4) the creation of new value through the introduction of new information-based products and services.^[11]

INFORMATION AS A RESOURCE

Fundamental to the organizational perspective of information management is the view and treatment of information as a strategic resource—one that needs to be managed like any other critical organizational resource, such as people, equipment, and capital. Many organizations recognize the potential value of information and the need to be aware of what information resources exist in an organization, and the costs associated with acquiring, storing, processing, and using that information. Forward-looking companies consider information as a strategic asset that can be leveraged into a competitive advantage in the markets served by the company.^[8]

The concept of managing information as an organizational resource emerged in the late 1970s with the creation of the US Paperwork Reduction Act that was established to resolve the huge costs of managing and handling information by bidders for government contracts.^[12] It was at this time, that the term information resources management (IRM) became popular and was used to signify the management of both information and information technologies.^[13,14] In reality though, the IRM concept was more about the management of data than it was about the management of other types of information, such as documents or reports. Case in point, an empirical assessment of the IRM construct defines IRM as “a comprehensive approach to planning, organizing, budgeting, directing, monitoring and controlling the people, funding, technologies and activities associated with acquiring, storing, processing and distributing data to meet a business need for the benefit of the enterprise”—not other types of information.^[15]

This entry suggests that within the organizational perspective of information management, IM is more than just the management of data (e.g., raw facts stored in transactional databases). Rather, information management in organizations involves the management of a varied set of information resources, ranging from data to information.^[16]

Data can be considered as “raw facts” that reflect the characteristics of an event or entity. Examples of data items held in an organization would be a customer name, an order quantity, or a unit price. The management of data in organizations traditionally belonged to the “Data Management” function or “Data Resource Management” department. Data management deals with the storage and processing of low-level data items found in transaction processing systems. The focus of data management is to maximize the speed, accuracy, and efficiency of processing and maintaining transactions or records.^[1] Data management is concerned with the creation and “acquisition of data, its storage in databases, its manipulation or processing to produce new (value-added) data and reports via application programs, and the transmission (communication) of the data or resulting reports.”^[8]

Information can be viewed as “meaningful data” where data have been converted into a meaningful and useful context, such as in the identification of a best-selling or worst-selling product item from historical sales data for a company. Examples of information items managed in a company would be summarized information pulled from transactional database systems and stored in data warehouses and/or data marts. Such “business intelligent” systems support decision-making, and allow for the slicing-and-dicing of summarized transactional information to find patterns and trends in operational data (e.g., sales data, supplier data, customer profile data) of importance and relevance to the organization. For example, the mining of summarized transactional data stored in a data warehouse could be used to justify the promotion of a product or service to an untapped market segment, or to measure the effectiveness of a marketing campaign in a certain geographical area.^[16]

Though more emphasis in organizations is placed on the management of data and information that is structured and formalized (i.e., data neatly contained within a transactional database or summarized information stored in a data warehouse), the amount of unstructured information that is created, acquired, organized, stored, distributed, and used within an organization usually far exceeds the amount of structured data and information an organization manages.

Unstructured information is the type of information that can be found in reports, documents, e-mail messages, and PowerPoint presentations, among others. This includes reports and documentation generated internally with a company and outside the enterprise as well. For example, managing information concerning industry trends, legislative regulations, competitor happenings, news bulletins, etc. is key to helping an organization stay abreast and react to competitor threats and environmental concerns. This is where information management intersects with a company’s “competitive intelligence” initiatives and “environmental scanning” activities. Competitive intelligence refers to the analysis of information about competitors and competitive conditions in particular industries or regions.^[17] Environmental scanning refers to the analysis of information about every sector of the external environment that can help an organization plan for its future.^[18,19]

Often document management systems, also known as content management systems, help manage unstructured information that is created, acquired, organized, stored, distributed, and used within an organization. These systems support the electronic capture, storage, distribution, archiving, and accessing of documents. These systems typically provide a document repository where documents can be housed, indexed, and accessed. In these repositories, metadata information is maintained about each document that tracks each document’s editorial history and relationships to other documents. Indexing of documents within these systems, through both manual keyword classifications and automatic indexing techniques, facilitates easy document retrieval by organizational workers at later points in time.^[16]

Unstructured information also includes records of transactional information processing such as invoices, contracts, order requisitions, bills of lading, and financial statements. The actual data may be resident in a transaction database and managed elsewhere, but the entire record of the event, often with signatures, are handled as a complete entity that must be generated, stored, processed, retrieved, and archived. The management of such information is referred to as “records management” and involves the application of systematic and scientific control to all recorded information that an organization needs to do business.^[20]

MANAGING INFORMATION PROCESSES

A good information management program in an organization will manage the full life cycle of information ranging from creation to use. For example:

- When generating transactional data, steps will be taken to ensure that the data will be stored following database “normalization rules” to promote data integrity, the single sourcing of data, the reduction of wasted database space, and fast transaction processing.

- When acquiring information, such as the purchase of market research data or competitor intelligence information, steps will be taken to reduce duplicate purchases and to increase the accessibility of any purchased data and information across the enterprise.
- Any data or information that is stored will be adequately protected against unauthorized access, as security, privacy, and copyright concerns exist.
- Data and information stewardship programs will be set up to identify those organizational workers or units who are responsible for the quality and management of certain data and information items.
- Data and information will be regularly backed up for recovery purposes.
- Duplicate or mirror copies of data and information items will be created to facilitate access, and reduce network congestion and/or an overload of requests on the servers on which the data and information reside.
- Old data and outdated information will be archived and/or deleted.^[16]

INFORMATION TECHNOLOGY'S ROLE

Information technology plays a critical role in the management of information in organizations. However, having said that, it should be understood that IT is the technical medium upon which information is housed, accessed, retrieved, distributed, and used, and not the primary entity that is being managed under the information management rubric. Information processes are.

Nevertheless, much confusion exists over the role IT plays in the management of information in organizations, with some equating information management primarily to the management of IT itself. For example, the mission of the Society for Information Management (SIM), a professional society for chief information officers and senior IT leaders based in the United States, is “to provide international leadership and education in the successful management and use of IT to achieve business objectives.”^[21] Likewise, Davis and Hamilton define information management as “the management of information systems and information resources by an individual, a group, or an organization” and describe IM as a “new business function with responsibility to define organizational informational requirements, plan, and build an information infrastructure and information systems applications, operate the system, and organize, staff, and manage these activities.”^[22]

This entry suggests that such descriptions of the information management function in organizations are inappropriate and are better served by alternate terms such as information systems management or IT management. Information systems management refers to the control over the development, design, rollout, and support of information systems applications that support business processes and workflow. Information technology management refers to the management and control over IT (e.g., hardware, software).

A variety of factors need to be considered when managing technology in organizations. The appropriate hardware and software needs to be installed and implemented. Repair and maintenance of hardware and software systems must be done. Licensing concerns must be taken into consideration. Information systems professionals (e.g., programmers, network specialists, database administrators, systems analysts) need to be hired and trained.^[16]

Information technology is very important to how information is managed in organizations, and the management of IT should be an integral part of any IM plan.^[11] After all, IT facilitates the ability to support information creation, acquisition, organization, storage, distribution, and use in both efficient and effective manners.

ORGANIZATIONAL INFORMATION PROCESSING

Information management is concerned with organizational information processing. Two broad orientations exist in the organizational information processing literature that have implications on information management.^[23]

The first, mainly developed at Carnegie Mellon University and represented by Herbert Simon, James March, and Richard Cyert, sees an organization's ability to process information as the core of managerial and organizational competencies and organizations as bounded, rational systems.^[24] Here, strategies to improve information processing capabilities concern increasing an organization's information processing capacity and reducing an organization's need for information processing.^[25]

The second, represented by Karl Weick and Richard Daft, sees organizations as loosely coupled systems where individual participants collectively develop shared interpretations of the environment and then act on the basis of this interpretation. In this sense, information processing is about reducing the equivocality or ambiguousness of information about the organization's external environment.^[26]

Taking these two orientations of organizational information processing together, information management then is about increasing an organization's information processing capacity and reducing both information processing needs and information equivocality.

THE LIBRARY PERSPECTIVE

Outside of the organizational context, information management also has meaning in the library world and to other information provision organizations whose central mandate is to provide clientele with access to information resources and services. Information management is of concern to all kinds of libraries, such as public libraries and academic and research libraries. This includes corporate libraries that manage serial subscriptions and electronic access to industry reports and electronic databases for workers and employees in their organizations.

From a library perspective, information management concerns the management of information collections, such as books and journals. The goal of information management from a library perspective is to help library patrons access and borrow information items held in the collection. A variety of activities surround the management of a library collection, including the development of collection policies and materials budgets, the selection of collection items, the analysis of collection usage and end user collection needs, training of collection staff, preservation of collection items, and the development of cooperative collections with other institutions.^[27]

Since libraries are neither the creators nor the users of information, this perspective deals with the management of a subset of information processes involved in the information life cycle. For example, Wilson identifies six information-related processes pertaining to IM that are applicable to the library perspective: 1) acquisition; 2) organization; 3) storage; 4) retrieval; 5) access/lending; and 6) dissemination.^[7]

Information acquisition involves the process of buying or securing information from sources external to the library for the collection. Care must be taken to ensure that the correct information items are acquired (those that match the information needs of end users) and at reasonable costs.

Information organization pertains to the process of indexing or classifying information housed in the collection to support easy retrieval at later points in time. Today, this process is sometimes referred to as "knowledge organization" in the library world. For example, the International Society for Knowledge Organization (ISKO) utilizes the term "knowledge organization" to represent the process of how documents, document representations (such as bibliographic records), and concepts (keywords, constructs) are ordered and classified.^[28] Likewise, Anderson defines knowledge organization as the "the description of documents, their contents, features and purposes, and the organization of these descriptions so as to make these documents and their parts accessible to persons seeking them or the messages that they contain" and describes how knowledge organization encompasses "every type and method of indexing, abstracting, cataloging, classification, records management, bibliography and the creation of textual or bibliographic databases for information retrieval."^[29] In this sense, knowledge organization deals with both processes and systems.^[30] With respect to information management, this entry suggests that the use of the "knowledge organization" label is misleading. In actuality, information items (such as documents and document representations)

are being organized or ordered (i.e., indexed, classified, cataloged) rather than knowledge items (concepts, constructs) per se. As such, “information organization” is suggested to be a better label to use when talking about the indexing and classification of items held in a collection, at least from a library perspective of IM.

Information storage refers to physically housing items in the collection. This encompasses the housing of both paper and electronic documents (and their document representations). This could be accomplished by storing books and journals in physical stacks in the library or storing full-text electronic versions of documents in an electronic or digital library.

Information retrieval involves the process of searching and finding information in the collection. Typically end users will conduct a search query using electronic tools that are Web-based to find items of interest in the collection. In the information retrieval process, end users themselves, or information search intermediaries (librarians) working on the behalf of end users, “pull” the information that is needed from the collection.

Information access/lending involves the process of providing physical or electronic access to the collection and the ability to check out information items of interest. For paper-based documents, this may involve signing out and borrowing a physical information item. For information in electronic form, this may involve viewing the information item online and/or downloading a copy of the information item from the collection.

Information dissemination is the process of circulating physical information items of interest from the collection to end users. Traditionally, methods of dissemination involve the distribution of photocopied journal content pages, or the circulation of documents. Today, this has largely been replaced with electronic alert services that “push” information items in the collection of interest to specific users based on stored end user profiles.

The management of library information collections is a complex and ever-evolving process.^[31] The extensive digitization of information resources has placed new pressures on libraries to respond to securing the requisite skills, resources, and competencies to successfully manage digital library collections.^[32] In order to manage the transition and reliance on digital information collections, it is important more than ever for libraries to consider the expectations and needs of end users, as well as limitations in library staff and budget adaptability.^[31] These are critical factors in rolling out any successful information management program from a library perspective.

THE PERSONAL PERSPECTIVE

Information management from the personal perspective refers to how individuals create, acquire, organize, store, distribute, and use information for personal purposes. This can concern the management of information for everyday use (e.g., personal calendars, schedules, diaries, news items) or work-related reasons (e.g., work schedules, things-to-do, project files). As such, personal information management involves the handling and processing of information over the entire information life cycle, just as in the case of the organizational perspective. However, the personal information management perspective differs from the organizational perspective of IM in that personal information management concerns items of interest to the individual, not the organization.

Central to the personal perspective of IM is the need to manage the information processes of information creation, acquisition, organization, distribution, and use so that the right information is accessible and available in the right place, in the right form, and of sufficient completeness and quality to meet personal information needs. A variety of information items are created, acquired, organized, stored, distributed, and used by people for personal purposes. These include personal notes, personal journals, Web pages, e-mail messages, news articles, address books, calendar dates, reminders, fax communications, etc.

Technologies and tools, such as electronic personal information managers like PDAs, help people carry out these processes more efficiently and effectively.^[33] Cheap and fast search and storage technologies bring stability and order to people's often chaotic and messy personal information

environments, and facilitate making the most of people's personal information collections.^[34] Despite the benefits of these technological tools, there is some evidence to suggest that technological advances are less important in determining how individuals organize and use information than are the tasks that people perform.^[35]

It is interesting to note that the phrase "personal information management" was first used in the 1980s with the advent of personal computers and their capacity to process and manage information.^[36] However, the concept of "personal information management" was most likely first implied back in 1945 by Vannevar Bush when he envisioned using the "Memex" machine for personal information management—a machine that allowed an individual to store all his or her books, records, and communications and consult those personal information sources with great speed and flexibility.^[37]

As the proliferation of personal computing devices and the explosion of the amount of information that is created, generated, and used continues to increase, the relevance of information management from the personal perspective will likely gain more attention and importance from both researchers and practitioners alike.

CONCLUSION

This entry provided a review and discussion of the information management concept. Importantly, a process-oriented definition of IM was proposed and three distinct perspectives of information management were described: the organizational, library, and personal perspectives. It was suggested that the organizational perspective was the most predominant, though the other two were shown to have relevance and meaning in today's world.

This entry emphasized the importance of information management. IM allows organizations to reach strategic goals and make sound decisions. IM enables libraries and other information provision organizations to manage information collections effectively and efficiently. IM helps people manage their own personal information collections.

This entry also described how IM is not so much about the management of technology, but rather more about the management of the processes of IM, specifically the creation, acquisition, organization, storage, distribution, and usage of information.

According to the current scope of the *International Journal of Information Management*, the challenge for information management now is less about managing activities that collect, store, and disseminate information, and more about placing greater focus on managing activities that make changes in patterns of behavior of customers, people, and organizations, and information that leads to changes in the way people use information.^[38] This may not be such an easy task. People and their information behaviors are hard to change. So are organizational cultures.^[11] IM is less about solving technical problems and more about addressing the human side of information management. Humans add the context, meaning and value to information, and it is humans who benefit and use this information. As such, "mastering information management is an essentially human task."^[39]

REFERENCES

1. Choo, C.W. *Information Management for the Intelligent Organization: The Art of Environmental Scanning*, 3rd Ed.; Learned Information: Medford, NJ, 2002.
2. Macevičiūtė, E.; Wilson, T.D. The development of the information management research area. *Inform. Res.* **2002**, 7 (3). <http://InformationR.net/ir/7-3/paper133.html> (accessed July 2008).
3. Wilson, T.D. Information management. In *International Encyclopedia of Information and Library Science*; Routledge: London, 2003; 263–278.
4. <http://www.GAO.gov> (accessed July 2008).
5. Davenport, T.H. *Process Innovation: Reengineering Work Through Information Technology*; Harvard Business School Press: Boston, 1993.
6. McGee, J.V.; Prusak, L. *Managing Information Strategically*; John Wiley & Sons: Toronto, ON, 1993.

7. Wilson, T.D. 'The nonsense of knowledge management' revisited. In *Introducing Information Management*, Macevičiūtė, E., Wilson, T.D., Eds.; Facet Publishing: London, 2005; 151–164.
8. Karim, N.S.A.; Hussein, R. Managers' perception of information management and the role of information and knowledge managers: The Malaysian perspectives. *Int. J. Inform. Manage.* **2008**, *28* (2), 114–127.
9. Detlor, B. *Towards Knowledge Portals: From Human Issues to Intelligent Agents*; Kluwer Academic Publishers: Dordrecht, the Netherlands, 2004.
10. Robertson, J. Ten principles of effective information management. *KM Column* **2005**, November. http://www.steptwo.com.au/papers/kmc_effectiveim/index.html (accessed July 2008).
11. <http://choo.fis.utoronto.ca/Imfaq/> (accessed July 2008).
12. Commission on Federal Paperwork, *Information Resources Management*; US Government Printing Office: Washington, DC, 1977.
13. Horton, F.W. *Information Resources Management: Concept and Cases*; Association for Systems Management: Cleveland, OH, 1979.
14. Horton, F.W. *The Information Management Workbook: IRM Made Simple*; Information Management Press: Washington, DC, 1982.
15. Lewis, B.R.; Snyder, C.A.; Rainer, R.K., Jr. An empirical assessment of the information resource management construct. *J. Manage. Inform. Syst.* **1995**, *12* (1), 199–223.
16. Baltzan, P.; Phillips, A.; Detlor, B. *Business-Driven Information Systems*, 1st Canadian Ed.; McGraw-Hill Ryerson: Whitby, ON, 2008.
17. Sutton, H. *Competitive Intelligence*; The Conference Board Inc.: New York, 1988.
18. Aguilar, F.J. *Scanning the Business Environment*; MacMillan Co.: New York, 1967.
19. Choo, C.W.; Auster, E. Scanning the business environment: Acquisition and use of information by managers. In *Annual Review of Information Science and Technology*, Williams, M.E., Ed.; Learned Information: Medford, NJ, 1993.
20. Robek, M.F.; Brown, G.F.; Wilmer, O.M. *Information and Records Management*, 3rd Ed.; Glencoe Publishing: Encino, CA, 1987.
21. <http://www.simnet.org> (accessed July 2008).
22. Davis, G.B.; Hamilton, S. *Managing Information: How Information Systems Impact Organizational Strategy*; Business One Irwin: Homewood, IL, 1993.
23. Choo, C.W. Towards an informational model of organizations. *Can. J. Inform. Sci.* **1991**, *16* (3), 32–62.
24. Simon, H. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*, 3rd Ed.; The Free Press: New York, 1976.
25. Galbraith, J.R. *Organization Design*; Addison-Wesley: Reading, MA, 1977.
26. Weick, K.E.; Daft, R.L. The effectiveness of interpretation systems. In *Organizational Effectiveness: A Comparison of Multiple Models*, Cameron, K.S., Whetten, D.A., Eds.; Academic Press: New York, 1983; 71–93.
27. Branin, J.J., Ed. *Collection Management in the 1990s*; American Library Association: Chicago, 1993.
28. <http://www.isko.org/ko.html> (accessed July 2008).
29. Anderson, J.D. Organization of knowledge. In *International Encyclopedia of Information and Library Science*, Feather, J., Sturges, P., Eds.; Routledge: London, 1996; 336–353.
30. http://www.db.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization.htm (accessed July 2008).
31. Branin, J.; Groen, F.; Thorin, S. The changing nature of collection management in research libraries. *Libr. Resour. Tech. Serv.* **2000**, *44* (1), 23–32.
32. Sreenivasulu, V. The role of a digital librarian in the management of digital information systems (DIS). *Electron. Libr.* **2000**, *18* (1), 12–20.
33. Jones, W. *Keeping Found Things Found: The Study and Practice of Personal Information Management*; Morgan Kaufmann Publishers: Burlington, MA, 2008.
34. Teevan, J.; Jones, W.; Bederson, B.B. Personal information management. *Commun. ACM* **2006**, *49* (1), 40–43.
35. Barreau, D. The persistence of behavior and form in the organization of personal information. *J. Am. Soc. Inform. Sci. Technol.* **2008**, *59* (2), 307–317.
36. Lansdale, M. The psychology of personal information management. *Appl. Ergon.* **1988**, *19* (1), 55–66.
37. Bush, V. As we may think. *The Atlantic Monthly* July **1945**, *176* (1), 101–108.
38. http://www.elsevier.com/wps/find/journaldescription.cws_home/30434/description#description (accessed July 2008).
39. Davenport, T. Putting the I in IT. In *Mastering Information Management*; Pearson Education: London, 2000.

This page intentionally left blank

10 Digital Asset Management

James E.P. Currall and Michael S. Moss

CONTENTS

Introduction.....	135
Where to Start?.....	135
Managing Assets.....	136
A Suggested Approach.....	137
Challenges.....	138
Digital Objects	138
Characteristics	139
Collections.....	141
Management.....	142
Logistics	144
Conclusion	145
Acknowledgments.....	145
References.....	146

INTRODUCTION

There is a tendency to suggest that managing digital assets is radically different from the management of traditional stock in archives and libraries and therefore requires the establishment of separate units with specialist technical staff. While it is true that there are technical challenges and that digital assets behave differently from their analogue equivalents, stock selection and collection building remain constant. The specialist skills of curators of different types of assets continue to be needed. This is not just because they have different semantic and technical metadata, but because, just as in the analogue, they require specialist knowledge in their conservation and use. In any strategic approach to digital asset management the relationship of the content to organizational goals must be given priority over the technical, just as in any established archive or library environment. What at present sets the digital apart from the analogue is the scale of resources required for hardware, software, and technical support. When confronted by the digital and technical experts that call themselves information specialists, archivists and librarians need to remember that is what they are and that they have centuries of experience behind them.

WHERE TO START?

On the subject of preserving digital content there is a “bewildering array of essays, reports case studies, conference proceedings, workshops, and training venues with an equally confusing ranges of recommendations, proposals, predictions, and procedures” with major initiatives in many parts of the world. Some examples are Digital Preservation (Library of Congress) that is home to The National Digital Information Infrastructure and Preservation Program established in 2000,^[1] the InterPARES project based at the University of British Columbia,^[2] Preserving Access to Digital Information at the National Library of Australia,^[3] Ryukoku University’s Digital Archives Research

Center in Japan,^[4] Digital Preservation Europe supported by the European Commission,^[5] and the Digital Preservation Coalition in the United Kingdom established in 2001.^[6] In January 2008 OCLC announced the formation of a “Task Force on Sustainable Digital Preservation and Access” “to explore dimensions of the sustainability challenge that have never been addressed before, and make the societal and institutional cases for supporting data repositories.”^[7] Although we all create information for our individual needs, the information that mostly concerns these initiatives is produced in an institutional or organizational context. Much of the literature and most projects take as their starting point “the complexity of preservation and long-term retrieval” of digital content,^[8] rather than the processes involved in its creation and the business needs that might drive such investment, the traditional point of departure of the information professions in the analogue. This has consequences for both the analysis of the problem and any recommendations that may result.

MANAGING ASSETS

Much of the established literature concerns itself with preserving objects, where the objects are encoded in file formats that have technical properties and require processes of “life support” to remain accessible and therefore (presumably) useful. More recently there is a growing body of literature and initiatives around aggregations of objects in what are termed “repositories” and the processes that should be in place for a repository to be regarded as “trusted”^[9] [*sic*] or perhaps better described as “trustworthy.” Such discussion, which is centered on the question “How” is of course necessary, but it is a long way downstream of the really important issues which all begin, as they always have done, with “W”—What?, Why?, Who?, Where?, When? Central to any management task in any setting is the purpose for which it is taking place. This is often lost in the excitement of processes and technicalities. The Research Roadmap Working Group of Digital Preservation Europe recently examined the challenges of preservation from five vantage points: digital object level, collection level, repository level, process level, and organizational environment.^[10] Only the last of these is concerned with creation and use and none are really centered on information content or purpose.

An instance of what happens if a narrow technical process view of management is taken is the treatment of risk in relation to digital material. It is invariably viewed as a set of threats to the integrity, renderability, or understandability of files, in other words downstream risks that require formalized process to mitigate. For example such risks involve server failure, fire, and flood—threats not opportunities. As Ross avers

where preservation is concerned, the “risk amplifiers” have taken the high ground from “risk attenuators,” as is evident from the growth in the number of publications, conferences, and conference presentations during the past ten years that stress how essential it is that we overcome the obstacles to the longevity of digital objects.^[11]

As Harvey points out this rhetoric “is lamentably short on supporting specifics and evidence.”^[12] Such an approach both exaggerates the curatorial imperative and gives the impression that all digital objects, whatever their provenance and purpose, have more in common than the fact that they are represented as bit patterns (1s and 0s). This is fundamentally different from experience of the analogue paper world where no one would claim that all printed matter or manuscript material shared a common ontological status or an imperative to preserve. This confusion has been compounded by the development of mass storage devices that appear to offer the opportunity to keep “everything” irrespective of its current or likely future value and a pervasive postmodernism that problematizes the dominant discourse thereby suggesting that every information object has equivalent value.^[13] As Greene puts it “who has the objectivity necessary to identify with certainty what is a turd and what is a ‘critically important record?’ ”^[14]

If we think of all information in the digital world as “objects,” with no thought as to what information they contain or of the fact that they have more or less, rather than equal value then we are not so much managing assets as commodities. The purpose of this entry is to explore the issues that are important in Asset Management, with an eye to the fact that the assets are Digital, rather than the issues surrounding the Management of the Digital with the fact that they are Assets being incidental. It would be rather like managing a Wal-Mart store and forgetting that the stock all has different shelf lives and patterns of demand.

A SUGGESTED APPROACH

Information has always been generated by processes that involve technology whether it be a stylus to write on slivers of bark on the frontier of the Roman empire in the first century AD^[15] or a PC to generate a bit pattern in the twenty-first century. The relationship between process, available technologies, and societal change has been much researched by those interested in the history of material culture, particularly the printed book, who warn against a techno-deterministic presentism.^[16–21] Heeding such advice and by investigating workflow, commentators are beginning to advance a different approach with the emphasis firmly on business needs and what are termed “artful processes [that] depend on the skills, experience, and judgment of the primary actors.”^[22] Those responsible for the creation and use of information objects, in turn, determine curation and preservation strategies and policies rather than these being imposed by technologists and records managers. Established practice in the analogue information world, accumulated over hundreds of years, is being restated in part by Web 2.0 technologies with Really Simple Syndication (RSS) and Atom feeds. This runs counter to the determination of information and Information and Communications Technology (ICT) professionals to formalize process and maintain control of information distribution, in ways that do not map well onto the needs and expectations of users. As long ago as 1997, the Effective Records Management (ERM) Project at the University of Glasgow in Scotland studied how information might best be managed in the digital world, with an emphasis on the information content and the needs and practices of those creating and using it. Using a demonstrator system, the ERM Project showed that such an approach to information handling effectively integrates technology and addresses the “risk amplifiers” in the creation and curation of digital information.^[22] Gilland-Swetland reached much the same conclusion in a study in 2000.^[23] If the goal is no longer to work towards rigidly integrated and formalized information assets but towards “looser and more dynamic couplings between service providers and consumers,” then the individual characteristics and purposes of information objects becomes easier to describe.^[24]

An approach to Digital Asset Management based on content and its value (What? and Why? questions) immediately raises the prospect of selection which clearly has to be based on some sort of rationale. This will vary from one setting to another; perhaps business need or collection policy and development or in the case of an individual acting in a private capacity; this may be purely sentiment. For all the talk of the emergence of a new preservation paradigm defined by the digital, this is the way it is in the analogue where information professionals have always had to work in a variety of institutional contexts and collaborate with “collectors.”^[25]

Managing assets has a cost associated with it and compared to paper information objects, digital information objects are less likely to survive being dumped on a shelf and forgotten about for years—in other words, benign neglect is unlikely to be a good strategy. This suggests that management of digital assets requires management of costs and inputs also. It is essential to grasp that preservation of any information object is an investment in order to realize future benefits from it. There is clearly an art to this in order to ensure that what is preserved does in fact bring benefits that justify the costs that have been incurred in preserving it, even if these do not become apparent for a very long time and are difficult to calculate.

CHALLENGES

One of the challenges for information professionals in the management of digital assets in relation to their nature as objects is that their very ontology appears to sweep away the curatorial boundaries that have separated libraries from archives, archives from museums, museums from galleries. A pattern of bits is relentlessly the same, whether it represents a new installation or an old master, a most secret report from the security services to an e-mail to a lover, a book, or this entry. The technical actions required by curatorial staff to preserve the pattern of bits are precisely the same, even if there is as yet no clear agreement about precisely what these should be. The reaction of the curatorial professions has been to attempt to translate their difference into the digital often by insisting on different standards of metadata or different approaches to appraisal or privileging that will be discussed later, rather than by focusing on the characteristics of the objects themselves and then concentrating on the actions required to preserve them. The other challenge is, paradoxically, almost the reverse and is based on the notion that precisely because all the objects are composed of 1s and 0s, that a single repository with a single metadata description framework is equally suitable for all types of object, whether they represent motion picture, sound, still image, engineering drawing, or character-based text, or whether they are records of banking transactions, minutes of meetings, art works, computer games, or state secrets. Such a technical approach ignores any underlying ontological characteristics that require different approaches to their curation and management. In the analogue a music score may be held in the same repository as a medieval manuscript, but the each require quite different curatorial approaches.

It is their status as information objects that determines their selection, description, and management as assets, whilst their status as digital objects determines how they are managed at a technical level. The former requires recognition of their difference, whilst the latter requires an acknowledgment of their sameness. We will distinguish between these two sorts of management as “Asset Management” and “Technical Management.”

DIGITAL OBJECTS

There are broadly two origins of digital objects that lead to different approaches to their value as assets and the way in which they are managed: those that are born digital and those that are digitized from an analogue original that already forms part of a collection. From a technical management perspective, this is often confused and for convenience they are treated as if they have the same status from the perspective of asset management. This is largely because archives, libraries, and museums have or are investing heavily in digitization, often with the encouragement of external agencies with other agendas, for example, the Japanese Government is committed to make publicly available all the documents relating to the Second World War in an attempt to seek closure with their neighbors for wartime atrocities.^[26]

The digitized object is a surrogate for an analogue original that has been appraised, privileged, or selected to form part of an institutional or private collection. Even if the digital surrogate permits exploration of features of the object not visible in the original, for example, in the *Beowulf* manuscript^[27] or the Durham *Liber Vitae*,^[28] it is not the original, and it may be that in the future with further technical development we will be able to see even more. Colin Webb, Director of Preservation Services at the National Library of Australia, warned “we shouldn’t assume we can automatically convert handwritten, or even printed text with a variety of type faces in poor condition, into highly accurate searchable text.”^[29] Some digitization projects are the equivalent of analogue microfilming projects that sought to conserve heavily used originals by making the contents available in an alternative media, census returns and newspapers are good examples. If the digitization process involves either manual indexing, as with census returns, or optical character recognition, or old newspapers, it will be much more expensive and problematic than microfilming. Others and possibly the majority are not. They are usually conceived as learning resources that contain selected objects from

collections that are situated in semi-structured environments to provide online users with a range of learning opportunities. The management of the assets in such projects can easily become haphazard, if care is not taken from the outset to ensure that preservation-quality images are made usually in Tagged Image File Format (TIFF) with appropriate technical metadata and, as importantly, they are properly referenced and cross-referenced to the analogue original.^[30]

An essential criterion for all digitization projects is good project planning that includes a sustainability strategy and above all does not abandon the editorial controls familiar in the analogue. Digital objects still need to be described and cataloged in exactly the same way as their analogue equivalents, particularly where their context is within a learning resource. It is often tempting to reduce the budget for such management processes, as inevitably it drives up the unit cost, and to assume wrongly in most cases that income generation will cover at least part of the outlay. This is only true for a very small fraction of projects, where the digitized objects themselves have to be paid for by the customer, as is usually the case with census returns. Much also is made of the potential of repurposing digital surrogates in other learning environments, but there is little evidence to suggest this happens except perhaps on a very limited scale, partly because of advances in digitizing and distribution technologies, and partly because digital objects used in different settings have different metadata requirements.

The born digital object is “original” in that there is no version available on other media to fall back on if it becomes damaged and, since vast quantities of these are produced as a product or by-product of daily work in organizations, they have usually not been subject to selection for a collection. In managing such assets it is important to recognize that the process of selection from the moment of creation can work in reverse to the analogue. The retention of an object in an e-mail folder, for example, is often automatic and requires deliberate intervention to remove it, albeit sometimes with warning barriers, whereas in the analogue the decision to file rather than throw in the wastepaper basket demands a deliberate decision by the creator or recipient or their agents. Born digital objects have existed since the mid-1950s and with the advent of the PC from the 1970s information migrated rapidly into the digital.^[31] A result of these developments was that processes familiar in the analogue world disappeared. For example, printed text was no longer “set” by a compositor, but simply converted into type from a word processed file and as a result often established editorial processes were abandoned. With the development of client–server architecture, from the 1980s and later Web-based architecture distribution of information objects was radically transformed and instead of copies being made, as for example in the printing of books, users viewed a single instantiation many times from a multiplicity of sites. This development finally demolished the intellectual distinction, which was already being eroded by technology, between archives that held unique objects and libraries that held multiple copies. There were now increasingly only unique objects, differentiated by the degree of intermediation—an online journal can be expected to have processes for quality review and refinement, whereas a personal Web site is unlikely to. This in a sense was the Gutenberg revolution the other way round, instead of copies of original being made available in print and widely distributed, the original itself could be viewed many times from many places.

CHARACTERISTICS

Digital information assets, whilst having a great deal in common with their analogue cousins, have a number of characteristics that set them apart, and present challenges for their management and in particular for their long-term survival.^[32] The most obvious is that individual objects are internally dynamic and inherently unstable in ways that their analogue equivalents could only be partially so.^[33] A book or journal article used to have a critical apparatus that referred to other stable analogue resources that could be discovered usually in libraries, archives, museums, and research institutes. Now, as in this entry, they include “hot links” to Web sites and blogs that we can expect to be dynamic and in some cases to disappear. This is the problem that Brewster Kahle

has sought to address in the Internet Archive.^[34] Although a useful tool, in itself it is full of such broken links.

Other types of dynamic “behaviors” are often embedded in digital documents such as macro code or graphics that depend on specific versions of rendering software to function correctly. For the library community it is now difficult, if not impossible, to distinguish between editions or versions and, as anyone who has used track changes knows, to observe with any clarity the gestation of the “original.” Wikipedia, for example, is edited continuously with little of the process of an analogue equivalent, although in this case there is at least a mechanism to cite a particular instantiation of an article. As problematic is the rendition of an “original” in a browser, word processing or other type of program that depends on the user settings, software version, screen size, printer configured for the system, and perhaps, the speed of an Internet connection. These may seem trivial problems to a technologist, but to those responsible for asset management, they present intellectual, technical, logical, and management challenges. In the analogue they are of vital importance, where we might describe this as a beautiful edition and that as a cheap edition or where we want to know precisely what an artist intended or the precise process of gestation of a literary work. This is why, although we may place a wide range of material in a digital repository, curatorial practice and user needs must continue to be differentiated.

It is often asked if all digital information assets share characteristics sufficient to make the idea of a single digital repository for a wide variety of assets a feasible proposition. In essence this question is about whether the management needs of music files are the same as committee papers, or whether the needs of scientific data files are the same as video. In-so-much as these are all files containing bit patterns of 1s and 0s, they might superficially appear to require the same treatment. Such an analysis fails to take account two very important characteristics of digital assets: the two differing types of metadata required adequately to describe them. Basic Dublin Core^[35] metadata superficially appears to provide a simple level of description that can be applied to a wide range of digital assets, but whilst it tackles basic description, it does not deal with technical attributes that are essential for rendition, calibration, and the physical context of the object. Consider the 40–50 attributes detailed in the JPEG images stored automatically by digital cameras or the similar number of attributes that can be stored in an MP3 file on a MP3 player. Dublin Core represents just the attributes that file types may have in common, but the considerable number of attributes (stored within the file itself or externally) that differentiate formats may require very different treatment. In this context, it should be remembered that the purpose of information assets revolves around their intellectual content rather than their technical attributes and different types of content need different curation and management regimes, irrespective of differences in media type, file format, and such like.

The file formats used for storing digital information have an intimate relationship with both their intellectual content and the software required for their rendition. The management needs of file types vary in relation to the ownership of the format, quite apart from the content or the medium of the file. For example, the “shelf life” of files that are in proprietary formats (and thus under the control of a single commercial organization)^[36] is significantly different from those that are nonproprietary^[37] or proprietary,^[38] but can be rendered and edited by a range of different pieces of software written by different organizations or individuals. Paper documents left in boxes in a storeroom decay relatively slowly (unless the environmental conditions are particularly unfavorable) and gradually, so that inspection will reveal when action needs to be taken to preserve them. Digital material, which needs hardware and software to render it, is very different. The media on which digital material is recorded are relatively unstable and a programme of transfer to new media is an essential management action. File formats suffer from technical obsolescence as hardware and software is developed and the environments that once rendered the file disappear from everyday use. This tends to happen more quickly for proprietary formats, as commercial pressures require constant change to ensure that products continue to produce healthy revenue streams.

COLLECTIONS

Harvey in his useful treatment of the subject in a digital environment is unequivocal in declaring “Selection decisions are essential in managing collections of heritage materials.”^[39] Not everyone would agree that selection to form collections is desirable. Lynch claims “in the humanities scholars are using the digital medium to circumvent the haphazard historical gerrymandering of knowledge into institutional collections belonging to communities.”^[40] He believes that objects simply by being available on the Web will “find their own unexpected user communities” and there is, therefore, no need to select.^[41] Although it is difficult to quarrel, this approach implies a keep everything “for ever” mentalité that is expensive and almost certainly unrealistic, and also that search engine algorithms are value free. Over a decade ago Miksa corrected the mistaken impression that even then the whole content of the web could be thought of as akin to a giant library or archive, preferring instead to see it from a library perspective as just an extension of the bibliographic universe from which to choose.^[42] Selection of content to form collections depends on the survival and availability of material that accord with the intentions of the collector that may not of course be consistent over time. It is often assumed by information professionals that they are the sole arbiters in the process of selection rather than being, as it were, the bar tenders in the “Last Chance Saloon.” The survival of all objects depends on chains of decisions that, even if haphazard, are rarely accidental. Such intentions are more often than not the context from which the object is chosen by a library, an archive or museum or even individuals to form part of their collections. For example a library that specializes in the French language will buy books about French not Afrikaans, an archive that collects business papers will be unlikely to acquire papers of bishops or imams, and an individual who collects broadsheets will not buy encyclopedias. We must never lose sight of the fact that collections are about content, something that enthusiasts for digital repositories sometimes seem to fail to grasp in trying to build up their holdings as part of a numbers game or a quest for a genre with unusual technical features.

The very concept of a collection is not straightforward and is open to a number of interpretations.^[43] A library might call the whole of its holdings a collection or identify separate collections of significant material as special collections. An archive will invariably define collections in its holdings. Such differences often reflect ownership or provenance or typology or combinations of them. In the analogue the curator can simply point to the shelves that contain a collection and the same applies even when a collection is distributed. In the digital, boundaries are much more fluid and single objects can be selected from any source to form a collection and never need to be physically assembled in a library, archive, or museum let alone a single digital repository. This may benefit users, but only serves to complicate asset management. The virtual collection will inevitably be worth more than the parts, but there will always be a risk that some parts may disappear at the flick of a switch. It is easy to magnify such risks. They were there in the analogue, particularly when collections were in private hands. In 1972 the Earl Fitzwilliam burnt 19 tons of family papers on a bonfire at Wentworth Woodhouse in Yorkshire in northern England that lasted 3 weeks.^[44] Famously Honoré Jaxon’s huge collection of papers about Louis Riel’s Resistance in Canada in 1885 became landfill after his eviction from his New York home in 1952.^[45] Data loss, like so much else, depends on the context in which it was created and held. It would be reprehensible if data that supported an important scientific advance that was funded by external agencies was deliberately destroyed, whereas a poet or artist has every right to destroy or never to keep their working notes or drawings. This may be regretted and efforts may be made to persuade them otherwise; but every individual or organization has a perfect right to destroy data providing they are not contravening regulation or have assessed the risk of so doing.

A fundamental purpose of libraries is to provide access to information that has been collected and preserved. In the paper world and the early years of the digital world, the best way to achieve this was by owning copies of the relevant material and lending this to patrons or allowing them reference access within the library. Increasingly in the digital world, libraries are negotiating access agreements to digital assets on the servers of publishers, aggregators, and other types of collection owners or agents,

and which members can access on their desktops.^[46] The collections are becoming (at least partly) virtual and there is increasing concern as to what happens to access at the termination of agreements. In the paper world when subscription to a journal series is terminated, patrons can still obtain access to the issues subscribed to, but this is not guaranteed in the e-journal world. There are a number of projects and services that are being developed to try to address this problem. One such is the open source LOCKSS (Lots Of Copies Keeps Stuff Safe)^[47] project that provides a mechanism for libraries to cache copies of journal content if their contract with the publisher permits this and furthermore to compare the cached content with other participating libraries and, if necessary, repair it to take care of problems that may result from corruption of the files. Another approach is being taken by the Portico service which “provides a permanent archive of electronic scholarly journals,” to which libraries can subscribe.^[48] The former approach gives libraries a type of ownership of the bit patterns representing the e-journals, but is designed to be a very low maintenance system which requires a low level of technical expertise, and therefore cost. The latter “outsources” the maintenance to the Portico service, but does not give the library any direct form of ownership of the bit patterns.

In the academic world, libraries have become increasingly interested in institutional digital repositories as a means to “own copies of” and manage the research output of their institutions. There is an interesting debate as to whether academics see their institution as the natural route for publishing their work or if discipline-based repositories fit more naturally with their culture and content. Davis and Connolly found that faculty at Cornell University showed relatively little interest in the institutional repository, tending towards a disciplinary culture rather than an institutional one.^[49] It will be intriguing to see whether or not institutional repositories become established within academic culture and libraries continue to have an “ownership” role for institutional content as part of the development of their collections in the digital age. Attempts to force members of the academy to deposit research outputs and core data are unlikely to succeed as they immediately raise questions about academic freedom and intellectual property rights. There is no shortage of open source software for the creation of such repositories, for example, e-prints, Fedora, and DSpace are readily available.^[50] Although these systems are intended to make the management of digital assets relatively easy and cost-effective, much time, effort, and resource is required to master them and put them to use and this is really only practicable in large, well-resourced organizations where a clear payback can be demonstrated.

In developing digital collection policy digital asset managers need to establish, as information managers have always had to, who the most appropriate custodians of the material under consideration are. Is this collection a place where such material should be housed? Such questions are never easy and the long-term issues around cost and management of digital material make the consequences of such decisions ever more involved. This is, of course, how it has always been and digital storage and access does not change this. What does change is the ability of third parties to discover material that may be held by a library or archive either illicitly or more commonly being made available without permission of copyright holders. This is particularly true of any performing arts data that by their very nature embody a maze of subsidiary rights where even the best intentioned curators may have trouble tracing ownership. Such complex IPR issues increasingly extend to all types of media, for example, a Web site will incorporate graphic design and images with associated rights as will any document with embedded behaviors that depend on licensed software. Creators of content, including supporting software, and their professional associations and advisers are becoming much more adept at recovering revenue for its use and in policing abuse. Curators who call for some distinction to be made in international copyright law between popular and academic content fail to grasp both how difficult it might be to draw the boundary and the sensitivities of the academic community about IPR and freedom of expression.

MANAGEMENT

The management of digital assets, like that of any other asset, cannot be separated from the overall management of an organization, whether it be a library, museum, archive, a university, or a

commercial organization. This means that the senior management of the organization is responsible for the management of digital information assets just as much as they are for other types of asset. With the widespread introduction of the International Financial Reporting Standards (IFRS) organizations are required to declare such assets in their balance sheets, particularly if there is an associated revenue stream, for example, the sale of digital copies, depreciate them and report any contingent liabilities, such as the ownership of content by third parties.^[51] For the digital asset manager, this is both an opportunity and a challenge. The opportunity derives from the fact that digital information assets have potential “value” for the organization, the challenge is to be able to present the opportunities to senior management in a way that enables the managers to understand the potential and the risks of not managing them well. In these circumstances auditors can be helpful both in determining “value” for accounting purposes and putting in place appropriate curatorial practices.

Information assets are normally classified in a balance sheet as intangible if they are tangential to an organization’s primary goals. Valuing intangibles is tricky, because there has always been a reluctance on the part of accountants to attribute more than a nominal value to them even if a great deal of investment has been made in them, particularly since the dot-com bubble of 1995–2001 which attributed unrealistically large values to such assets. The *espida* project at the University of Glasgow developed a methodology that is widely applicable for this purpose and places emphasis on business need and by extension artful processes.^[52]

The priority management gives to managing its digital assets will depend on organizational goals. A university research library that subscribes to many digital resources is likely to take it more seriously than a public library whose principal activity is to lend books. A company in the financial services sector will give it the highest priority if it wishes to remain in business. The ICT community assured management that the problems of storage, access, and preservation that were becoming acute in the analogue environment could be easily resolved by developments in technology. At the most extreme this could be expressed as “digitize the lot and destroy the analogue originals.”^[53] In some parts of the commercial world, particularly financial services, this strategy was pursued with success, albeit at a cost as much in workflow and data analysis as technology, whereas it led to very heated debates amongst librarians, especially about newspaper collections that were often printed on fragile paper that was both expensive and difficult to conserve and access.^[54]

Digitization programs may no longer be so radical, but they often are predicated on the use of cheap off-site storage of the analogue originals. Such an equation can only apply where the analogue objects are already heavily used, as would be case with census returns, or where information can be more easily unlocked from a “machine readable” digital resource such as local newspapers. Initiation costs for making digital copies is expensive and it is often easy to overlook the potential costs of maintenance and exaggerate revenue streams.^[55] Increasingly digitization projects result in the provision of online resources as part of a suite of digital services, for example, in the Digital Collections and Programs in the Library of Congress^[56] and the Catalogues et Bibliothèque Numérique at the Bibliothèque Nationale de France.^[57] Evidence suggests that, as in these two examples, where digital objects are in active use, their curation will be guaranteed as long as they are considered to be mission critical, but the same is true of the analogue. Where there is a business need to keep data alive human resourcefulness and investment will be found.^[58]

Problems arise when information objects cease to be needed for current business. This is where a clear distinction needs to be drawn between the content of different objects. A library may be quite happy to delete an online book or cancel a subscription in the confidence that an exact copy will be held elsewhere, for example, in a copyright library, whereas in other contexts there may be requirements to retain noncurrent data for long periods, for example, under the 2002 Sarbanes–Oxley Act in the United States or the introduction in many jurisdictions of the UNECE Århus Convention on Access to Environmental Information.^[59] It is no longer possible to assume that such controls only apply in a corporate context, particularly if a research library is in receipt of original electronic data that will need to be reviewed on a case by case basis, and in the United States where libraries and

archives are subject to the Patriot Act. The degree to which any organization complies with the regulatory environment is a matter of balancing the risk of noncompliance against the costs of complying. If the penalties for noncompliance are relatively trivial, there is little incentive to invest in the retention of noncurrent information for such purposes. Regulation very rarely requires permanent retention (records relating to fissile material are one instance) and normally stipulates a period after which information assets can be destroyed. These will vary between categories and jurisdictions, and over time. All strategies for the management of digital assets must include procedures for secure destruction that are not as technically straightforward as they may seem. Where assets are retained beyond stipulated retention periods, any contingent liabilities that may reside in them remains and there have been several high profile cases where records that could have been destroyed have been discovered by the courts resulting in heavy penalties.^[60] If, however, the noncurrent information has an identifiable ongoing “value” to an organization and the contingent liability arising from retention is considered low then a case may be made to the management to invest in long-term preservation.

In some instances, such as observational data from the solar system or the global environment or digital art, advancing a case for complete permanent retention as a business need is relatively straightforward. This does not mean arguments do not have to be advanced for costs to be subsumed within the overall resource allocation. In the majority of instances long-term, let alone permanent retention, will involve appraisal, privileging, or selection, principally to avoid any contingent liability and to remove objects considered of little value. This is much more complex than appraisal in the analogue because of the very granularity of many digital resources and associated IPR issues. Although there has been much discussion of macro and automated appraisal, the weakness of both approaches is that they embed the formulization of business process by information professionals; for example: “Macro-appraisal asserts that archivists not researchers or creators are society’s professional agents appointed by law to form its collective memory.”^[61] Even for government records it is hard to envisage how this could be the case and in the private sector such an approach would be incredible. What it overlooks is the long tail of intentionality that underpins all information objects from the instant of their creation to their ultimate disposal. In rejecting any “scientific formula” for appraisal decisions Greene was unequivocal in stating that just because it is “localized and contingent”; “Appraisal is messy and complex, and its results are neither predictable nor uniform.”^[14] Nevertheless management will need to be confident that the process respects the due diligence that stakeholders expect.

A final challenge of the management of digital information assets is to ensure that they are managed in a secure way, such that they are not accidentally lost or fall into the wrong hands. Digital asset managers need to have in place systems and processes that provide an appropriate level of surety that deposits entrusted to their care will be managed in a way that is commensurate with their value. What this means is not that there is a single defined standard for digital asset management, but that the effort and cost expended on management should be proportionate and that depositors have realistic understanding of what to expect. There is no such thing as an absolute guarantee that a digital asset will be available, accessible, and interpretable for all time. Some assets are required for a set period of time and this is followed by another, perhaps longer period, during which they have value but their survival is less “critical.” Such distinctions should be reflected in management practices.

LOGISTICS

Increasingly the tasks involved in digital information asset management are complex and require access to specialist services such as file type identification, format conversion, normalization, encryption/decryption, digital signatures, bit-pattern validation, persistent identifiers, and so forth. There are clear trends toward managing assets through a collection of loosely coupled services over the Internet, rather than via monolithic applications.^[62] This trend toward Service Oriented

Architectures is described in a very readable article by Barbara Reed.^[63] Whilst this article is written in the context of records management, it applies equally in other areas of information science. Digital asset managers will need to design environments that are appropriate to their needs from such tools and services or contract out this work to specialist consultants. Systems must be designed to suit the type of assets to be managed, and the objectives of the parent organization, providing

- Adequate storage media of appropriate types
- Context, metadata, and finding aids
- Access mechanisms
- Management and preservation services

taking account of regulation and compliance and both management and user perspectives.

Once a business case has been established for preservation, accepted and resources allocated, the capture of digital assets is not as straightforward as the analogue equivalent. There are those who argue that preservation requires intervention by librarians and archivists as the ultimate curators in the design of information systems. There are attendant dangers in confusing these two functions. Publishers may not welcome interference or advice from librarians, preferring to continue to control their digital assets so that they can be repurposed and repackaged, either in part or in whole, in the future. In an academic context it is for senior management and funding bodies to negotiate rights of access to and the deposit of assets that they have financed. This is contested territory that, as we have discussed, raises issues of IPR and freedom of expression that are difficult to resolve. In both the public and private sector the task of records managers is to manage current and noncurrent information within the context of corporate objectives. Sometimes this will result in either the creation of an archive or the transfer of assets to an archive, but often it will not and in these circumstances deposit will be a matter for negotiation. Since the constant is the information object, there needs to be consultation, but neither the records manager nor the archivist or librarian can dictate processes that must fulfill the business needs of the organization and not their own expectations in the way much of the literature on digital preservation suggests.

CONCLUSION

The best advice to a librarian or archivist confronted with the management of digital assets is to remember that it is the content that matters, not the technology. It is salutary to remind those who proclaim that the digital environment represents an information paradigm shift that there have been other revolutions in the management and distribution of information and knowledge with accompanying enabling technologies, such as printing in the fifteenth century and photography in the nineteenth. Librarians and archivists learned to deal with the far-reaching implications of these changes without losing sight of their core objectives to manage, curate, and make available information for the benefit of user communities. They have regularly had to reform their intellectual framework and the ways in which they curate, organize, and reference their collections to accommodate technical change and to match societal expectations which in themselves are constantly modified by technologies. With such a pedigree there is no reason to suppose they will not be a match for the challenge of the digital.

ACKNOWLEDGMENTS

The authors would like to thank their colleagues and the doctoral students in the Humanities Advanced Technology and Information Institute at the University of Glasgow, particularly Dr. Susan Stuart, and successive students on the M.Sc. Information Management and Preservation programme for their help and inspiration.

REFERENCES

1. <http://www.digitalpreservation.gov/> (accessed November 2007).
2. <http://www.interpares.org/> (accessed November 2007).
3. <http://www.nla.gov.au/padi/> (accessed November 2007).
4. <http://www.afc.ryukoku.ac.jp/english/english.html> (accessed November 2007).
5. <http://www.digitalpreservationeurope.eu/about/> (accessed November 2007).
6. <http://www.dpconline.org/graphics/index.html> (accessed November 2007).
7. <http://www.oclc.org/uk/en/news/releases/200692.htm> (accessed March 2008).
8. 'Importance of Digital Preservation' Digital Preservation (Library of Congress) website, <http://www.digitalpreservation.gov/importance/> (accessed November 2007).
9. See for example: *Trusted Digital Repositories: Attributes and Responsibilities*; An RLG-OCLC Report. Mountain View, CA, May 2002, <http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf> (accessed March 2008).
10. http://www.digitalpreservationeurope.eu/publications/dpe_research_roadmap_D72.pdf.
11. Ross, S. *Digital Preservation, Archival Science and Methodological Foundations for Digital Libraries*; keynote address at Ubiquitous Digital Libraries—11th European Conference on Research and Advanced Technology for Digital Libraries, Budapest 2007, author's draft available at: http://www.ecdl2007.org/sross_draft_ECCL_paper.pdf (accessed November 2007).
12. Harvey, D.R. *Preserving Digital Material*; Saur: München, 2005; 25.
13. Himmelfarb, G. Revolution in the library. *Libr. Trends* **1999**, 47 (4), 612–619.
14. Greene, M.A. Not magic, not science, but art: Comment on 'archival appraisal alchemy' (response to Richard Cox). Choices & Challenges, Collecting by Museums and Archives 2002, <http://www.hfmvgv.org/research/publications/symposium2002/papers/greene.asp> (accessed November 2007).
15. 'Vindolanda on Line' <http://vindolanda.csad.ox.ac.uk/> (accessed November 2007).
16. Burke, P. *A Social History of Knowledge from Gutenberg to Diderot*; Polity Press: Cambridge, U.K., 2000.
17. Headrick, D. *When Information Came of Age: Technologies of Knowledge in the Age of Reason and Revolution 1700–1850*; Oxford University Press: Oxford, U.K., 2000.
18. Levy, D.M. *Scrolling Forward: Making Sense of Documents in the Digital Age*; Arcade: New York, 2001.
19. McLuhan, M. *The Gutenberg Galaxy: The Making of Typographic Man*; University of Toronto Press: Toronto, ON, 1962.
20. Webster, F. *Theories of The Information Society*; Routledge: London and New York, 1995.
21. McKitterick, D. *Print, Manuscript and the Search for Order 1450–1830*; Cambridge University Press: Cambridge and New York, 2003.
22. This work is described in: Currall, J.; Johnson, C.E.; Johnston, P.; Moss, M.S.; Richmond, L.M. "No Going Back?" *The Final Report of the Effective Records Management Project*; Project funded under the JISC Technology Applications Programme JTAP—375, 2001, <http://www.gla.ac.uk/infostrat/ERM/Docs/ERM-Final.pdf>. (accessed November 2007).
23. Gilliland-Swetland, A.J. *Enduring Paradigm, New Opportunities The Value of the Archival Perspective in the Digital Environment*; Council on Library and Information Resources: Washington, DC, 2000.
24. Hill, C.; Yates, R.; Jones, C.; Kogan, S.L. Beyond predictable workflows: Enhancing productivity in artful business processes. *IBM Systems Journal – Business Collaboration* **2006**, 45 (6), 674, <http://www.research.ibm.com/journal/sj/454/hill.html> (accessed November 2007).
25. Harvey, D.R. *Preserving Digital Materials*; Saur: München, 2005; 22–23.
26. <http://www.jacar.go.jp/english/> (accessed March 2008).
27. Kiernan, K. *Electronic Beowulf*; British Library: London, 2004.
28. <http://www.dlv.org.uk/> (accessed November 2007).
29. Webb, C. *Limitations and Pitfalls of Digitisation*; National Library of Australia, Staff Papers, 2000, <http://www.nla.gov.au/nla/staffpaper/2000/webb5.html> (accessed November 2007).
30. See for example Carpenter, J.; Antonowicz, L. *Digitisation of Learning Materials and Community Grids for Learning: Final Evaluation Findings*; Big Lottery Fund, 2006; http://www.biglotteryfund.org.uk/er_eval_digi_final.pdf (accessed March 2008), No.26.
31. A useful timeline, admittedly with a US perspective can be found in the Digital Preservation Management tutorial compiled by Cornell University at <http://www.library.cornell.edu/iris/tutorial/dpm/timeline/index.html> (accessed November 2007).
32. An overview of some of these characteristics may be found in Currall, J.; Johnson, C.E.; Johnston, P.; Moss, M.S.; Richmond, L.M. "No Going Back?" *The Final Report of the Effective Records Management*

- Project*; Project funded under the JISC Technology Applications Programme JTAP—375, 2001, [http://www.gla.ac.uk/infostrat/ERM/ Docs/ERM-Final.pdf](http://www.gla.ac.uk/infostrat/ERM/Docs/ERM-Final.pdf) (accessed November 2007).
33. Allison, A.; Currall, J.; Moss, M.; Stuart, S. Digital identity matters. *JASIST* **2005**, *56* (4), 364–372.
 34. The Internet Archive can be accessed via the ‘Way Back Machine’ at: <http://www.archive.org/index.php> and a good description of its philosophy and how it operates is given in Wikipedia at: http://en.wikipedia.org/wiki/Internet_Archive (accessed March 2008).
 35. The Dublin Core Metadata Initiative (DCMI) is described at: <http://dublincore.org/> “The Dublin Core Metadata Initiative provides simple standards to facilitate the finding, sharing and management of information.”
 36. The file formats used by the Microsoft Office Suite of products. <http://www.microsoft.com/> is a good example. (accessed March 2008).
 37. LaTeX. <http://www.latex-project.org/> (accessed March 2008) is a non-proprietary document preparation system that has maintained stable file formats over a period of two decades having been first developed in 1985 by Leslie Lamport.
 38. The Portable Document Format (PDF) belongs to Adobe Systems (<http://www.adobe.com/>), but its definition is published and there are many tools that have been written that can read, write and edit the format.
 39. Harvey, D.R. *Preserving Digital Materials*; Saur: München, **2005**, *55*, 53–75.
 40. Lynch, C.A. Colliding with the real world: Heresies and unexplored questions about audience, economics, and control of digital libraries. In *Digital Library Use: Social Practice in Design and Evaluation*; Bishop, A., Butterfield, B., House, N., Van., , Eds.; The MIT Press: Cambridge, MA, 2003.
 41. Lynch, C. Digital collections, digital libraries and the digitization of cultural heritage information. *First Monday* **2002**, *7* (5). http://www.firstmonday.org/issues_7/lynch/ (accessed November 2008).
 42. Miksa, F.L.; Doty, P. *Intellectual Realities and the Digital Library*; University of Texas: Austin, TX, 1994; 16. <http://www.csd.tamu.edu/DL94/paper/miksa.html> (accessed November 2008).
 43. Lee, H.-L. What is a collection? *J. Am. Soc. Inform. Sci.* **2000**, *51* (12), 1106–1113.
 44. The mansion of mystery and malice. *The Sunday Times* February 11, 2007.
 45. Smith, D.B. *Honoré Jaxon: Prairie Visionary*; Coteau Books: Regina, SK, 2007.
 46. Kane L.T. Access versus Ownership.
 47. LOCKSS, in *Encyclopedia of Library and Information Science*, 2nd Ed., p. 50. <http://www.lockss.org/lockss/Home> (accessed March 2008).
 48. Portico, <http://www.portico.org/> (accessed March 2008).
 49. Davis, P.M.; Connolly, M.J.L. Institutional Repositories: Evaluating the Reasons for Non-use of Cornell University’s Installation of DSpace. *Dlib Magazine* March/April **2007**, *13* (3/4), <http://www.dlib.org/dlib/march07/davis/03davis.html> (accessed March 2008).
 50. Eprints – <http://www.eprints.org/>, Fedora – <http://www.fedora-commons.org/>, DSpace – <http://www.dspace.org/> (accessed March 2008).
 51. http://www.icaew.co.uk/library/index.cfm?AUB=TB2I_25594. (March 2008).
 52. <http://www.gla.ac.uk/espida/> (accessed November 2007).
 53. Baker, N. *Double fold: libraries and the assault on paper*; Random House: New York, 2001.
 54. McKitterick, D. *Do we want to keep our newspapers?* Office for Humanities Communication Publication: London, 2002.
 55. Moss, M.; Currall, J. Digitisation: taking stock. *J. Soc. Archivists* **2004**, *25* (2), 123–137, (accessed March 2007), from <http://eprints.gla.ac.uk/1059/>.
 56. <http://www.loc.gov/library/libarch-digital.html> (accessed November 2007).
 57. <http://www.bnf.fr/pages/catalogues.htm> (accessed November 2007).
 58. Waller, M.; Sharpe, R. *Mind the gap—Assessing digital preservation needs in the U.K.*; Digital Preservation Coalition: London, 2006, <http://www.dpconline.org/docs/reports/uknamindthegap.pdf> (accessed November 2007).
 59. Buckley, J.W. *New Life for Dead Storage Technologies*; Storage Networking Solutions (SNS): Europe, 2007, <http://www.sns europe.com/snslink/news/printer-friendly.php?newsid=5759> (accessed November 2007).
 60. *FDA V. Brown & Williamson Tobacco Corp.* (98–1152) 529 U.S. 120 (2000) 153 F.3d 155, affirmed, and *British American Tobacco Documents Archive* (2004), <http://bat.library.ucsf.edu/> (March, 2004).
 61. Cook, T. *Appraisal Methodology: Macro-Appraisal and Functional Analysis*; Library and Archives of Canada: Ottawa, 2001. <http://www.collectionscanada.ca/information-management/007/007007-1035-e.html#tphp> (accessed November 2007)
 62. Kim, Y.; Ross, S. Genre Classification in Automated Ingest and Appraisal Metadata. In *Proceedings UK e-Science All Hands Meeting 2006: Achievements, Challenges, and New Opportunities*, Nottingham, Cox,

- S.J., Ed.; National e-Science Centre; Glasgow and Edinburgh, 2006, http://eprints.erpanet.org/110/01/genre_extraction_KIM_ROSS_2006_ECDL.pdf (accessed November 2007).
62. A good example are the services provided by the UK National Archives based on its PRONOM file format registry (<http://www.nationalarchives.gov.uk/pronom/>), which can be accessed via the DROID (digital record object identification) tool <http://droid.sourceforge.net/wiki/index.php/Introduction> (accessed March 2008).
 63. Reed, B. Service-oriented architectures and recordkeeping. *Records Management Journal* **2008**, 18 (1), 7–20.

11 Network Management

Robert J. Sandusky

CONTENTS

Introduction.....	149
What Is Managed in a Network?	150
Dimensions of Network Management	150
Perspectives on Network Management Work	153
Patterns of Work.....	153
Patterns of Work and Dimensions of Network Management.....	156
Information and Activity in Network Management	157
Conclusion	159
References.....	160

INTRODUCTION

Network management includes activities related to the design, construction, and operation of communications networks, which are one type of distributed infrastructure. Network management activities are performed by a network management organization (NMO). An NMO is one of many types of information technology service organizations and, as discussed here, is responsible for the design, monitoring, and evolution of distributed infrastructure (an assemblage of hardware, software, and services) in its particular organizational context. The network itself is the primary object of the NMO's work and the NMO generally has little influence over how the network is used by other members and affiliates of the organization. An NMO may be a single organization but will typically interact with multiple organizations including customers, peer organizations collaborating to provide network services, and vendors. Network management, in all but the simplest cases such as a home or small single-site local area network (LAN), is inherently a distributed activity: a network usually includes equipment distributed in multiple locations. At the smaller end of the scale, equipment may be located in multiple rooms or on multiple floors of a building (referred to as a LAN). In more complex cases, the equipment may be located in multiple buildings (on a campus or across a metropolitan area, often referred to as a MAN, or metropolitan area network), in multiple cities, or even multiple countries (a WAN, or wide area network). This entry will assume, for purposes of illustration and discussion, a MAN or WAN unless otherwise noted.

This entry begins with a brief description of the types of components that make up a MAN or WAN and are the network components that require management. The five major dimensions of network management, as defined by the International Standards Organization (ISO) as part of the Open Systems Interconnection model, are used to introduce the range of activities associated with management of complex networks. Next, how design work is approached in NMOs is discussed in terms of the classic "waterfall" software development life cycle. The dynamic and collaborative nature of work within an NMO is described in terms of two interrelated modes of working: design work and real-time supervisory control work. Finally, the critical role of information management is described in the context of a disaster recovery/business continuity test.

WHAT IS MANAGED IN A NETWORK?

Networks, including the Internet, closed or private networks built by organizations, and smaller-scale LANs, are composed of several classes of components. Applications (e.g., Web sites, enterprise databases, digital libraries, e-mail systems) made accessible to remote users, messages carried by the network (i.e., frames, packets), end-user computers (i.e., client computers), server computers (e.g., Web servers, file servers, proxy servers), routers, switches, hubs, equipment racks, link signaling devices (i.e., link encryptors, digital signaling units/digital “modems”), access lines (e.g., the Ethernet cabling from client computers to a local switch, router, or hub), wireless access points, and long-distance telecommunications circuits (e.g., the higher-capacity, leased digital line between network locations usually connecting routers to each other).^[1] All of these network components, except the cables, contain software and configurable parameters and all must be managed in order to provide effective network services that evolve to meet demand and keep pace with technical innovations. Vendor-provided software for configurable components must be tested and implemented to fix bugs, patch security holes, and to add functionality. The NMO also monitors the network in real-time for failures, errors, traffic bottlenecks, etc., and corrects problems by issuing commands to change component state and by swapping failed components with replacements. The NMO also monitors actual usage patterns and takes requests for new services into account as part of its network design and provisioning work.

The NMO must have the means to organize information about all the elements of the network in order to manage these components. The NMO must also have a set of operating parameters that, in combination with the network components, can enable the NMO to create appropriate standard operating procedures (SOPs) and ancillary support systems such as trouble ticket, configuration management, test scheduling, and reporting systems.

DIMENSIONS OF NETWORK MANAGEMENT

A conceptual framework associated with the ISO’s Open Systems Interconnection model for communications systems is used to organize the full range of network management activities.^[2] While originally focused on network management, this framework may be used to characterize the management of distributed computing systems more generally. The OSI framework is broken down into five functional areas: fault management, configuration management, performance management, accounting management, and security management.

Fault management is the detection, diagnosis, and correction of problems, or faults, in a network. Network components asynchronously generate fault data as messages, referred to as alarms, which are categorized as major, minor, and informational depending upon their significance or impact. Major alarms are generated when a network component fails due to hardware failure, software failure, or because the component can no longer be contacted by the network management systems (i.e., it has become isolated from the rest of the network). Minor alarms are generated when less critical events are detected, such as a performance threshold being exceeded, a burst of transmission errors is detected, or a failure of a subcomponent (e.g., a single interface card failure on a large router). Informational messages are generated for a variety of reasons, do not require direct attention by the real-time supervisory control (RTSC) staff, but are logged and available if needed (e.g., an anomalous but automatically recoverable condition, such as a communications protocol error). As alarms are generated, they are propagated to the NMO via both a real-time and an archival path. Alarms are sent to network management workstations that are continuously monitored by RTSC staff and to databases for long-term alarm storage. Propagation of alarms to the network management workstations supports the real-time work of isolating and correcting faults requiring direct intervention by NMO staff. Examples of faults requiring real-time intervention include failures that do not automatically recover or large-scale failures that trigger the performance of a SOP. Propagation of alarms to databases provides comprehensive historical data for postevent, trend, or long-term fault analysis.

Faults are logged in a database using a record format that includes date, time, location, fault codes, and other information of potential use in diagnosis, correction, or to support subsequent, long-term analysis. To support fault diagnosis, NMO staff use diagnostic test tools and component status query commands to retrieve real-time component status, and various component activation/deactivation commands to alter component status or component operating parameters. NMO staff are trained in troubleshooting procedures, and have access to SOPs for monitoring the network, documenting fault conditions in trouble tickets, and problem escalation (i.e., when to contact next-level expertise and/or management if a fault cannot be corrected in a timely manner). Trouble tickets are created for all faults requiring intervention. The trouble ticket is a locus for a textual narrative of the situation as it unfolds, providing capabilities to combine descriptions of evidence, activity, causal hypotheses, as well as a record of interpersonal, intra-, and inter-organizational communications. To correct faults that have been diagnosed, the NMO staff use commands to alter physical configurations (e.g., to replace a failed component with a backup component), manually activate a deactivated component, or to modify the operating parameters of affected components.

Configuration management is the set of capabilities that allow the NMO to set up and modify the network's physical and logical configurations. Following initial implementation the network's physical and logical configurations will change "to alleviate congestion, isolate faults or meet changing user needs."^[2] Configuration changes may be unplanned, caused by network faults, or as a result of NMO staff action in response to a fault or changing, real-time network conditions, or as part of a preplanned configuration change. Faults often cause spontaneous network reconfigurations, such as removal of a long-distance telecommunications link from the network topology. Networks must be fault-tolerant: their internal complexity (the large number of varied and interconnected components) and reliance on long-distance telecommunications links result in frequent (e.g., several per day) single component failures. Therefore, networks must be designed to tolerate all single-component failures in order to be considered minimally reliable. In order to provide availability on the order of 99.9% or higher, the design must tolerate the simultaneous failure of any two components. In some situations, networks are designed to tolerate all dual-component failures as well as a subset of three-component failures. Network components are designed to attempt automatic recovery following failures, and many failures are corrected automatically. In a significant percentage of cases, NMO diagnosis and correction are required to either effect component recovery (e.g., manually reactivating a component) or to temporarily reconfigure the network so connections and traffic flow around the failure (e.g., moving a LAN's connection from its failed default router to an available connection on an alternate router). Planned network reconfigurations take two forms: predesigned modifications that are intended to be permanent, such as changes made to address new requirements, and predesigned modifications that are intended to be temporary, such as those that are part of disaster recovery/business continuity plans. Adding, modifying, or removing routers, interface cards to a router, or telecommunications circuits between routers are examples of common planned permanent configuration management activities. Network planning and design work is required prior to implementation of either a new network or a significant revision to an existing network. One of the implications of designing a fault-tolerant system is the need to preplan the organization's responses to a range of anticipated fault events including normal, single-component failures (e.g., loss of a long-distance telecommunications link) or high-impact failures (e.g., complete loss of connectivity for a network site, whether it is caused by natural disaster, massive telecommunications link failure, fire, simultaneous failure of both primary and backup routers, etc.). Preplanning work is complex, and results in the creation of SOPs that guides NMO reconfigurations in response to classes of failures (see discussion below in the section on "Perspectives on Network Management Work").

Performance management refers both to the capabilities in the network that support measurement of network use and the expertise within the NMO to analyze data, justify recommendations for change, and plan the configuration changes needed to support emerging or anticipated network uses. NMOs are interested in understanding and documenting past usage patterns, comparing

current usage patterns to historical patterns, and using both historical and current data to support the capacity planning process. While performance monitoring occurs in real-time as the RTSC staff continuously monitor utilization of network components like routers and long-distance telecommunications links, the bulk of performance management activities are performed “off-line” by design staff. Network components periodically—every few minutes—log levels of usage by writing data to network management servers for later analysis. Network components may also generate fault messages or alarms when utilization levels exceed certain thresholds to enable real-time supervisory staff to respond with network configuration changes. Networks and their protocols are designed to be tolerant of fluctuations in network traffic or topology, so real-time intervention is not usually required. Performance management is part of the longer-term network design process and involves analysis of the logged performance data in order to detect changes in usage patterns that lead to modifications in the network’s configuration. For example, increasing network traffic between two locations may lead to installation of higher capacity routers, telecommunications links, reconfiguration of the network topology, or a combination of the three.

Accounting management is the set of capabilities that allows the NMO to collect and manage usage and configuration data, track costs, and charge network users according to the tariffs, service level agreements, or contracts that are in place. Several network components may be capable of generating and transmitting usage data that may be fed into an off-line accounting or billing system (e.g., routers and servers running shared applications may generate data that are fed into a billing system). Costs of other components, such as cables, routers, or equipment racks, may have their costs allocated on some other basis (such as employee head count or a count of connections), or may be factored into usage-sensitive charges. NMOs evaluate the costs, benefits, and complexity of collecting and correlating large amounts of data and negotiating customer agreements in order to produce detailed bills against other approaches that don’t require detailed usage-based calculations (e.g., simpler, flat-rate or percentage-allocated cost-recovery arrangements). The simpler approaches are particularly useful for networks built to support the internal operations of an organization. NMOs providing services to the public, such as land-line and mobile telecommunications organizations, are more likely to develop usage-sensitive procedures.

Security management is concerned with control of physical access to network components, end-point authentication, user authentication, key management for encryption, end-to-end data encryption, security log management, monitoring emerging security threat vectors and developing organizational responses, and firewall management. Security management has two objects in the network management context: first, the network itself must be secured (e.g., control the access to router management ports which are used to issue reconfiguration commands to the router; ensure that physical access to network equipment at all locations is controlled; manage NMO staff authentication and access); second, the network must support the security requirements of the users and applications transmitting data across the network (e.g., providing encryption of telecommunications circuits; end-to-end encryption of data streams; management of encryption keys). In environments with high security needs, including financial institutions, health care organizations, and commercial organizations in highly competitive markets, security management responsibilities are concentrated in an independent part of the organization, with a management chain distinct from the RTSC and technical staffs. A separate management chain encourages, enforces, and makes a high degree of transparency possible through establishment of arms-length relationships between different parts of the NMO. NMOs with these types of stringent needs will also employ a comprehensive internal auditing process, augmented by recurring external audits. Security logs, like fault logs, identify date, time, location, and diagnostic information. Security logs should be reviewed regularly in order to identify unusual usage patterns or to correlate successful and unsuccessful access attempts by users and intruders alike. Several of the activities included in security management, including encryption key management and developing responses to emerging security threats, are complex management challenges in a distributed environment.^[3]

PERSPECTIVES ON NETWORK MANAGEMENT WORK

Network management depends upon the blending of actions performed by people with a range of skills and expertise, and these actions depend upon information management practices. Design and RTSC work patterns are defined, and the central role of information management in network management is described. These two work patterns are discussed in the context of the five dimensions of network management. The section concludes with a description of a complex yet typical activity, a disaster recovery test (DRT), which illustrates the relationship between information and activity in distributed infrastructure management.

PATTERNS OF WORK

Network management combines patterns of both design and RTSC work. Research has shown that design work and RTSC work have different rhythms and patterns that, in network management, are coordinated to perform fault, configuration, performance, accounting, and security management. RTSC work is event driven, dynamic, and involves the interaction of multiple workers with complex systems.^[4] Malfunctions in the operation of these systems can result in serious consequences such as the loss of human life or financial loss. Signs of malfunction must often be recognized and responses formulated within minutes or seconds in order to prevent adverse consequences. By contrast, design work has different patterns and rhythms than RTSC work and is characterized as being self-paced (as opposed to being event-driven), reflective (as opposed to being time-constrained), and creative.^[4–6] The object of design work is usually to iteratively create, study, or revise some artifact like a document, a set of specifications, a product, or a system, such as a data communications network. In network management, there are frequently occasions when workers who usually do design work join the RTSC staff to perform event-driven and time-constrained work, particularly with regard to fault and configuration management. This section will first describe how an NMO for a WAN is organized. How new network implementations are managed is described next, followed by a description of the role of information management in network management work. Finally, the five dimensions of network management (above) are discussed in terms of design work and RTSC work.

The NMO for a large network, such as a national-scale WAN, includes a staff of technical specialists who provide depth of knowledge in particular technologies, the RTSC staff, budget, accounting, and administrative support staff, and executive management. Categories of technical specialists include network hardware engineers, network software engineers, data security specialists, and circuit/topology experts. Network hardware engineers are specialists in network equipment, signaling equipment, electrical requirements, lower-level network protocols, and troubleshooting. The most skilled network hardware engineers will have degrees in computer or electrical engineering. Network software engineers are more likely to be trained in computer science, and will have expertise in higher level network protocols, software debugging, software development, and software engineering processes (e.g., the waterfall systems development life cycle discussed below). Circuit/topology experts manage the leased long-distance telecommunications circuits and the relationships with the vendor or vendors from whom they are leased, and are responsible for managing the network topology. Security specialists may have backgrounds in either engineering or computer science, but will focus on authentication, authorization, dedicated security hardware (e.g., link encryptors and firewalls) and encryption key management processes. In large NMOs, the security experts will be in a separate group and report to a different manager than do the hardware and software engineers in order to maintain arms-length relationships and help ensure that security procedures and standards are not compromised by expediency (e.g., cost or schedule pressures). In total, these technical experts are referred to here as “members of the technical staff.”

The RTSC staff is organized into three shifts, with shift supervisors who report to the supervisory control manager. The RTSC staff are the people who work most directly and most often with the network itself, monitoring the network’s health and configuration (fault and configuration

management). The RTSC staff also act as the human interface between the network and the NMO's internal and external customers. The RTSC staff work in one or more centralized network control rooms. RTSC staff backgrounds are likely to vary from some who have university degrees in a variety of technical and nontechnical fields, to those who have experience in real-time operations but little or no postsecondary education.

The budget and accounting staff rely upon information generated by and about the network to manage billing, budgeting, and financial planning. The executive managers often have backgrounds that include a degree in a technical field and experience as a member of the technical staff. As managers, they will have a well-developed model of the technology and the ability to perform cross-community coordination with the various NMO groups (members of the technical staff, RTSC staff, accountants), internal and external customers, and vendors.

Network design and software engineering processes have many parallels, and can be described by well-known life cycle models such as the "waterfall" model of software development (Figure 11.1). In both situations, the goal is to create a new object: the network or piece of software. The life cycle begins with the identification and documentation of *requirements*. Given a set of relatively stable requirements, one or more *design* alternatives can be produced to satisfy the requirements. Given agreement on a design, *implementation* of the object can begin. As the object is created, each discrete piece undergoes *testing*, and the object is ultimately tested as a whole. After successful testing, the object is released and goes into *maintenance* (software) or goes into *production operation* (network). The life cycle is repeated as new requirements emerge. Each of these subsequent iterations can be managed using the same life cycle model: new requirements are documented, a network design revision is created, and the changes are implemented, tested, and made operational. The NMO is responsible for managing the entire implementation/revision life cycle for its network, and the cycle can be reiterated many times throughout the useful life of the network. Typical types of new requirements for a network include expanding the network into new locations, increasing capacity, or adding support for new data transmission protocols.

In order to design a new network (or modify an existing network), requirements must be gathered to inform the network design (network locations, traffic patterns and types, security and reliability requirements, service level requirements, cost parameters, etc.). During the network design process, engineers use a combination of formal, algorithmic, and professional heuristic processes. For example, given a certain expected traffic load, algorithmic tools can be used to define the number, arrangement, and sizes of routers and the circuits connecting them. The arrangement of routers and circuits is referred to as the network topology. WANs are never designed as a completely interconnected mesh (i.e., each site linked directly to each other site, Figure 11.2), so the topology must be

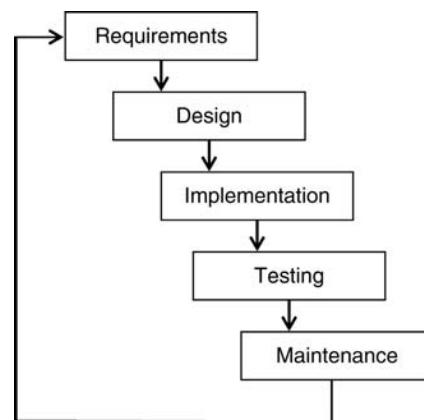


FIGURE 11.1 The "waterfall" systems development life cycle.

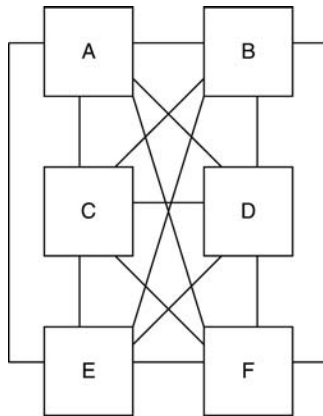


FIGURE 11.2 Full mesh topology, where A, B, C, D, E, and F are network sites.

carefully designed to provide adequate support for the traffic patterns and service-level expectations while minimizing operating costs: long-distance circuits between cities are often leased from a telecommunications provider and have recurring monthly costs (Figure 11.3). The set of equipment and services needed can then be acquired, installed, and tested. Any moderately complex network is, when built, an ecosystem comprised of a wide variety of equipment and services, including cables, equipment racks, signaling devices (e.g., encryptors and digital “modems”), routers, and network management servers and workstations. Each type of equipment (e.g., routers) can be thought of as a species within the ecosystem. Engineers plan and manage installation and testing of hardware and services at each site (the routers, link signaling devices, switches, hubs, racks, long-distance telecommunications circuits, wired and wireless access points, and cables). Software engineers define and test connection parameters and system options within routers, link signaling devices, and any other software-configurable network components. RTSC staff monitor and control the systems as they are installed and tested, perform routine tasks (on both a daily and per-shift basis) and respond to faults. Documentation, policies, and SOPs are developed and refined throughout the planning, implementation, and testing phases. Customer endpoints (LANs, hosts, servers, and applications) are connected to the network, tested, and the new network is in production operation and maintenance mode. Following initial implementation of a network, configuration change tends to be

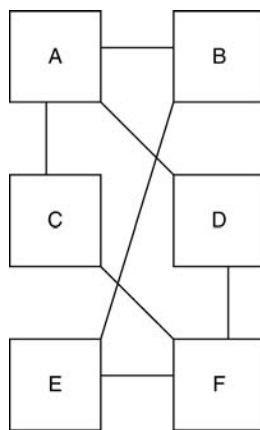


FIGURE 11.3 Optimized network topology, where A, B, C, D, E, and F are network sites.

gradual, to replace failed equipment, replace a species (e.g., routers) with an upgraded version, or to make modifications in response to changing user or organization needs. Whether initial installation or incremental upgrade, network reconfigurations include both physical changes (replacing a hardware component), software changes (modifying software parameters related to the component, also known as provisioning), and changes to the documentation.^[7]

Management of distributed infrastructure, including network management, is information-intensive work. Each kind of equipment included in the network is managed by its own workstation-based suite of applications tailored to manage that layer of technology. Each of these systems contains a database (the Management Information Base, or MIB) that provides a representation of the configuration of that layer of the network, providing an accurate, one-to-one correspondence to the physical network configuration. Information in each MIB is, however, isolated and few tools exist to provide effectively merged views of data from all layers of the network. SOPs are created, using an iterative process of proposal, negotiation, and testing, in order to define preplanned and vetted operational sequences for use in routine and nonroutine but anticipated situations. SOPs are organization and context specific, reflecting organizational values and priorities as well as the combinations of technology that make each WAN unique (e.g., kinds of equipment and locations). Some (SOPs) are so commonly used that the RTSC staff has them memorized and the paper documents are only used by newcomers or as referents during discussion. Other procedures, like disaster recovery procedures, are hopefully never actually used. Versions of these types of procedures are often modified into test scripts, which are used once and then discarded.

The NMO staff creates formal databases and other information management tools to serve the specific needs of the NMO, such as trouble ticket, configuration management, and test scheduling systems. Other formal information systems include the vendor provided MIBs and the NMO's own SOPs. Other more situated and less formal information systems included are handwritten notes and collections of information assembled by individuals, created spontaneously by the NMO staff in order to allow them to do their work more effectively. The informal systems often combine information drawn from one or more formal information sources. The combination of information from multiple sources results in new information compounds^[8] that highlight, through processes of extraction, juxtaposition, and integration of domain semantics, what is salient in a domain rich with information.^[9] People from all parts of the organization spend a great deal of time seeking and gathering bits of formal data, adding the organization-specific stuff of meaning and context, and producing useful, but often difficult to find and manage collections of information. These informal information compounds are often expressed in tabular format in order to represent relationships between physical and logical objects. The relationships of logical objects (like the symbolic or mnemonic name of a network device) to physical objects (like a port on a router or a cable) are enormously important during troubleshooting or in planning for network changes. These information compounds are a means by which the members of the NMO can introduce the organization-specific data about customers, purpose, and meaning into both the real-time supervisory control and design environments. However, maintaining the integrity of these informal representations over time is difficult due to the high rate at which changes are made to the network configuration.

Schematic drawings of relationships between network devices are another common information compound useful in highlighting important relationships between discrete elements in the network. These, too, are prone to problems caused by change and poor document maintenance.

PATTERNS OF WORK AND DIMENSIONS OF NETWORK MANAGEMENT

The degree to which members of the technical staff, who are primarily responsible for network design work, and the RTSC staff, who are responsible for monitoring and controlling the network in real-time, share and coordinate the work associated with the five dimensions of network management varies by dimension, and is influenced by the values, priorities, and other characteristics of the NMO. The work patterns described below are based upon network management work as practiced

in a large financial institution,^[10] but are typical of any NMO operating a WAN in a context where high security, reliability, and availability are expected.

In terms of fault management, members of the technical staff will be engaged in the requirements definition and implementation of equipment and software related to fault management. Software engineers may develop or customize systems for managing alarm data in either real-time or for fault diagnosis or post hoc fault analysis. Members of the technical staff do not participate routinely in fault monitoring activities: this is the responsibility of the RTSC staff. However, under certain conditions, such as during real-time troubleshooting, members of the technical staff merge with the RTSC staff—usually joining together in the network control room—to isolate and resolve a difficult problem as quickly as possible.

Faults requiring intervention occur frequently in large networks, often because of the complex hardware and software configurations located at each network site or because of problems with the telecommunications circuits that connect network locations to each other. An NMO creates procedures (SOPs) that define the canonical responses to typical fault situations. RTSC staff use the SOP as a reference when responding to typical faults, which should result in rapid service recovery performed in a standardized manner. If the standard response does not work, the RTSC staff moves into troubleshooting mode in order to restore service to affected users (by perhaps reconfiguring the network so that redundant or alternate components can be used by the affected users), isolate and characterize the outage, document the outage in a database (by creating a trouble ticket that serves as a collection point for alarms, actions taken, actors involved; a trouble ticket created in response to nonroutine events becomes a narrative of the NMOs collective actions), and escalate the situation (i.e., notify management and next-level technical support, including hardware and software engineers who are responsible for network design and are typically expert in relatively narrow areas such as circuits, routers, or security; or possibly equipment or service providers).

Configuration management is primarily performed by members of the technical staff using a reflective, self-paced, and creative design process, although the RTSC staff are involved for the implementation, testing, and ongoing maintenance. A standard implementation process, such as the waterfall process described above, is typically used to manage preplanned, permanent configuration changes. The RTSC staff may make temporary configuration changes in response to network faults, or as part of the execution of a SOP (e.g., see the description of a DRT below).

Performance management is almost entirely done by members of the technical staff in a reflective process where data collected by the network is compiled and analyzed. In rare cases, real-time network congestion may bring RTSC and design staff together in the control room for real-time analysis and intervention. These rare real-time situations are likely to cause a network capacity review by members of the technical staff.

Accounting management is also almost entirely a post hoc, off-line activity. RTSC have little to do other than ensure that the accounting data is being collected continuously. Accounting data is usually important to the NMO even in cases where detailed usage tracking is not done; usage information is still likely to be used to demonstrate NMO and network relevance as well as to help apportion network costs to customer organizations.

Security management involves members of the technical staff most directly and includes managing access control, user accounts and passwords, and encryption key management procedures. The RTSC staff does get involved, particularly in diagnosing suspected link encryption or data encryption problems. In a 24-hour, high-security environment, additional control procedures are required to manage access to “superuser” passwords that may need to be accessed from a secure location, such as a safe, when required for fault resolution.

INFORMATION AND ACTIVITY IN NETWORK MANAGEMENT

An example of how an NMO plans and conducts a disaster recovery, or business continuity, test (DRT) is indicative of the ways the design and RTSC staffs interact and of the reliance of the

NMO on effective information management. The DRTs in a WAN are complex. The DRT scenario assumes that the entire network site, its servers, and LANs at that site have all been disabled (e.g., due to a fire, flood, etc.). A private network with high standards for availability would set as a network design requirement the ability to survive the complete loss of any network site by providing connectivity at an alternate, backup site. This requirement directs many other decisions and activities. For example, the need to redirect huge flows of data from a failed site to a backup site has an impact on the initial design of the network and its topology, as well as the way recovery mechanisms within the network are customized. The NMO also commits significant human resources from the technical and RTSC staff to create SOPs that will support this sort of recovery within the target times, which require complete network and server recovery within a few hours of the outage. The SOP likely contains scores of steps that are to be performed in a preplanned sequence, often by multiple organizations working in coordination. Such a complex document must be created and reviewed by both RTSC and technical staffs: the members of the technical staff are subject matter experts with deep knowledge of a narrow technical specialty, while the RTSC staff have broad working knowledge of the range of network technologies, but with less technical depth. The RTSC staff's "big picture" view of the network and the organizational environment allows them to contribute information about how activities should be sequenced and coordinated, both within and across the participating sites. People from multiple organizations and sites must be familiar with the procedures and rehearse them regularly. Backup data must be transmitted regularly, often on a continuous transaction-by-transaction basis, or by using daily data transfers. In either case, these data transfers become part of the RTSC staff's ongoing responsibilities.

The planning for a DRT begins a couple of months before the test date. A DRT director, usually a manager, is assigned responsibility for planning and running the test. Draft test scripts are distributed, which are modified versions of the SOP. The modifications reflect the goals and exceptions for the current test. (It is, as a practical matter, nearly impossible to take a major site out of the network for the 18–24 hours it takes to run a disaster simulation and test all of the recovery and fallback procedures. The tests are invariably limited to some degree to accommodate other competing requirements.) Conference calls are held where the site being tested, the NMO, and any other sites involved in recovery of servers "walk through" the script, looking for errors, sequencing problems, negotiating modifications, and inserting proper articulation and control points (e.g., periodic management conferences to review progress and status when the test is executed). Members of the technical staff and RTSC staff are present at all of the planning meetings and participate in the discussions and negotiations. The SOP is designed to serve as a canonical text in that it embodies the actions to be taken in the event that an "ideal" disaster occurs: a complete, total disaster at the site that would render it unusable. As such, the SOP embodies a "best practice" approach to dealing with a disaster, providing guidance on the most reliable way to effect the recovery of each piece of the failed environment at the backup sites. At the start of the test, the disaster itself is simulated, and how the simulation is to be performed is part of the test script. Specific dates and times are included in the test script to provide a rough sense of what it means to be on or behind or ahead of schedule during the test, values which are not included in the canonical SOP.

As the test date nears, it is discussed during the normal change and scheduling meetings to ensure the readiness of the NMO and other direct participants, to allow new information to be integrated into the schedule and plan, and to ensure awareness by a wider range of people and organizations who may be conducting other activities that overlap with the DRT. The day before the test begins, a final meeting is held, attended by representatives from all locations with responsibility for conducting the DRT. This meeting is the last opportunity for cancellations, late modifications to the test script, discussion of resource conflicts, times, etc.

The DRT begins a couple of hours before dawn, and RTSC staff from the NMO are responsible for many of the first steps to reconfigure the network to support movement of servers and other systems from the site under test to various recovery sites at other locations across the network. The third shift NMO staff cannot participate in daytime planning meetings that have occurred during

the past week, and so they rely upon the written test schedules and scripts, plus any information passed verbally to them by the previous shift staff. They are also likely to have other network reconstructions to perform during the same time period, near the end of their shift. If this part of the day goes well, the NMO will be in good shape for the entire weekend. If not, the political fallout could last for weeks. The site disaster will be simulated by disabling all of the network connections from other network sites to the site under test, in effect isolating that site from the rest of the network. The NMO RTSC staff issue status commands to ensure that the circuits are inactive. Alarms printers and displays record the disables as well as hundreds of other, cascading events. Network elements, represented by icons, change from green to red on many of the workstations. The NMO RTSC staff contact the test director to report that their current task is completed.

Contingencies are bound to arise in any event this complex. One site may get a late start; there may be problems reading a tape when trying to restore a software image at the backup site. The NMO is also likely to be dealing with normal problems, such as circuit problems that are causing high error rates and affecting network performance. In real time, processes move forward at the many participating sites. Some parts of the process get stuck and require intervention; some succeed; others simply fail, and the test script is modified on the fly in a continuous series of phone calls and negotiations. When the reality of the test day proves problematic, people can refer to the plan and make on-the-spot decisions about how to work around the difficulty: steps may be done in the “wrong” order, certain steps skipped, or certain steps repeated. These variations must be noted at the NMO as they occur, both to keep track of the test status in real time and as information to be used to modify the underlying SOP and subsequent similar test scripts.

The duration the DRT and other tests may span shift boundaries, making it necessary for accurate status to be passed from shift to shift. Managers from the NMO need status and may be on site or at home. The technical staff members supporting this and other simultaneous tests need to be aware of the status of the major test and adjust their weekend activities accordingly. The NMO must coordinate its activities through the DRT manager. As the test runs, problems may occur which must be resolved. Some parts of the test may be omitted in response to unexpected problems to avoid compromising the entire test.

The final event associated with a DRT is the postmortem held the following week. The participants in the postmortem are the same as those who participated in the many pretest meetings. The DRT director in effect grades the performance of the participants. The NMO and other organizations are held responsible for problems arising during the test. Responsibility for resolving each problem is publicly assigned. Resolution of problems assigned to the NMO may involve modifying the SOP or making configuration changes to the network. Even though problems occur during the DRTs, they are, due to the careful preplanning and test-day coordination, almost always considered successful.

CONCLUSION

Network management is not exclusively a technical issue of routers, telecommunications links, servers, applications, and the interconnections of LANs via metropolitan or WANs. Effective network management depends upon the blending of design and RTSC work to support all dimensions of network management: fault, configuration, performance, accounting, and security management. Effective organization and control of complex, dynamic information is also important to the NMO, as it enables the recording and communication of large amounts of information in forms that incorporate the values of the NMO and the larger organization within which it is embedded. While the five dimensions of network management form a common basis for organizing network management across a wide variety of organizations, the details of network management practice are situated and influenced by security requirements, available technical expertise, and budget. The concerns of network management apply as well to other situations in which distributed infrastructure must be managed.

Standards published by the Internet Engineering Task Force (IETF) have come to dominate network management practice. The most important of these are the Simple Network Management Protocol Version 3 (SNMPv3), defined in Requests for Comment 2571–2575.^[11–15] Further information about the technical aspects and best practices of network and infrastructure management, including how SNMPv3 is used in practice, can be found in a number of works, including Limoncelli–Muller.^[16–19]

REFERENCES

1. Panko, R.R. *Business Data Networks and Telecommunications*, 6th Ed.; Pearson Prentice Hall: Upper Saddle River, NJ, 2007.
2. Klerer, S.M. The OSI management architecture: An overview. *Network* **1988**, 2 (2), 20–29.
3. Millman, D. Authentication and Authorization. In *Encyclopedia of Library and Information Sciences*, 3rd Ed.; Taylor & Francis: New York, 2009; 413–419.
4. Jones, P.M. Cooperative work in mission operations: Analysis and implications for computer support. *Comput. Support. Coop. Work Int. J.* **1995**, 3 (2), 103–145.
5. Bucciarelli, L.L. An ethnographic perspective on engineering design. *Des. Stud.* **1988**, 9 (3), 159–168.
6. Schön, D.; Bennett, J. Reflective conversation with materials. In *Bringing Design to Software*; Winograd, T., Ed.; ACM Press: New York, 1996; 171–184.
7. Wang, H.H. *Telecommunications Network Management*; McGraw-Hill: New York, 1999.
8. Paepcke, A. Information needs in technical work settings and their implications for the design of computer tools. *Comput. Support. Coop. Work J. Collab. Comp.* **1996**, 5 (1), 63–92.
9. Goodwin, C. Professional vision. *Am. Anthropol.* **1994**, 96 (3), 606–633.
10. Sandusky, R.J. Infrastructure management as cooperative work: Implications for systems design. *Comput. Support. Coop. Work J. Collab. Comput.* **2003**, 12 (1), 97–122.
11. Harrington, D.; Presuhn, R.; Wijnen, B. *An Architecture for Describing SNMP Management Frameworks*, RFC 2571; Internet Engineering Task Force, 1999, <http://www.ietf.org/rfc/rfc2571.txt> (accessed August 7, 2008).
12. Case, J.; Harrington, D.; Presuhn, R.; Wijnen, B. *Message Processing and Dispatching for the Simple Network Management Protocol*, RFC 2572; Internet Engineering Task Force, 1999, <http://www.ietf.org/rfc/rfc2572.txt> (accessed August 7, 2008).
13. Levi, D.; Meyer, P.; Stewart, B. *SNMP Applications*, RFC 2573; Internet Engineering Task Force, 1999, <http://www.ietf.org/rfc/rfc2573.txt> (accessed August 7, 2008).
14. Blumenthal, U.; Wijnen, B. *User-based Security Model (USM) for Version 3 of the Simple Network Management Protocol (SNMPv3)*, RFC 2574; Internet Engineering Task Force, 1999, <http://www.ietf.org/rfc/rfc2574.txt> (accessed August 7, 2008).
15. Wijnen, B.; Presuhn, R.; McCloghrie, K. *View-based Access Control Model (VACM) for the Simple Network Management Protocol (SNMP)*, RFC 2575; Internet Engineering Task Force, 1999, <http://www.ietf.org/rfc/rfc2575.txt> (accessed August 7, 2008).
16. Limoncelli, T.A.; Hogan, C.J.; Chalup, S.R. *The Practice of System and Network Administration*, 2nd Ed.; Addison-Wesley: Upper Saddle River, NJ, 2007.
17. Burke, J.R. *Network Management: Concepts and Practice, A Hands-on Approach*; Pearson Education: Upper Saddle River, NJ, 2004.
18. Morris, S.B. *Network Management, MIBs and MPLS*; Prentice Hall: Upper Saddle River, NJ, 2003.
19. Muller, N.J. *LANs to WANs: The Complete Management Guide*; Artech House: Boston, 2003.

13 Search Engine Optimization

Nicholas Carroll

CONTENTS

Introduction.....	172
Definition of Higher Rankings.....	173
Origins and History of SEO.....	173
Practitioners and Software.....	174
Early Search Engine Promoters and Reporters.....	174
The Current SEO Industry.....	174
SEO Software (Automating the Process).....	174
SEO Worldwide.....	175
Black Hat vs. White Hat SEO.....	175
SEO Methods.....	175
Elements of a Web Page to Be Optimized.....	176
Choosing Keywords to Target—The Core of SEO.....	176
Content.....	178
Density.....	178
Placement.....	179
Inclusion of Related Keywords.....	179
Consistency.....	179
Technical.....	180
Keyword Loading, Stuffing, and Spamming.....	180
The Meta Keywords Tag.....	180
Entry Pages.....	180
Cloaking.....	181
Linking.....	181
Link-Weighting Fundamentals.....	181
Strategies for Leveraging Links.....	182
Web Site and Web Page Structure.....	183
Web Analytics.....	183
Combining Content, Technical, and Linking SEO.....	184
Broad Issues in SEO.....	184
Building Downwards vs. Outwards.....	184
Broad vs. Narrow Targeting and Long-Tail Terms.....	184
Balancing Targeting and Serendipity.....	185
Lead Time and Longevity.....	185
Geographic Targeting.....	186
SEO Source Tracking.....	186
Increased Competition.....	186
Search Engines Only Index Words.....	186
Trends.....	187
Fading Trends.....	187
Current Trends.....	187

Trends Affecting SEO	188
New Search Engine Presentation Methods	188
Specialized Searches	188
Alternate Search Channels	189
Standards and Regulation in SEO	190
HTML	190
Certification.....	190
SEO Industry	190
Search Engines.....	190
Government Regulation	191
User Behaviors in Search.....	191
Basic Behaviors.....	191
Searching Popular Subjects.....	192
Use of Boolean Syntax.....	192
Search Engine Loyalty	192
Hardened User Behaviors.....	192
User Behavior Worldwide.....	192
Conclusion	193
Acknowledgments.....	193
References.....	193
Bibliography	195
Users Section	195

INTRODUCTION

Search engine optimization (SEO) is the craft of elevating Web sites or individual Web site pages to higher rankings on World Wide Web (WWW) search engines through programming, marketing, or content acumen. The definition often includes specifications for increased traffic to a given Web site, improved quality of traffic, increased profits, or brand awareness.

In a typical online search on a topic, anywhere from hundreds to millions of articles may be extracted and ranked by probability of relevance on a search engine results page (SERP). From the earliest appearance of Web search engines, creators of Web sites realized that ranking high was vital for visits to their Web sites, since users seldom look on even second or third screens of rankings, let alone the hundredth screen. The numbers of Web sites on the Web increased so rapidly from the beginning that the competition for the most desirable top positions developed rapidly.

In this entry, we move through several key aspects of SEO. First, we define “rankings,” and review the origins and history of SEO. Then we discuss what is known as “Black Hat” and “White Hat” SEO, as the competition for rankings can lead to distortions and misleading rankings, depending on what techniques are used.

The heart of the entry concerns SEO methods. We review keyword selection and then attend to various aspects of the content of the Web pages, such as keyword density, placement, and consistency. We then go on to the technical aspects and review techniques such as keyword stuffing, meta tags, and cloaking, as well as strategies for boosting page rank through links from and to other Web sites, with brief discussion of combined strategies and the analytic methods applied to track results.

The entry expands into more general issues such as broad vs. narrow targeting, page longevity and rank, geographic targeting, and increasing competition. Trends in SEO practices are described from early to present time, as are external trends affecting SEO.

We conclude with a description of SEO regulation, mostly in the form of search engines policing their listings, and an overview of typical behavior patterns of search engine users.

All the methods described have been or continue to be effective SEO to some degree. Changes in search engine indexing protocols as well as their methods for displaying search results mean that

no strict description of the best method or methods can remain entirely accurate indefinitely. This entry does not cover pay-per-click, other forms of online advertising, or the resale of Web site traffic or links.

Because the discipline originated in the mid-1990s, terminology is still in flux. General references at the end of this entry point to the more authoritative Web sites and definitions.

DEFINITION OF HIGHER RANKINGS

“Higher rankings” in the popular press or lay discussion generally equates with the goal of having a Web page appear in the first 10 or 20 search records for a particular search term, as it is broadly established by tracking that few search engine users will click through to any link beyond the 20th record. SEO professionals usually discuss rankings in terms of SERP position (Figure 13.1).

Before search engines accepted paid advertisements, SEO was considered a unique form of promotion, radically different from all traditional forms of advertising. It is now more often considered a subset of search engine marketing (SEM), and is sometimes referred to as “organic search” or “natural SEM,” as opposed to paid advertisements placed on the pages of search engines or their affiliates.

Online commerce was the originating force behind search engine promotion and remains the primary driving force behind SEO. Nonprofits and government bodies apply some SEO methods but tend to rely on their unique identity to assure them a prominent SERP position, e.g., the Red Cross, Amnesty International, the Vatican, the New York Department of Motor Vehicles, or the Peoria Public Library.

ORIGINS AND HISTORY OF SEO

The term “search engine optimization” came into popular use in 1997–1998 and is frequently attributed to Danny Sullivan,^[1] then operating “Search Engine Watch,” though Sullivan states he is uncertain who coined it. Previous terms included “search engine placement,” “search engine ranking,” “search engine positioning,” and “search engine promotion,” the latter attributed to Jim Rhodes,^[2]

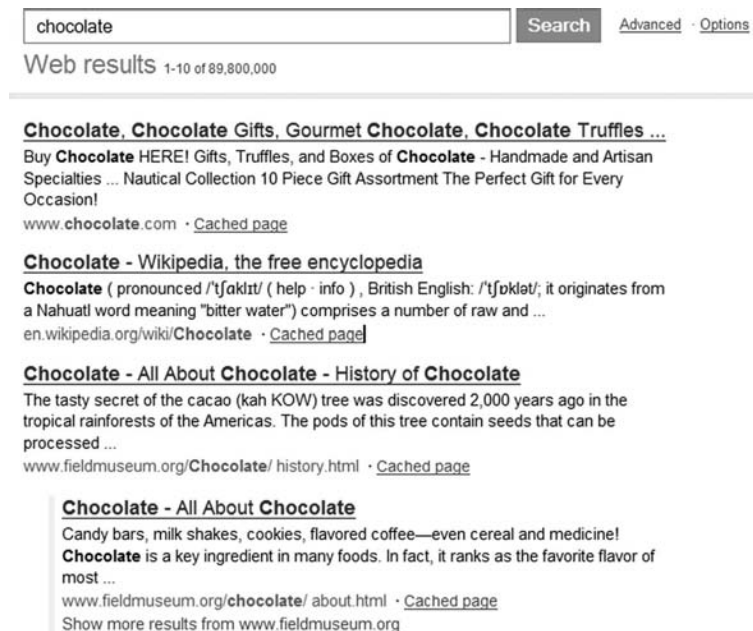


FIGURE 13.1 Typical SERP in search for chocolate, showing the first four results.

author of “The Art of Search Engine Promotion.” Predecessors such as Heath^[3] in his 1995 article “Pointers on how to create business Web sites that work” did not have a formal name for SEO.

The period from the mid-1990s to about 2000 was characterized by broad experimentation on the part of both search engines seeking a business model and Web site creators intent on promoting themselves. Search engines were relatively under-powered and minimally staffed; their primary focus was on keeping pace with the growth in new Web sites. SEO quickly became part of the American “wild, wild Web” metaphor, with more active Web site owners engaging in a huge variety of methods to gain higher rankings, as described in “SEO Methods.” In a period when many multinational organizations did not have Web sites at all, smaller and more nimble organizations and individuals aggressively practiced SEO to establish a beachhead on the WWW.

The early days of search engines were in some ways a struggle against pornography. Visitors who started using search engines after 2000 have little conception of how pervasive the online sex industry was at one time, with their records appearing among search results for cooking, art, quilting, travel, and other innocuous subjects.

As search engines have grown and become more sophisticated, and the number of Web sites has increased more than tenfold, aggressive maneuvering to rank well in the SERPs has to some extent given way to a focus on ranking well in particular niches, and executing well on fundamentals rather than exercising brilliance in manipulating search engines.^[4]

PRACTITIONERS AND SOFTWARE

EARLY SEARCH ENGINE PROMOTERS AND REPORTERS

Early SEO innovation was an individualistic endeavor, primarily developed by small-to-midsize businesses ranging from small hotels to makers of custom sports equipment. Several such niche and popular products were financially successful, demonstrating that SEO was a path to profit.

Web sites with forums for exchange of SEO strategies began to appear by 1996, including still-existing Web sites such as <http://virtualpromote.com>, <http://www.searchengineforums.com>, and the archived <http://www.deadlock.com/promote>.

Web sites reporting on search engines and optimization began about the same time, including <http://www.wilsonweb.com>, <http://www.searchenginewatch.com>, and <http://www.webmasterworld.com>. More recent additions include <http://www.seobook.com/blog>, <http://www.searchengineland.com>, and <http://www.toprankblog.com/search-marketing-blogs>.

THE CURRENT SEO INDUSTRY

Whether offered as a subset of SEM services or sold alone as organic search, SEO has become a niche industry with its own sales forces and conventions. The service may be sold as a one-time audit or on an ongoing basis with monthly billing and performance reporting.

Pure SEO consulting firms now number in the thousands, primarily concentrated in North America and the United Kingdom. Tens of thousands more Web designers and developers offer the service as ancillary to building Web sites, and an unknown number of Webmasters and Web site owners apply SEO methods to their own Web sites.

Gross expenditures on SEO in 2006 were estimated at over USD one billion, with steady annual growth anticipated. This figure describes organic search and does not include paid advertising or most in-house work in smaller companies.^[5]

SEO SOFTWARE (AUTOMATING THE PROCESS)

Automated Web page submission tools came into being in the mid-1990s soon after search engines began indexing the WWW. For some years, they were a vital part of SEO strategy for large Web

sites, as manually entering thousands of URLs into the submission pages of a dozen search engines was extremely time-consuming. Submission tools fell into disuse as search engines became more adept at “deep crawling” (following links into the lower levels of Web sites’ file hierarchies). The ones currently in use typically offer additional features such as help in choosing keywords and automated reporting on SERP positions.

Web analytics tools move beyond automated reporting to features such as tracking visitors’ paths through Web sites, and integrating the data with financials.^[6]

Most content management systems, originally focused on intranets, now include management features for Web sites, and many are programmed to allow a significant modification of variables important to SEO.^[7] Likewise most blogging software includes features that encourage users to add keywords to the title tags, the Web page body, and the filename of a given Web page.

SEO WORLDWIDE

While search engines are now used worldwide, most SEO reporting is about the United States, Canada, and the United Kingdom. Activity in other countries must be inferred. Australia, Ireland, and Russia in particular have significant indications of entrepreneurial SEO, including blogs on the subject.^[8]

BLACK HAT VS. WHITE HAT SEO

“White Hat” SEO typically refers to strategies and tactics that are in concordance with the policies of online search engines, in a loose tacit agreement to provide Web surfers with “relevant content.” “Black Hat” SEO describes tactics that ignore generally accepted conventions of ethical WWW behavior to advance an agenda or commercial interest. The focus on search engine policies before other interests flows from the pervasive position of search engines, at the mid-point in the process of aggregating Web page data and delivering it to users. The tilted point of view is buttressed by the oligopoly of major search engines, with a small number of them processing the vast majority of searches, while millions of Web sites vie for high rankings.

Technically sophisticated tactics are often equated with Black Hat SEO, and a focus on high-quality content with White Hat SEO. However, there is substantial overlap; few if any tactics can be inherently classified as good or evil. From the perspective of a user, the main criteria of legitimate SEO is whether a given search return is relevant to their interests, regardless of how it achieved its ranking. Most experienced SEO practitioners consider intent to be the defining factor.

Typically, methods considered Black Hat SEO develop high page rankings faster, while those known as White Hat SEO tend to create longer-lasting rankings. A preference for one approach or the other is not entirely a matter of the SEO practitioner’s personal preferences; it also depends on the business model. Web sites or pages promoting products and services with short life cycles are suited to Black Hat methods, as they do not suffer from being “burned,” meaning they have drawn the attention of search engine administrators and been banned from that search engine’s index entirely. (This has also happened to sites of long-view organizations such as auto manufacturers when SEO subcontractors acted with an excess of zeal.)

White Hat SEO is better suited to Web sites that offer products with long life cycles. It is also suited to academic or government Web sites, where there is likely to be a consistent focus over decades—such Web sites often gain high and enduring SERP rankings simply by publishing high-quality content about a particular subject.

SEO METHODS

In describing SEO it is necessary to distinguish between automated search engines and manually created directories such as Yahoo! Directory (<http://dir.yahoo.com>) or the Open Directory Project (ODP) (<http://www.dmoz.org>), where listings are added by human editors rather than automated protocols.

SEO describes strategies and tactics for influencing page rank on search engines that use “robots” (<http://www.robotstxt.org/wc/faq.html>) or “spiders” to “crawl” Web pages, traveling from page to page through hyperlinks, and indexing those pages by algorithms and protocols. While directories such as the original Yahoo! Directory can be searched from within Yahoo!, and the ODP pages can frequently be found in SERPs, it is the manner of creating the index that differentiates search engines and directories.

ELEMENTS OF A WEB PAGE TO BE OPTIMIZED

The following example of HTML (Hypertext Markup Language) is referred to throughout this entry. The markup language and page structure used are explained in Table 13.1 below the example. (Web page structure can be viewed in almost all Web browsers by selecting the “View > Source” or “View > Page Source” option.)^[9]

```
<html>
<head>
<title>ELIS—Encyclopedia of Library and Information Sciences 2008-2009</title>
<meta name="description" content="The definitive reference to library and information sciences.">
<meta name="keywords" content="encyclopedia, marcia, bates, mary, niles, maack, mack, optimise">
</head>
<body>
<h1>ELIS—Encyclopedia of Library and Information Sciences</h1>

<p>Body content written for SEO is rich in keywords and also readable. Proper nouns such as "encyclopedia" are used instead of pronouns. Initializations such as "SEO" are spelled out as "search engine optimization," both to include alternate search terms and to increase keyword density in the Web page. To target both sophisticated and lay searchers, common terms such as "acronym" are used in addition to precise terms such as "initialization." Concepts are described by all likely variants, such as "SERP," "search engine position," or "ranking."</p>
<h2>Further Encyclopedia Resources</h2>
<p>See the <a href="http://informationr.net/ir/12-4/colis/colis29.html">encyclopedia description</a> for further information on ELIS.</p>
<p><font size="-1">Keywords: encyclopaedia, taylor, francis, marcia, bates, mary, niles, maack, mack, optimise</font></p>
</body>
</html>
```

CHOOSING KEYWORDS TO TARGET—THE CORE OF SEO

Regardless of preferred methods, White Hat or Black Hat, content or technical, the most critical part of effective SEO is deciding which keywords to target for high rankings. While search engines continually work to improve their ranking algorithms, they are not clairvoyant; targeting the right keywords is still the foundation of reaching the right audience.^[10]

A self-centric viewpoint will usually lead an organization to aim at high rankings for its own name, services, agendas, or products. For example, the Raffles Hotel or the band U2 might focus on targeting searches for their own name, in a reasonable belief that most potential customers are specifically seeking them, rather than hotels in Singapore or rock music in general.

TABLE 13.1**Tags and Text in a Simple Web Page****Tags and Text in a Simple Web Page, as Seen Through an HTML Editing Tool**

```

<html>

<head>
<title>ELIS: Encyclopedia of Library
and Information Sciences</title>

<meta name="description" content="The
definitive reference to library and
information sciences.">

<meta name="keywords"
content="encyclopedia, marcia, bates,
mary, nils, maack, mack, optimise">
</head>
<body>
<h1>ELIS: Encyclopedia of Library and
Information Sciences</h1>



<p>Body content written for SEO is
rich in keywords. . . [etc.]</p>
<h2>Resources</h2>

<p>See the <a href="http://
informationr.net/ir/12-4/colis/
colis29.html">encyclopedia
description</a> for further
information on ELIS.</p>
<p>Keywords: encyclopaedia, taylor,
francis, marcia, bates, mary, nils,
maack, mack, optimise</p>

</body>
</html>

```

Explanation of HTML Tags and Structure

Flag to user's browser: Begin parsing for HyperText Markup Language

Begin the head content (which is not visible in a browser)

The page title is the single most important part of Web page text for search engine ranking. While not shown in the body text in a Web browser, it can be seen in the title bar at the top, and is shown as the title of the Web page in almost all search engine listings

Most search engines will use the meta description as the default description in SERPs. SEO practitioners consider this tag a minor influence on SERP position, and usually orient the text towards readers rather than search engines

The meta keywords tag, originally intended as a field for general metadata, was widely abused by SEO practitioners and is now a minor or irrelevant element of SEO

End of the Head section

Begin the text visible in a browser

Headings tags, from H1 down to H6, are the equivalent of chapter and section headings in books, and are usually considered influential in SEO

IMG tags define what image to display as part of a Web page. The "alt" attribute is used to store a brief description, typically keywords

The paragraph tag is the fundamental "container" for body content

The second-level heading tag may have influence on SERP position, as well as significance for metadata and document readability

An outbound link where the "anchor text" contains the relevant keyword "encyclopedia"

This is an unsightly but often necessary addition to the body content, displaying related keywords in the body text, where the search engines will read and index them. This practice evolved when search engines stopped indexing meta keywords

End of the body element

Flag to user's browser: end of HTML file

Further information on HTML is available at <http://www.w3.org/MarkUp/Guide/>.

While this may be a successful strategy for broadly known organizations or people, it is considered poor-quality SEO when an unknown product, service, or agenda is being advanced. Products and agendas without name recognition are better served by targeting a generic search term such as "stainless steel ball bearings" than a term like "Smith Ball Bearing Company." A multinational company such as General Electric or Hitachi—that manufactures a huge variety of products—might likewise target searches for the products rather than its own company name.

Regardless of whether the focus is on the organization or the products, SEO can fail when names are chosen without forethought to online search, particularly when they compete with long-established names. Two examples:

1. Any product named “Guardian” must compete in the SERPs with dozens of well-established newspapers throughout the English-speaking world.
2. Organizations using initializations that might be unique in a local telephone directory frequently face obscurity when competing for recognition in a global medium—“ABC” is not only an initialization for the American Broadcasting Company, but hundreds or thousands of other organizations throughout nations that use the Roman alphabet.

This leads to a fundamental truth about the limits of SEO as a promotional avenue: it can only succeed if people are searching for relevant keywords. If an idea or product is beyond the public’s conception, it cannot be promoted through search engines—in contrast to promotion through traditional untargeted media such as radio, TV, or print. This usually makes SEO a poor method for promoting radically new ideas. For example, “anti-gravity belts” is a search only used by 10-year-old boys.

When possible, experienced SEO practitioners perform keyword analysis before a Web site is built, expanded, or redesigned, using paid or free online tools such as <http://www.keyworddiscovery.com>, <http://www.wordtracker.com>, <http://inventory.overture.com/d/searchinventory/suggestion/> or <https://adwords.google.com/select/KeywordToolExternal> to view aggregated searches conducted on search engines.

CONTENT

At its simplest, the content-focused approach to SEO is to tell the story thoroughly and precisely, in hopes that the words (content) will find a fit with search engine indexing algorithms, and rise in the rankings for particular search terms. The presumption in this strategy is that the writer’s words will also find harmony with the terms used by searchers, and if the writer does have the same interests—and frame of reference—as the target audience, this by itself can result in highly successful SEO. (Conversely, Web pages written in ignorance of either subject matter or audience may achieve high SERP positions, but not necessarily for the keywords that draw the desired audience.)^[1]

A highly content-oriented strategy is often seen in Web sites that were conceived and designed by an individual or small group. Content-oriented Web sites or sub-sites may also come into being in an organization with no SEO strategy at all, such as a department of a university, where the authors are discussing the same or related subjects.

Density

Keyword density receives attention because it is easy to understand and simple to calculate through online density-checkers such as <http://www.keyworddensity.com>. (Such tools are easily found through search engines with the term “keyword density analyzer.”)

Though the effectiveness of intentionally repeating keywords throughout a Web page is debated by SEO practitioners, a search for highly promoted products such as “loans,” “chocolate,” or “discount shoes” will return numerous high-ranking Web pages with the core keyword repeated dozens of times, and keyword densities reaching up to 7% or higher of the page’s content. (However, exceptionally high-density pages can suddenly plummet in the SERPs or be dropped from the index entirely if search engines classify them as examples of “keyword spamming,” described in the Technical strategies section below.)

Conversely, using a set of keywords only once in a Web page usually results in SERP obscurity, even if the set is extremely precise. For example, even “dodo bird nesting” returns tens of thousands of Web pages, and of the first few dozen search engine records, almost all will have the three keywords used several times.

Many SEO practitioners and analysts favor well-written content over keyword density, as search engines become better at differentiating between conventional prose and deliberately enhanced Web pages. This school of thought says that content should be primarily aimed at humans, not search engines, and that good writing not only entices users to read and act, but is better search engine bait in the first place. The belief is often expressed as “Content is king.”

Placement

The placement of keywords in a Web page is a significant factor in SEO, since search engine algorithms accord more weight to certain positions, whether visible or within tags (see Table 13.1). Notable areas for placement include

Visible in Web page

Inside heading tags <h1, h2, h3. . . .>

Placed at the beginning (top) of a Web page.

Within outbound links, i.e., “anchor text” (example in Table 13.1).

Placed NEAR other keywords (generally within 10 words).

Visible in SERPs and browser title bar

Inside the HTML <title> tag.

As part of the URL

Competition over search terms and SERP positions has expanded the use of keywords to the point that they are now routinely made part of file names (e.g., <http://www.mysite.com/used-software-for-sale.html>), or may even form the domain name itself, e.g., <http://www.usedsoftwareforsale.com>.

The title tag is almost universally agreed to be highly important by SEO practitioners. Most other placements are debated. The <meta=“keywords”> tag is largely considered of no importance. The “alt” content in image tags is considered significant in image SERPs, less so in the overall ranking of a Web page.^[12]

Inclusion of Related Keywords

Due to the unpredictability of user search behavior, many SEO practitioners include synonyms, cognates, tangential terms, and misspellings in Web pages. In the mid-1990s, SEO practitioners placed such related keywords in the meta keywords field (see Table 13.1), but massive abuse and misuse of that field led virtually all search engines to stop indexing meta keywords. As a result, SEO practitioners reluctantly turned to placing them as a visible footer on the Web page, though this compromises the esthetics of page design. Some examples of keyword (or “tags”) line listings:

Synonyms, cognates, and closely related terms for a medical page about trigeminal neuralgia, formerly called tic douloureux:

“Keywords: tic douloureux, neuropathy, facial pain, face pain.”

Tangential terms (statistical outliers) for a site on vegetarian diet:

“Keywords: low sodium, low fat, organic.”

Variants on the author’s name for a Web page written by Jon Smythe:

“Keywords: John, Smith, Smyth.”

Common misspellings for a travel Web site about Colombia:

“Keywords: accommodation, Columbia.”

(However, when a keyword tag is used to repeat words already in the body content, it simply becomes “keyword stuffing”—see “Keyword Loading” below.)

Consistency

Building a Web site with a consistent pattern of clearly related keywords is commonly known as creating a “theme.”^[13] A Web site with a clear theme stays on a single topic or closely related topics

throughout its pages (barring such standard pages as “Contact Us”). This strategy builds the overall ranking of the Web site and, as a consequence, the SERP positions of the individual Web pages.

The concept “clearly related” restricts thematic rank-building to keyword relationships that search engines are capable of recognizing. Where “ship” and “marina” might begin to build a theme relating to recreational sailing, search engines would be less likely to recognize a theme in a Web page with the separate keywords “ship” and “space.” Once combined into the more specific keyword “spaceship,” the page becomes identifiable as related to science fiction or space exploration, and if the Web site also contains the words “galaxy” and “parsec,” a theme begins to build.

The interest in themes peaked around 2004, alarmists suggesting that Web sites without a keyword theme would plummet in the SERPs.^[14] However, there are exceptions to the rule: newspapers, encyclopedias, and many news blogs will never have a clear theme, yet they still can reach high SERP positions for a variety of keywords, based on meeting search engine criteria other than consistency of subject matter.

TECHNICAL

A technically oriented SEO strategy emphasizes programming skill or ingenuity over command of language, familiarity with the target audience, or interweaving related content. Tactics vary from the simple, which can be executed by anyone familiar with HTML, to the sophisticated, which require knowledge of programming, Web site servers, or WWW and Internet protocols.

Keyword Loading, Stuffing, and Spamming

Keyword density can be taken to an extreme. Efforts to increase keyword density or variety that involve excessive or awkward repetition are known as “keyword loading.” An example would be replacing every instance of “it” in a page with the noun in question.

When this repetition reaches the level of incoherence, with the same word or words used dozens of times in the body, title, or tags of a Web page, it is usually called “keyword stuffing” (a.k.a. “cramming”), “spamming the index” (“index” referring to the search engines’ databases of Web pages), or “spamdexing.” One of the earliest SEO tactics, it began with simply repeating keywords hundreds of times, generally at the bottom of the page, and frequently with the font color the same as the background color, thus rendering the text invisible to humans. As search engine algorithms began to discount this tactic, keyword spamming evolved into a more precise metering of keyword density. This obsolete technique periodically sees a resurgence as search engine administrators let down their guard.

The Meta Keywords Tag

Located in the head section of a Web page’s HTML (see Table 13.1) and invisible to users viewing a Web page in a browser, the meta keywords tag was created purely for metadata.

In the mid to late 1990s meta keywords were highly popular as a quick path to higher rankings, and indeed stuffing the meta keywords field showed some success when only a few million pages were being indexed. By the time the tactic became broadly known—with “keyword-jacking” lawsuits over copyright and trademark infringement—search engines were on their way to down-ranking the meta keywords tag contents, and usually not indexing the keywords in the field at all.

Entry Pages

“Entry page” is the broadest descriptor for the strategy of creating particular pages to rank well with search engines, sometimes finely tuned to rank well with a particular search engine. “Doorway page” and “gateway page” imply that a page contains some relevant content that has been tuned for high SERP position. “Bridge page” or “jump page” imply that a page may be little more than an uninformative landing page that either urges the visitor to click through to the rest of the Web site, automatically redirects them to the rest of the Web site, or redirects them to an entirely different

Web site.^[15] Most media coverage of such pages classifies them as Black Hat SEO. However, there is no clear agreement on the definition of these pages or their relative level of deception. Many SEO practitioners observe that while bridge and jump pages may be technical and possibly Black Hat as well, all pages visible to search engines are in a real sense entry pages.

Cloaking

“Cloaking” is used to describe a broad range of tactics. The common element is that the human visitor and the search engine spiders “see” different content when visiting a Web page. In the simplest form, the SEO practitioner creates a page tuned for high SERP position, and replaces it with a human-readable page after the page has been indexed by search engines (of course the artificially high ranking only lasts until search engine spiders visit the page again).

Two of the more technical methods, which attempt to deceive search engines on a continuing basis:

1. *User-agent-specific page delivery*, in which the Web page server “sniffs” the incoming page request, extracts the data that identifies what browser the visitor is running (e.g., “googlebot” or “msnbot”), and delivers a special Web page tuned to gain higher SERP on that SE.
2. *IP-specific page delivery*, in which the Web page server delivers a special Web page based on the visitor’s Internet Protocol address (e.g., “127.0.0.1”).

Cloaking is broadly considered one of the most aggressive and sophisticated forms of Black Hat SEO, but search engine analysts point out that even cloaking has legitimate purposes, as when a Web site is migrating to a new domain name; the owners may want to keep the old Web site available to the public until the new one becomes established in SERPs.^[16]

LINKING

Links have become such an important determinant of SERP position that under some circumstances they may override all other elements of ranking, and catapult a Web page of little or no relevance to a top SERP position.

Link-Weighting Fundamentals

Since the mid-1990s, search engines have routinely tracked the links from (and to) Web sites, often assigning a relative value based on the links’ source or destination. Links pointing to a Web site are generally called “inbound links,” and links pointing to other Web sites “outbound links.” (Inbound links are frequently called “backlinks,” though the term has other technical meanings in computing.)^[17]

“Link-weighting” is frequently described as “link popularity” or even “popularity” measurements by popular media. More accurately, search engine link analysis algorithms attempt to infer the value of a Web page or entire Web site based on four factors:

1. The number of inbound links.
2. The “quality” of the inbound links, based on the quality of the originating Web site. Assessing the quality of Web sites is where link-weighting moves into more sophisticated and sometimes arcane mathematics.^[18,19] In simple terms, an inbound link from a major university has more value than a link from an obscure small business, and has a greater beneficial effect on SERP position.
3. The “relevance” of the inbound link. A Web site about a baseball team benefits more by inbound links from other baseball-oriented Web sites than by links from the personal pages of fans.

4. The keywords contained within inbound links (the linked text visible in a browser), known as “anchor text.” The most famous example was the 2003 link manipulation of Google SERPs (known as “Google bombing”) in which hundreds or more Web site owners inserted the link `<a href="http://www.whitehouse.gov/president/gwbbio.html">Miserable Failure</a>` into their pages, causing the U.S. White House biography of George W. Bush to rise to the top ranking for the term “miserable failure.”^[20] (The gwbbio.html page rose to #1 ranking on Google for that term even though it did not contain either of the keywords in question.)

The cumulative importance of these factors in SERP position has led to naming the collective effect of inbound links “link juice” (American “slang definition of” “juice”: influence, clout, <http://www.m-w.com>.)

Web sites with many relevant or high-quality links pointing to them (inbound links) are known as “authority” Web sites; those with many links pointing to other relevant Web sites (outbound links) as “expert” or “hub” Web sites. The presumption built into the search engine algorithms is that a Web site with many inbound links from high-quality Web pages is an authoritative source, and that pages with many outbound links to authority Web sites serve as a WWW resource. (The mathematics can become circular and even self-reinforcing, as when <http://www.wikipedia.org> Web pages briefly started to dominate #1 SERPs positions for thousands of subjects.)

Strategies for Leveraging Links

Link-weighting was implemented to in effect create a weighted voting scheme and take some of the control of rankings away from a Web site’s designer. Since weighting means that not all inbound links are of equal value to SERP position, the methodology also opened the door to SEO opportunities of bewildering complexity, in which Web sites buy, sell, and trade links to and from their Web sites in order to improve their SERP positions.

Always a factor in SEO, “link building” is now broadly considered a core element of strategy, and in some schools of thought the most important element.^[21] Increasing media attention brought linking as a SERP ranking factor to the attention of Web site owners and the quickly growing number of SEO practitioners, and today organizations spend substantial time and effort on link building aimed at high rankings.

Non-reciprocal links (a.k.a. “generosity links”) were the first evolution, in the mid-1990s. At that time Web site creators linked to almost any remotely related Web site, including their competitors. By the late 1990s “reciprocal links” had become the standard, though the exchange was offered with a view to a direct increase in traffic rather than SERP position.

Soliciting inbound links, at first done casually, now often means assigning employees or sub-contractors to solicit links from high-ranking Web sites. The return on investment is questionable—high-ranking Web sites have little to gain by giving an outbound link, and the site owners may be concerned that linking to a low-quality Web site will harm their own rankings. (Soliciting links should not be equated with link building, a broader term that covers all strategies for gaining inbound links.)^[22]

The practice of posting links in discussion groups also originated with the goal of direct traffic rather than influencing SERP positions. “FFAs” (Free For All Web sites), “link farms,” and “link rings” were crude early link-exchange schemes, most variants indiscriminately exchanging thousands of links without regard to relevance. While such sites are still active, their value in SEO has largely been eliminated by increasing sophistication of search engine ranking protocols. Most SEO practitioners today consider participating in them to be useless at best, and at worst, possibly injurious to a Web site’s SERP positions.

“Paid links”—inbound links that have been purchased, sometimes masking as editorial recommendations—existed before the WWW on Internet bulletin boards, always with the intent of publicity or direct financial gain. With the broadening awareness of SEO, using paid links to influence

SERP position has become both a business strategy and a source of contention between Web site owners and search engines.

Good content, the oldest link-building strategy of all, is a somewhat indirect way to build links. Coupled with even a modest amount of self-promotion beyond good SEO—or occasionally just with good content that perfectly targets a popular search—competent writing on a particular subject can generate hundreds or thousands of inbound links to a given Web page, often without any communication at all with the Web sites that are giving the links. Because content-inspired linking may produce results slowly (in months or years), and is often difficult to quantify, few organizations devote serious effort to the method.

WEB SITE AND WEB PAGE STRUCTURE

The directory (folder) structure of an SE-friendly Web site looks similar to a clearly and logically organized hard drive on a personal computer, with the additional proviso that every document is directly or indirectly linked from the home page or some other prominent Web page. Ideally the structure is “shallow” (three or fewer sub-directory levels), to make it easier for search engines to spider, though that has become less important as all major search engines now perform deep crawling.

Actual page structure of HTML pages is in theory dictated by an adherence to W3C standards. In practice, Web sites use almost any markup code that can be rendered by a Web browser and leave the difficulties of indexing to the search engines.

There are drawbacks to unorthodoxy where search engine rankings are concerned. Use of highly irregular Web site structure, page structure, or file naming conventions can seriously harm SERP position. In extreme cases, search engines simply do not add a Web site to their indexes; poor Web site structure can be as destructive to Web site rankings as the most extreme Black Hat tactics, and “site maps” generated specifically to aid search engines in indexing a Web site are not a substitute for logical Web site structure.^[23]

Despite a broad disregard for standards among Web site owners, most SEO practitioners consider disciplined site and page structure fundamental good practice, though these are seen as a foundation for SEO rather than an SEO strategy in themselves.

WEB ANALYTICS

The most thorough SEO practitioners use “Web analytics” to analyze Web site traffic for patterns that can lead to enhanced SEO. While the term embraces areas more concerned with usability and user behavior while on a Web site, analytics sweeps in SEO functions such as keyword analysis and SERP position monitoring.

Data can be collected by “tagging” individual Web pages with Javascript and other programming, or by placing cookies on visitors’ browsers.

“Keyword analysis” is often initiated at the server level through “log analysis.” Because WWW communications protocols usually pass the full URL of the previously visited page to the destination Web site, the headers can be processed for search terms, and those terms then organized and further analyzed.

In either case, data is then analyzed manually, or with in-house programming, or with one of the many commercial Web analytics programs.

The data below is a sample of some of the information that can be extracted from Web site server logs: visitor’s ISP, visitor’s IP address, number of pages visited, time and date, visitor’s browser and operating system, the Web site the visitor came from, the search terms they used, and the landing page.

A visitor from “dynamic.dsl.com” (88.104.65.226) was logged twice, starting at 12:46:36 on Sunday, October 14, 2007.

The initial browser was Firefox/2.0.0.7 (Windows XP; en).

This visitor first arrived from <http://www.google.co.uk> while searching “ecommerce business models 1–10” and visited “digitalenterprise.org/models/models.html.”

There has been a general shift in Web analytics from log analysis to page tagging, in part due to the limits of log information compared with information gained by tagging, and in part driven by vendors of analytics software.

COMBINING CONTENT, TECHNICAL, AND LINKING SEO

Strategies that combine all three forms of SEO may be the most effective in gaining high SERP positions. Combined strategies are uncommon because Web sites are normally designed without regard to SEO (although the most effective SEO is begun at the conceptual level), and because Web site owners and managers rarely allow significant changes to the structure or content of their Web sites after they are built and online.

An alternate way of viewing SEO is “on page” or “off page” SEO, the former being methods applied directly to Web sites and pages, the latter focusing on methods external to a Web site, such as link strategies.^[24]

BROAD ISSUES IN SEO

A number of strategic issues that may affect a Web site go beyond fundamental SEO tactics or current “best practices.”

BUILDING DOWNWARDS VS. OUTWARDS

Large or growing organizations face the question of whether to house all their Web pages under a single WWW domain, or to establish separate Web sites for different services, product lines, or agendas. An established organization with a strong brand will usually lean keyword targeting towards its own brand name in SERP positions, and thus benefit from housing all its Web pages under a single domain, whereas a newer organization with separate divisions would be more likely to establish separate Web sites for each product line or goal.

This is not strictly a business issue; a university establishing a new campus in a different city would probably create a new Web site for that campus, just as a business that sells to both architects and game designers might choose to divide its product lines into two separate Web sites. The university would certainly link its two Web sites for the link-weighting benefits; the company selling unrelated products might not link its Web sites at all.

BROAD VS. NARROW TARGETING AND LONG-TAIL TERMS

The potential benefits in targeting highly specific searches were known to SEO practitioners by 1995–1996. Bed and breakfasts were one of the earliest and most evident examples, showing how highly specific keywords (e.g., “bed breakfast Ireland Kilkenny”) could reach precisely the right audience.

Since the Web was still sparsely populated, the more skilled practitioners were able to gain high SERP rankings for broad searches at the same time as they targeted narrow niches. For example, in 1997, it was possible for an Irish bed and breakfast’s Web site to gain a high SERP position for both its own locale and Irish B&Bs in general. That grew more difficult with increasing competition, and today a search for “bed breakfast Ireland” will generally return a SERP dominated by bed and breakfast directories and associations.

The changing situation became somewhat better understood in 2003, when Zipfian distributions were mentioned in an article by Clay Shirky about “power laws” as applied to blogs.^[25] Power laws

were later popularized as “The Long Tail” by an article in *Wired* magazine.^[26] With use of the term growing, many clients and SEO practitioners now refer to any three- to four-word term as a “long-tail term”; others use the description more correctly to describe an uncommon search term.

There are now indications that the pendulum of interest has swung too far towards uncommon terms, and that organizations are targeting long-tail terms without a clear view towards long-term benefits such as memberships or profits.^[27]

BALANCING TARGETING AND SERENDIPITY

While some organizations have only one easily described product or agenda, for most organizations SEO can be too successful when it targets particular search terms so tightly that their Web pages can be found by little else. When virtually all traffic to a Web site comes from a small number of terms, analyzing incoming search terms becomes an exercise in analyzing what one already knows. Ideally, improvements in SEO create an increase in targeted traffic along with an increase in unanticipated search terms. While this can be analyzed in terms of Zipfian distributions, in proactive marketing it is better represented by the APUPA (Alien–Penumbral–Umbral–Penumbral–Alien) bell curve (Figure 13.2).

SEO can also be too successful when a Web page captures a high SERP position on a major search engine for an extremely broad term like “health.” Where this might be satisfactory to a large organization like the U.S. National Institutes of Health, an organization focusing on a particular niche of health could be swamped with masses of unwanted Web site visitors and e-mail.^[28] Web sites deluged by unwanted traffic sometimes convert a liability to an asset by “reselling” the traffic or the entire Web site.

LEAD TIME AND LONGEVITY

Search engines rarely assign top rankings to newly indexed Web pages. A Web site can take months or even years to reach its “natural” position in SERPs. This is sometimes called the “sandbox,” referring to the period a new Web site may wait for a good SERP position or even to be listed at all by a search engine.

Equally, search engines assign value to longevity, and older well-ranked Web sites dating from the 1990s can be notoriously difficult to dislodge from their SERP positions by new competitors.

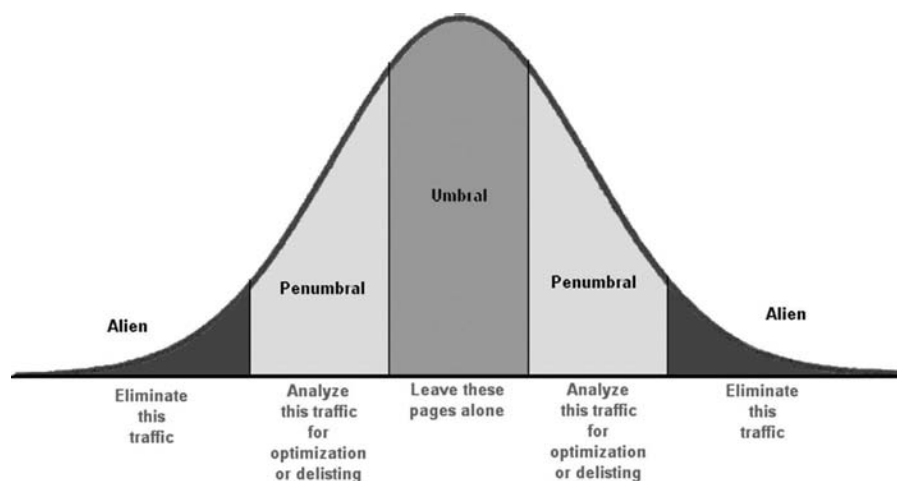


FIGURE 13.2 APUPA chart expressed in SEO terms. © Hastings Research, 2005.

GEOGRAPHIC TARGETING

Geographic targeting through SEO had limited success through the mid-00s. Few early Web site designers made a diligent attempt to anticipate all the geographic keywords searchers might use, and those who did target searches with geographic keywords often used so many keywords (e.g., “London,” “Pimlico,” “Knightsbridge,” “Belgravia,” “Dulwich,” “Stratford,” “Paddington,” “Fleet Street”—and every other village and neighboring town) that the pages were treated as index spam by search engines, and down-ranked to obscurity. Other designers created dozens or even thousands of “mirror pages” (pages with substantially similar content), each targeting a particular locale, with the result that the pages were down-ranked for “duplicate content.”

In a cause-and-effect loop, users learned that searching by locale was nearly useless, and abandoned the effort (with the exception of specific travel destinations), leading most SEO practitioners to abandon their efforts at geographic targeting.

With the now-growing success of geographic targeting efforts by search engines—typically displaying maps showing physical locations—users have again started searching geographically, and most SEO practitioners advise making some effort to target searches containing geographic keywords, even though geographic searches may be defined by commercial databases rather than SEO efforts.

SEO SOURCE TRACKING

The wealth of data that accompanied visitors to Web sites led many ecommerce pioneers to think the perfection of marketing analysis had arrived. While this sometimes bordered on true for new businesses that were entirely WWW-based, the opposite turned out to be the case for large, established organizations already being marketed through advertising, mail, public relations, and retail outlets; the WWW added another layer of complexity to their market tracking. Due to its 24-hr availability, the WWW also creates a “smoothing” effect on response to promotional efforts, which can hinder analysis of event-driven spikes in visitor traffic.

Refinements in Web analytics have not entirely compensated for these factors, and in many ways Web site traffic source analysis remains less precise than that of traditional promotion and advertising.

INCREASED COMPETITION

The continuing growth in the number of Web sites and Web pages means newly created Web sites have increasing difficulty gaining rankings for common keywords, so their approach is often to use a fanciful domain name and drive Web traffic via social networking or publicity instead of SEO.

Aside from the natural growth in Web sites, huge numbers of “made for ads” (MFA) Web sites have been created for no purpose other than to make money by hosting online advertisements—they have no products, services, or agendas of their own. Because advertising revenue indirectly comes from organizations that pay for online ads, MFA Web sites often specifically target keywords used by existing organizations. Nonprofit informational Web sites such as Wikipedia (<http://www.wikipedia.org/>) also compete directly for many search terms, often gaining very high search rankings.

SEARCH ENGINES ONLY INDEX WORDS

While search engines are often technically capable of distilling words when they are represented in images, and in theory from audio tracks, in practice they index words from readable text documents, file names, or metadata attached to other forms of electronic files. There is a persistent failure to understand this aspect of SEO, with many Web site owners believing that if the human eye can read the words or grasp the meaning, the search engines will correctly index the files.

Further restricting SEO, search engines by choice index only certain types of text documents or metadata. At one time that meant pure text or HTML only. Most search engines are now willing to crawl and index file formats such as PDF (Portable Document Format), Microsoft Word and Excel, and in some cases Flash.

TRENDS

FADING TRENDS

“Meta keyword tags” were the first major chimera in SEO (“The meta keywords tag” earlier). They have been followed by other imaginary fast-tracks to high SERP position, all beginning with a grain of truth then blown out of proportion.

“Google PageRank” has been one of the most persistent areas of focus in SEO. Because of media attention and the easy access to the publicly visible PageRank via the Google Toolbar, the scale has commanded a great deal of interest from Web site owners, many of whom consider the published PageRank a practical scale for performance-based SEO contracts. Most SEO professionals now consider public PageRank in itself a minor factor in a Web site’s actual SERP positions, more an effect than a cause. The latter opinion is to some extent corroborated by Google, Inc.^[29] Over the course of 2007, countless Web sites saw their Google public PageRank drop significantly, which may further diminish interest in the measurement. (“PageRank” in popular usage should not be confused with “pagerank,” an internal Google term.)

Web site “traffic”—whether measured in “hits” (requests for individual files, whether pages or images), “page views,” “unique visitors,” or the dot-com measure of “more eyeballs”—is slowly losing its popular connection to rankings, as organizations focus on conversion rate of users’ visits to desired actions. (In extreme cases, traffic volume will have greater influence on SERP positions than relevance; however, this is uncommon and often results from a situation such as a major news event, where the boost in rankings may be due to a proliferation of inbound links, rather than traffic volume itself.)

CURRENT TRENDS

Link building previously described in “Link-weighting fundamentals”

Pursuing “long-tail” searches (previously described in “Broad vs. Narrow Targeting and Long-Tail Terms”)

Bringing SEO in-house: Major companies in North America and Europe are increasingly bringing SEO in-house as their online sales grow to (USD) millions or hundreds of millions. Companies are hiring in consultants, training current employees, or both.^[30]

Using best practices: Clients and employers are now commonly asking that practitioners follow “best practices.” For SEO, the term describes such basic practices as prioritizing meta titles and heading tags, including keywords, and avoiding discredited (Black Hat) tactics.^[31]

Tuning metadata for search engine interfaces: Increasingly, practitioners structure Web pages so that titles, descriptions, and filenames are presented appealingly on SERPs. Eye-tracking studies, where eye motion is represented by printed “heat maps,” are currently the basis for most decisions in tuning metadata.^[32]

Tagging Web pages: adding keywords or allowing visitors to add keywords to the visible text—is common on many social networking Web sites and some ecommerce and media Web sites, but has shown mixed results in improving SERP positions, possibly due to the indiscriminately chosen keywords selected by lay users, Web site owners, and bloggers. Tagging may benefit SERP position most for Web sites that encourage commentators to use “controlled vocabularies.”^[33] It may also boost SERP positions for “long-tail” search terms.

Web 2.0 designs: Web 2.0 methods have been criticized as being detrimental to SEO. This criticism has arisen with each evolution in Web site development, from “dynamic page delivery” (assembling Web pages from databases “on the fly”) through Cascading Style Sheets and PDF files. In most previous instances, organizations’ design preferences have won out over SEO considerations, and eventually the major search engines have adapted their protocols to classify new Web page structures and document formats.

Since search engines do not immediately retool for new forms of Web page markup and structure, Web sites that employ Web 2.0 methods such as AJAX (Asynchronous Javascript and XML) or that stream together content from different sources (“mashups”) may drop in SERP positions until search engines decide how to interpret these new document structures.

TRENDS AFFECTING SEO

New Search Engine Presentation Methods

Clusters of search records such as those displayed by <http://www.clusty.com> or <http://www.ask.com> are likely to continue improving in relevance; if they increase in popularity, each cluster link will constitute a second-tier SERP, and SEO may begin targeting particular clusters.

Drop-down *contextual search suggestions* are now offered at <http://www.ask.com> and <http://www.yahoo.com>. Because users see contextual suggestions as they type, even before seeing the first SERP, “contextual position” may become a sought-after goal of SEO. Since contextual suggestions are all displayed on the first screen of a search engine’s Web site, without the need for scrolling down, position is not likely to be as critical as SERP position, and “contextual inclusion” may prove to be the desired goal. (Contextual search flows from a given search engine’s database of “most likely” requests; it should not be confused with the URL suggestions that Web browsers make by accessing a user’s own search history.) (Figure 13.3).

With exceptions, such as the European search engine Kartoo (<http://www.kartoo.com>), interfaces such as topic maps and link maps show little sign of entering the mainstream of WWW search, and thus do not affect SEO.

Blended search—the blending into the primary WWW SERPs of news, images, maps, videos, and other types of records once considered niche searches—have now been instituted by major search engines.^[34] This presents opportunities to the more aggressive SEO practitioners, some of whom are attempting to dominate blended SERPs with a mix of varied data formats. (Blended search has also been called “universal search.”)

Specialized Searches

Specialized searches have been offered for decades through pre-Web online information providers. They are often called “vertical searches” (i.e., niche search). Specialized search could disrupt current SEO strategies, whether it is offered by major search engines or niche suppliers. First, if niche searches or tools proliferate in the GYM group (Google, Yahoo, Microsoft)—each with its own

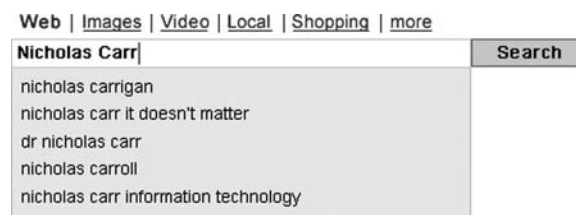


FIGURE 13.3 Example of drop-down contextual search suggestions.

classification protocols—they may dilute the value of SEO methods that address the GYM standard algorithms. Second, niche search engines such as <http://www.vivisimo.com> have entered the search race, and others such as <http://www.northernlight.com> (founded 1996) are reentering public search; these specialized search engines are usually less susceptible to technical methods, as they tend to use proportionately more human editing than broad reach search engines. Either trend could diminish the importance of technical- or link-based SEO.

“Personal search” has been a holy grail of the Internet industry since the mid-1990s. Early attempts at personal search included Yahoo! personalization features and numerous failures in “push technology” (Push technology: A means of automatically delivering information via the Internet to subscribers based on their choices for customized news, etc. *Webster's New Millennium Dictionary of English*, 2007. Described by H.P. Luhn as “SDI” (Selective Dissemination System), <http://www.ischool.utexas.edu/~ssoy/organizing/l391d2c.htm>, accessed February 2009). The current trend is towards analyzing individual users' search requests in the context of their previous searches (remembered by the server) and delivering customized results. Taken to an extreme, this could theoretically filter out a majority of available Web pages regardless of the SEO efforts invested, as search engines “learn” to focus on individual users' core interests.

“Local search” (geographic search) has become a serious goal for search engines. However other than travel destinations, few small retail businesses have Web sites, so search engines cannot easily find data for them by crawling the WWW. The primary solution to date has been to buy or barter for data such as traditional telephone book databases. Other new local search companies are aggregating local data for search engines, often crawling online local directories for source data search engines themselves are encouraging data entry from local organizations. Some trades with guilds such as law or medicine have niche Web sites with search functions; as yet few are comprehensive. Trades without guild organizations such as auto repair or beauty services may depend entirely on the success of aggregation if they are to be located through broad search. In few of these cases do current SEO methods provide a clear path into SERPs, and paid listings may vie with SEO in creating online exposure for local businesses.^[35]

Alternate Search Channels

Alternate search channels could have an even more dramatic effect on SEO practices than blended or niche search services:

SEO for mobile communications devices: Analysis of users' WWW search behaviors on mobile devices has shown an emphasis on local retail search followed by entertainment. Coupling mobile search with GPS navigation feedback opens the prospect of delivering content or ads raises the prospect of directing mobile users to the nearest restaurant or a movie showing, thus both of these searches are potentially lucrative advertising venues—search engines are pursuing them, and SEO will follow. It is not yet certain whether search engines will convert conventional Web pages to mobile-friendly formats, or whether organizations will have to create new pages specifically targeting mobile devices.^[36,37]

SMO (Social Media Optimization): It is arguable how much SMO involves “search” in the sense of users searching by keywords, because the spread of information on social networking Web sites is largely viral (self-promotion coupled with word-of-mouth). SMO is currently practiced by creating profiles on popular social media sharing and news Web sites, building a large base of “friends” and contributing unique promotional content with the option for other community members to vote in favor or against. Popular content that goes “hot” is placed on the high traffic home pages of the social media sharing and news sites, sometimes generating tremendous amounts of exposure, direct traffic, and secondary effect inbound links from bloggers that post about the content. SMO also involves using software to make posts automatically, this latter method being simply spamming. There may be a growing synergy between SEO and SMO, and SMO may become a parallel profession lumped under the umbrella of SEM.^[38,39]

STANDARDS AND REGULATION IN SEO

The practice of SEO is marked by the absence of standards, other than those imposed by the willingness of major search engines to index Web pages, their level of effort in returning the most relevant Web pages, and punitive actions (down-ranking or de-listing) taken against what search engine administrators consider unacceptable manipulation of SERP positions.

HTML

Standard Web pages (HTML) are in theory defined by W3C standards—which are largely ignored by both Web developers and amateurs. Other formats now being indexed by search engines—including PDF, Microsoft Word and Excel, and still images and videos—follow even fewer standards.

CERTIFICATION

Certification is in beginning stages. Some businesses have begun offering SEO credentials, and accredited colleges are now beginning to offer courses in SEO.

SEO INDUSTRY

The SEO industry may be loosely divided into SEO professionals of varying ability; Web designers and programmers who know the rudiments; and a much more visible group of low-cost SEO companies that operate through advertising and e-mail marketing.^[40]

While there are no formal standards or regulation, informal standards are slowly evolving through consensus of trade associations and Web sites reporting on SEO. (See “Bibliography” at the end of this entry.)

SEARCH ENGINES

Because it is difficult to mathematically determine whether a given Web page’s content accurately reflects the page’s relevance to visitors (much less the creator’s true agenda), search engines have faced and continue to face an impossible task in down-ranking deceptive pages without penalizing legitimate ones. (In the classic if now simplistic example, a medical Web site can unwittingly trigger an algorithm aimed at filtering out pornography pages.)

Attempts to “game” (manipulate) search engine rankings have been so relentless, from the first significant appearance of online search engines, that many experts consider search engine administration to be equally a process of excluding Black Hat pages and elevating relevant pages.^[41] While penalizing irrelevant content faces the same problem as returning relevant content—a struggle to develop artificial intelligence—most search engines strongly downgrade Web sites that are found to be using the more technically complex SEO tactics such as IP-specific page delivery (see “Cloaking” earlier).

This issue of controlling Black Hat SEO without penalizing White Hat SEO is a continued source of tension between search engine administrators and SEO practitioners. Search engines selling advertising space on their own SERPs has exacerbated the tension by raising the issue that pursues traditional media, of whether there is a true division between editorial and advertising departments (or search relevance and advertising, in the case of search engines).

Like the evolution of military tactics, regulation is a game of innovation and counter-measure. The counter-measures may make a given tactic so ineffective that it is completely forgotten by both attackers and defenders, at which point it may be reintroduced by the attackers. (A medical analogy would be the mutation of microorganisms in reaction to new antibiotics, accompanied by the resurgence of forgotten diseases when society no longer guards against them.)

GOVERNMENT REGULATION

At present there is no hint of SEO regulation. As subcontractors, SEO consultancies are in a position similar to advertising agencies—with no actual control over organizations or distribution channels—so any government regulation is likely to flow indirectly from regulation of Web sites themselves (trade practices) or of search engines (particularly in the United States, where anti-trust law may affect conglomeration of major search engines with software, hardware, or media companies).

USER BEHAVIORS IN SEARCH

User behavior while searching online shows more consistency than change in the period from 1995 to 2008. This section will briefly address the consistent behaviors and available research on the course of changes.

Users' search behavior has a strong effect on SEO, since many users click the first listing in a SERP. On the other hand, gaining a #1 SERP position for a particular search term does not guarantee that a user will click that link; they might click the #2 link or the #10 link if those page titles, descriptions, or URLs are more compelling.

In the same vein, there is no clear evidence that users who click the first record will take further action; many SEO practitioners believe the more motivated users will scan an entire SERP before deciding which link to click. Regardless, user behavior while visiting SERPs affects SEO decisions as well as the search engines' goal of relevance.

Research into specific behaviors has revealed a great deal about what users do in the specific environment of a search engine interface when tested in a laboratory setting. It leaves unanswered many questions about why users search the way they do—in part because most research has focused on what users do when looking at a SERP, without inquiring how or why they arrived at that particular SERP, or what actions they take after clicking a particular link. More research is required to develop consistent conclusions. Substantial research from information science has yet to be incorporated in SEO; when it is, it may transform strategies and tactics.

Variables that affect user search behavior:

- Users' level of subject knowledge
- Improvements in results returned by search engines
- Users' confidence in the quality of search results
- Degree of users' sophistication in search
- Presentation style of individual search results (title, description, URL, etc.)
- Motivation/curiosity/laziness
- Habituation among users (hardening behaviors)
- Search engine interface design
- Number of search engines available

BASIC BEHAVIORS

Few users methodically click search results in sequential order, from the first result to the bottom of the page. Typically they skip over unappealing titles or URLs, and may bounce back and forth between organic search listings and paid listings. (Depending on SERP design, users may not always know the difference between paid and organic listings.) If a relevant Web page is not found quickly, users may change search engines, change their search term, or migrate to a general information Web site such as a dictionary, encyclopedia, or user-fed Q&A Web site.

SEARCHING POPULAR SUBJECTS

Patterns of search seen on Wordtracker, Yahoo! search tool, or KeywordDiscovery continue to show a preponderance of searches for celebrities and popular news, as described since 1999 in the weekly “Lycos 50” listing of the most popular topical searches (<http://50.lycos.com/archives.asp>) with written analysis and opinion. Google Trends (<http://www.google.com/trends>), which shows the number of searches for chosen terms in chart format, is well-suited to display spikes in interest, relative interest in different subjects, or seasonal patterns in popular searches; like the Lycos 50, it demonstrates that a huge proportion of Web searches are for popular subjects rather than niche information.

USE OF BOOLEAN SYNTAX

In the 1990s, WWW search engines typically limited the use of search syntax to Boolean AND, OR, and NOT, with OR as the default. (With AND or NOT typically applied by +/- symbols.) Initially these options were only available by typing them into the search term entry box. HotBot search engine was one of the earliest to offer Boolean logic to untrained users, with the addition of a drop-down box offering AND as well as “exact phrase” options, but the options were rarely used.

In the late 1990s, search engines moved towards a default (or forced) Boolean AND; by early 2003, it was the default on all major search engines search engines.^[42] Although this change narrowed search results, it also hugely increased the relevance of results, and, at the same time, reduced the average user’s motivation to learn Boolean syntax. Some data samples suggest that Boolean search skills are known to a smaller percentage of users today than in 1997, though it is uncertain whether the decrease is due to a loss of interest in search syntax or an influx of less sophisticated users.^[43]

SEARCH ENGINE LOYALTY

Available statistics strongly indicate that users are now more likely to refine or expand their search term than switch to another search engine.^[44] The disinclination to switch is often called “loyalty” in a broad sense; motivating factors could be the generally improved quality of search returns over the last few years, or users’ increased recognition of the value of using precise search terms. In terms of user behavior, this could be described as a slight shift from “berrypicking” (wandering about and collecting bits of information) to more linear search behaviors.^[45]

HARDENED USER BEHAVIORS

Users’ search behavior search engines may have become “hardened,” or fixed, by their expectations, in particular in the United States and Canada, because the majority of popular search engines have presented returns as linear lists since the mid-1990s. Current SEO is predicated on the idea that users type keywords into a box, click a “Search” button, and scan a list from top to bottom.^[46] Because they represent a large conceptual shift, interfaces that present information in clusters, topic maps, or relationship maps face an entry barrier.

USER BEHAVIOR WORLDWIDE

Patterns of user behavior outside the United States and Canada do not show identical evolution. While less analysis has been done, in regions as disparate as England and China users are apparently more willing to examine entire SERP pages, as well as second and third pages, and less likely to click on the first or first few SERP positions.

CONCLUSION

Beginning in the mid-1990s, SEO evolved from placing random keywords in all possible parts of Web pages, to more focused doorway page strategies targeting particular keywords. By 2000, improvements in search engines were rendering most of such tactics obsolete, and SEO practitioners moved on to the integration of overall Web site structure, and then on to establishing a Web site's relationship with the WWW as a whole through linking.^[47] For most, the reluctant and secondary focus of SEO was on users or quality of content. For others, a focus on user experience as well as SEO provided long term, sustainable results throughout the evolution.

SEO now faces broad changes in search and user behavior. The growth of online information is outpacing the indexing rates of all search engines. Search engines are struggling to deal with that overload, and in the process changing their strategies about what information to present and how to present it. Hardware—notably mobile communications devices—is redefining the technical limits of information presentation and also creating niches in user search behaviors.

ACKNOWLEDGMENTS

With thanks to Jim Rhodes, Jim Heath, Lee Odden, Guy Shalev, Paula Sheil, and Kelly Bryan.

REFERENCES

1. Sullivan, D. <http://forums.searchenginewatch.com/showthread.php?t=78> (accessed February 2009).
2. Rhodes, J. Art of search engine promotion, 1997. <http://www.deadlock.com/promote/search-engines> (accessed February 2009).
3. Heath, J. *Pointers on How to Create Business Websites That Work*, Published online, 1995. Earliest available version at <http://www.viacorp.com/pointers1995.html> (accessed February 2009).
4. November 2006 Web Server Survey Netcraft **2006**, http://news.netcraft.com/archives/2006/11/01/november_2006_web_server_survey.html (accessed February 2009).
5. Search Engine Marketing Professional Organization (SEMPO), *The State of Search Engine Marketing 2006: Survey of Advertisers and Agencies*, 5. Search Engine Marketing Professional Organization (SEMPO): Wakefield, MA, 2007.
6. Web Analytics Association, <http://www.webanalyticsassociation.org> (accessed February 2009).
7. CMSWatch, <http://www.cmswatch.com> (accessed February 2009).
8. Wilsdon, N., Ed.; Global Search Report 2007, E3internet, 2007. <http://www.multilingual-search.com/global-search-report-2007-released-free-pdf/15/10/2007> (accessed February 2009).
9. The Global Structure of an HTML Document, <http://www.w3.org/TR/html401/struct/global.html> (accessed February 2009).
10. Grappone, J.; Couzin, G. *Search Engine Optimization: An Hour A Day*; Sybex: Hoboken, NJ, June 2006; 46–47, 101. <http://www.yourseoplan.com/book-inside-contents.html> (accessed February 2009).
11. Whalen, J. Avoiding clueless-is as clueless-does SEO, <http://searchengineland.com/070927-071933.php> (accessed February 2009).
12. Stamoulis, N. The top 8 SEO techniques (A dispute), <http://www.searchengineoptimizationjournal.com/2007/09/17/the-top-8-seo-techniques-a-dispute> (accessed February 2009).
13. Seda, C. *How to Win Sales & Influence Spiders*; New Riders: Berkeley, CA, February 2007, 5.
14. Bradley, S. Using keyword themes to structure your site content, 2006, <http://www.vanseodesign.com/blog/seo/keyword-themes-for-site-structure/> (accessed February 2009).
15. Sullivan, D. *What Are Doorway Pages?* Published online, 2007, <http://www.searchenginewatch.com/showPage.html?page=2167831> (accessed February 2009).
16. Sherman, C. *In Defense of Search Engine Cloaking*; Published online, 2001, <http://www.searchenginewatch.com/showPage.html?page=2157261> (accessed February 2009).
17. Nelson, T.H. *Literary Machines*; Mindful Press: Sausalito, CA, 1982. (This is the earliest specification of the term hyperlinking, ca. 1965.)
18. Langville, A.; Meyer, C. *Google PageRank and Beyond: The Science of Search Engine Rankings*; Princeton University Press: Princeton, NJ, 2006.

19. Kleinberg, J. *Authoritative Sources in a Hyperlinked Environment*, Proceedings of the 9th ACM-SIAM Symposium on Discrete Algorithms, San Francisco, 1998. Extended version in J. ACM **1999**, 46. Also appears as IBM Research Report RJ 10076, May 1997, <http://www.cs.cornell.edu/home/kleinber/auth.pdf> (accessed February 2009).
20. Sullivan, D. Google kills Bush's miserable failure search & other Google bombs; 2007. <http://searchengineland.com/070125-230048.php> (accessed February 2009).
21. Wall, A. *The SEObook*; 2007, 197–199. <http://www.seobook.com>.
22. Bodden, L. Making sense of linking and site promotion; 2007. <http://www.toprankblog.com/2007/09/linking-site-promotion/> (accessed February 2009).
23. Thurow, S. The right way to think about site maps; 2007. <http://searchengineland.com/070809-091501.php> (accessed February 2009).
24. Off Page Optimisation VS On Page Optimisation, http://www.directtraffic.org/on_page_optimisation.htm (accessed February 2009).
25. Shirky, C. Power laws, weblogs, and inequality; 2003. http://www.shirky.com/writings/powerlaw_weblog.html (accessed February 2009).
26. Anderson, C. The long tail. *Wired Mag.* **2004**. <http://www.wired.com/wired/archive/12.10/tail.html> (accessed February 2009).
27. Brynjolfsson, E.; Hu, Y.; Smith, M.D. From niches to riches: the anatomy of the long tail. *Sloan Manage. Rev.* **2006**.
28. Carroll, N. The anti-thesaurus: A proposal for improving internet search while reducing unnecessary traffic loads; 2001. <http://www.hastingsresearch.com/net/06-anti-thesaurus.shtml> (accessed February 2009).
29. <http://www.google.com/corporate/tech.html> (accessed February 2009).
30. *Search Marketing Benchmark Guide 2008*, 2008, 43–50. <http://www.sherpastore.com/Search-Marketing-Benchmark-2008.html> (accessed February 2009).
31. Fusco, P.J. SEO Best Practices: 20 Questions; 2007, <http://www.clickz.com/showPage.html?page=3624560> (accessed February 2009).
32. Whitepapers and videos on user eye tracking, <http://www.enquiroresearch.com/> (accessed February 2009).
33. Tagging advisory from DailyKos.com political blog, http://www.dkosopedia.com/wiki/Standard_Tags_-_alpha_list (accessed February 2009).
34. Sullivan, D. Search 3.0: The Blended & Vertical Search Revolution, <http://searchengineland.com/071127-091128.php> (accessed February 2009).
35. Guide on how to get your business listed on major local search engines, yellow pages sites and social local networks. <http://www.locallytype.com/pages/submit.htm> (accessed February 2009).
36. Carroll, N.; McGraw, M.; Brahms, S.; Rodgers, D. Wireless usability 2001–2002 report, <http://www.hastingsresearch.com/reports/62-wap-2001.shtml> (accessed February 2009).
37. Holahan, C. The battle for mobile search. *BusinessWeek* **2007**, http://www.businessweek.com/technology/content/feb2007/tc20070220_828216.htm (accessed February 2009).
38. Bhargava, R. 5 Rules of Social Media Optimization (SMO), 2006. http://rohitbhargava.typepad.com/weblog/2006/08/5_rules_of_soci.html (accessed February 2009).
39. Seda, C. *How to Win Sales & Influence Spiders*; New Riders: Berkeley, CA, 2007; 58–74.
40. Seda, C. *How to Win Sales & Influence Spiders*; New Riders: Berkeley, CA, 2007; 16–22.
41. Price, G. Personal communication about gaming search engines, <http://www.resourceshelf.com> (accessed February 2009).
42. Sullivan, D. Search Features Chart; 2001, <http://searchenginewatch.com/showPage.html?page=2155981> (accessed February 2009).
43. Hastings Research proprietary databases of real-time searches and web server logs, 1995–2008.
44. *Search Marketing Benchmark Guide 2008*, 2007, 135, <http://www.sherpastore.com/Search-Marketing-Benchmark-2008.html>.
45. Bates, M.J. The Design Of Browsing And Berrypicking Techniques For The Online Search Interface, 1989, <http://www.gseis.ucla.edu/faculty/bates/berrypicking.html> (accessed February 2009).
46. Nielsen, J. Mental models for search are getting firmer, 2005, <http://www.useit.com/alertbox/20050509.html> (accessed February 2009).
47. Rhodes, J. Personal communication on early SEO practices.

BIBLIOGRAPHY

1. A timeline of SEO. <http://blog.searchenginewatch.com/blog/060417-130526> (accessed February 2009).
2. <http://www.searchengineland.com> (accessed February 2009).
3. <http://www.searchenginewatch.com> (accessed February 2009).
4. <http://www.seopros.org> (accessed February 2009).
5. http://www.sempo.org/learning_center (accessed February 2009).
6. http://www.sempo.org/learning_center/sem_glossary (accessed February 2009).

USERS SECTION

1. Mezei, C. Website and search engine user behavior analysis, 2006, <http://www.searchnewz.com/latestsearch/senews/sn-4-20061101WebsiteandSearchEngineUserBehaviorAnalysis.html> (accessed February 2009).
2. Search Engine User Behavior Study, iProspect, 2006, http://www.iprospect.com/about/whitepaper_seuserbehavior_-apr06.htm (accessed February 2009).

This page intentionally left blank

14 Records Compliance and Risk Management

Bradley J. Wiles

CONTENTS

Introduction.....	197
Records Compliance	198
Risk Management	199
Records Compliance and Risk Management in Combination	200
Conclusion	201
Acknowledgments.....	201
References.....	201
Bibliography	202

INTRODUCTION

The concepts of *compliance* and *risk management* have been applied to various methods and strategies of business operations in a wide variety of industries over the past several decades.^[1] This entry identifies banking, investment firms, and insurance companies as being particularly sensitive to managing risk and maintaining compliance to ensure sound business management. Minimizing risk and meeting various levels of statutory obligations are consistently perceived as necessary for stability, growth, and longevity. This is particularly relevant in the *archives and records administration* field, where securing the human record depends on evolving legal requirements, professional norms, and institutional standards. This entry will examine the place of compliance and risk management within the context of archives and records administration. It will provide separate definitions of each concept and evaluate these as related and complementary aspects of the creation, capture, and use of records.

Perhaps first it would be useful to provide a comprehensive definition for records management, simply because the principles it encompasses are generally applicable to a wide array of records and information activities. Records management is

a professional discipline that is primarily concerned with the management of document-based information systems. The application of systematic and scientific controls to recorded information required in the operation of an organization's business. The systematic control of all organizational records during the various stages of their life cycle, from their creation or receipt, through their processing, distribution, maintenance, and use, to their ultimate disposition.^[2]

The purpose of records management is, "to promote economies and efficiencies in recordkeeping, to assure that useless records are systematically destroyed while valuable information is protected and maintained in a manner that facilitates its access and use."^[2]

Records management is also concerned with the analysis of any and all recorded information created, received, maintained, or used by an organization in accordance with its mission, operations, and activities.^[3] Furthermore, records management deals with records regardless of age, type,

or format, and is often an enterprise-wide undertaking that relies on cooperation at all levels of an organization.^[4] Present-day records management is an outgrowth of traditional archival practice, but typically deals with contemporary recorded information and documents.^[4] In the mid-twentieth century, a distinction was made between “historical” and “modern” documents, mostly because of the deluge of records flowing into traditional archives following World War II.^[4] However, this distinction is no longer prevalent, and records management principles are applied to records produced by corporations, businesses, and governmental units (organizations using and retaining documents in the course of regular business), as well as those kept by historical societies, manuscript libraries, and other cultural institutions (organizations who hold collections for reasons of scholarship or posterity). In nearly all instances, the concepts of compliance and risk management factor significantly into the modern practice of archives and records administration.

RECORDS COMPLIANCE

It is somewhat difficult to comprehensively define compliance in the context of archives and records administration, simply because compliance measures are constantly in flux. In most instances, recordkeeping and reporting are part and parcel of other core compliance areas within organizations. These might include promoting fair and ethical business conduct, requiring proper environmental, health, and safety policies, or ensuring consumer protection,^[5] and may or may not have a records component. In general, records compliance requirements derive from legislation, regulatory standards, codes of best practices, or community expectations (p. 10).^[4] These requirements exist at all social, political, and organizational levels, might be explicit or implied, and incorporate any number or combination of requirements (pp. 38–39).^[4] For example, explicit legal requirements for the creation and management of records in a particular industry could mandate the extended retention of certain types of records created in particular operational contexts. It might also dictate the form in which these records should be maintained, the rights of individuals or corporate bodies to have access to those records, and the protection of the intellectual property that the records contain (p. 38).^[4]

Compliance requirements affect records at all stages throughout their life cycle and can be subject to multiple legal jurisdictions. A significant amount of legal research is often necessary for organizations to determine the legal requirements for retention, destruction, discovery, and disclosure. The first step is to verify which authoritative entities the organization falls under. The degree of regulation varies by the type of industry and while the majority of businesses and organizations are not subject to any recordkeeping guidelines, those that are (e.g., financial services and pharmaceuticals) tend to be heavily regulated by multiple entities.^[6] In particular, those responsible for an organization’s legal and regulatory records compliance in the United States should be familiar with the Uniform Rules of Evidence (which apply to state courts), the Federal Rules of Evidence (which apply to federal courts), the National Conference of Commissioners on Uniform State Laws (which helps interpret and standardize laws across jurisdictions), the Uniform Commercial Code, the Uniform Electronic Transactions Act, the Photographic Copies as Evidence Act, the Federal Paperwork Reduction Act, the Code of Federal Regulations, and the statutes of the National Archives and Records Administration.^[6]

The notion of accountability factors heavily into an organization’s need to maintain compliance with strictures set forth by internal and external forces. In order to support accountability the records management department or representative must initiate a multilateral assessment of business needs, evidentiary requirements, legal obligations, and recognized best practices (pp. 155, 158).^[4] This will enable ready response to legal challenges, grievances, and complaints from within and outside of the organization. Typically, compliance is checked—and accountability is fostered—through any number of internal and external inspections and audits, which often consist of a close physical examination of an organization’s recordkeeping system measured against whatever compliance issues are at hand (p. 159).^[4] The most effective audit procedures are ones established at the

founding of the vital records management program and that remain sensitive to changing compliance regulations (pp. 75–76).^[3] Audits best support accountability through regular scheduling that matches other organizational troubleshooting efforts and through designating proper authority to carry out inspections, make assessments, and implement recommendations. Though records retention audits typically only consist of a sampling of an organization's records series, this is usually sufficient to determine that organization's appropriate level of need for records security, storage, duplication, and destruction (pp. 75–76).^[3]

RISK MANAGEMENT

According to the Association of Records Managers and Administrators, or ARMA International, risk management is a process that ensures an organization does not assume an unacceptable level of risk and includes both risk analysis and risk assessment procedures as key components.^[7] The level of risk is determined by the importance of the record. All records have relative value, but vital records—ones that are unique and irreplaceable—tend to take precedence over those that can be recreated or those that are not essential to normal business operations (p. 73).^[2] Risk analysis involves the evaluation of exposure of vital records to specific risks and provides a basis for protection planning by identifying threats and vulnerabilities (p. 133).^[3] There are several broad categories of threats, which include destruction, loss, corruption, and improper disclosure of records and information (p. 133).^[3]

In any organization it is imperative to discern acceptable levels of risk and typically this is accomplished through qualitative or quantitative risk assessment. The qualitative approach is most useful at identifying and categorizing physical security problems and other vulnerabilities such as malicious destruction, warfare, terrorist attacks, civil insurrections, purposeful sabotage, vandalism, and theft, as well as natural disasters or man-made accidents, errors, negligence, and incompetence (pp. 133, 136).^[3] Quantitative risk assessment is based on methodologies developed for product safety analysis and computer security applications (p. 136).^[3] It initially relies on qualitative measures such as site visits, and focuses on group discussions, but also uses numeric calculations to determine the annualized loss expectancy of a given institution (p. 136).^[3] This dollar amount is estimated by figuring the probability of a particular event (or events) occurring over a given time span (p. 137).^[3] Ultimately, it factors in the cost of record or information restoration, the loss of business as a result of not being able to perform specific business functions, and the costs of any litigation that might ensue (p. 137).^[3] Upon completion of the analysis and assessment, risk can be managed through a series of controls, protections, and preventative measures that include strategic location (or relocation) of facilities, limiting access to vital records storage areas, stringent personnel screening and training, effective physical and electronic security, maintenance of multiple copies of vital records and information, initiating disaster recovery and business continuity plans, and developing a system of internal audits to ensure structural integrity (pp. 138–141).^[3]

Though risk management largely addresses the prospect of lost business or capital, it would be insufficient to consider records risk management strictly in financial terms. One of the main tools of managing risk is the development of records retention policies and schedules. A records retention plan helps prioritize value by determining the risks involved in not keeping or destroying a set of records against the costs of maintaining them (p. 161).^[4] However, this often represents a convergence of competing interests between internal and external factors, which are not always financial. For example, cultural institutions such as archives repositories or special research libraries might feel justified spending an inordinate amount of resources to keep a certain records collection because of its social or historical importance, or because of the prestige it confers on the institution. Likewise, organizations that do not use records expressly for their scholarly or cultural value still might justify the expense of maintaining an historical archive because of its potential to provide materials for advertising, marketing, public relations, and in support of corporate memory (pp. 158–159).^[4] The length and quality of any retention program can be radically contingent among

the various types of organizations that use records and the functions that those records fulfill.^[8] This entry identifies the traditional retention schedules as following along departmental and functional lines, which can cause a tremendous amount of overlap and duplication. It recommends a process-based retention schedule that documents the work flow and follows legal retention requirements and business needs. This duration is based on the operational needs of the institution and the multiple levels of risk that accompanies those needs.

Though risk management procedures are partially based on what *might* happen—the worst-case scenario, the probability of disaster—it also involves active management of records and information assets from the earliest stages of production. Increasingly this means following a legally viable records retention policy, after taking steps to ensure authenticity in the production of records and information to meet all legal and regulatory requirements (p. 103).^[4] The possibility of litigation has emerged as a primary concern in the creation and capture of records, particularly in the digital age when most documents are not necessarily created, stored, and disseminated in static formats.^[9] The types of records that withstand legal challenges are those created as part of a transaction (as opposed to those created *ad hoc* or retroactively) during the regular course of business and that comply with legal guidelines for content, medium, or layout (pp. 103–104).^[4] Indeed, the landscape of records management, and thus the tools and procedures for managing risk, currently operate under a hyper-regulated atmosphere of compliance.

RECORDS COMPLIANCE AND RISK MANAGEMENT IN COMBINATION

According to a recent article in a financial services publication,

In recent years . . . risk management and compliance have begun to converge in a new paradigm. In what amounts to a subtle but important shift in mindset, compliance requirements are now considered a major source of risk . . . Such thinking leads to a natural prioritization of compliance initiatives: Those risks that pose the greatest threat . . . are highest on the list. Strategies to eliminate and/or manage such risks become the primary focus of the compliance function. Since compliance requirements that relate to protecting client interests pose some of the greatest risks . . . taking a risk-oriented approach to compliance naturally aligns the firm's interests with client interests.^[10]

The relationship of records compliance and risk management in archives and records administration is often one of reciprocity. The identification and management of risk is usually dictated by compliance measures that exist within and outside of an organization. Compliance measures—through legislation, best-practices initiatives, or organizational rules—are sensitive to business needs and generally reflect standards determined by industry-wide risk assessment. This reciprocal relationship is perhaps most fully embodied in the Australian National Archives, which adopted the Australian Standards for Records Management (AS-4390) in 1996.^[11] AS-4390 approaches government records from a risk management perspective and seeks to ensure impartiality and accountability among the creators and keepers of state records.^[11] AS-4390 also provides for a system of compliance audits and continual monitoring of documentary integrity (pp. 171–172).^[4] In this sense, AS-4390 demonstrates how, when applied to an organization, the principles of risk management and compliance are interrelated and complementary. In 2001, it was adapted by the International Organization for Standardization (ISO 15489-1:2001) as the preferred model for electronic records produced and maintained by governmental bodies (pp. 27, 249).^[4] The authors note that ISO 15489-1:2001 standard for sustainable records management is based on AS-4390, and these provide the “benchmark against which records management programmes and systems can be measured.”

In some instances the confluence of risk management and compliance complicates the rubric under which an organization operates. Risk often accompanies reward and sometimes the benefit of partial or noncompliance is perceived as outweighing the cost. For example, a multinational company may have branches in several jurisdictions that all require different retention periods for

vital records. To keep all vital records in every jurisdiction for the longest possible period to cover all bases may be cost prohibitive. To irregularly schedule and follow several retention plans may be logistically impossible. In this instance, the organization will likely turn to whatever cost/benefit analysis its risk and compliance cohorts have formulated. They would consider which records were most abundant and valuable, where these records were located, and exactly what compliance requirements it was absolutely necessary to fulfill. In the event that the compliance measures did not specifically address records, the organization would factor in limitations on legal recourse against any action they might carry out (p. 159).^[4] From there, it could be decided that the risk of less than full compliance of certain records (of relatively low value) in certain jurisdictions (with less punitive measures and enforcement) is acceptable. The likelihood of getting caught and receiving a fine, or of losing business and prestige, might be considered less than the cost of full compliance.

Though this may not present an ideal scenario, realistically the level of compliance becomes a matter of how an institution chooses to allocate its resources, particularly in a regulation-heavy climate.^[12] This entry asserts that despite the billions of dollars spent to accommodate laws like Sarbanes–Oxley and HIPPA, one hundred percent compliance is “fundamentally impossible.” However, compliance and risk management are increasingly no longer viewed as necessary evils or impediments to profitable business operations. More and more records are considered assets, whether they are used in everyday business transactions or for academic and cultural reasons. As a result, risk management and compliance are increasingly perceived as core operational functions, integral to the solvency and growth of an organization. In the wake of far-reaching banking, investment, and corporate accounting scandals of this decade, risk management and compliance in records administration should further gain cachet as necessary systemic enhancements. They should no longer merely serve as corrective measures only used after an organization gets caught doing something illegal or unethical.

CONCLUSION

This entry has attempted to identify the key elements of risk management and compliance in archives and records administration, both as individual concepts and as complementary components of recordkeeping. It should be reiterated that the management of risk requires an enterprise-wide assessment of threats and vulnerabilities increasingly aimed at securing dynamic documentary formats. As a result of greater reliance on digital mediums, the move toward more stringent regulation and standardization has found both support and controversy in the archives and records management world. It could safely be assumed that as governments, businesses, cultural institutions, and other organizations become more fully digitized, risk management and compliance will become even more closely connected.

ACKNOWLEDGMENTS

The author would like to thank the following: Dr. Ciaran Trace and Dr. Louise Robbins at the University of Wisconsin School of Library and Information Studies, and Dr. Peter Gottlieb at the Wisconsin Historical Society.

REFERENCES

1. Ludwick, K. Tackling risk-based compliance. *J. Invest. Compl.* **2006**, 7 (4), 61.
2. Robek, M.; Brown, G.; Stephens, D., Eds. *Information and Records Management: Document-Based Information Systems*, 4th Ed.; Glencoe/McGraw-Hill, Inc.: Woodland Hills, CA, 1995; 585.
3. Saffady, W. *Records and Information Management: Fundamentals of Professional Practice*; ARMA International: Lenexa, KS, 2004.

4. Shepherd, E.; Yeo, G. *Managing Records: A Handbook of Principles and Practice*; Facet Publishing: London, 2003.
5. Cellini, R.J. Compliance risk: A top-10 list. *Directors Boards* **2007**, *31*, 52–53.
6. Cogar, R.N. Legal 101 for RIM Professionals. *Inform. Manag. J.* **2005**, *39* (6), 49–56.
7. ARMA International, *Vital Records: Identifying, Managing, and Recovering Business-Critical Records*; ANSI/ARMA International: Lenexa, KS, 2003; 6.
8. Torres, T. Creating a process-focused retention schedule. *Inform. Manag. J.* **2006**, *40* (5), 62–69.
9. Arnold, J.R. We're not in Kansas anymore. *AIIM E-doc Mag.* **2007**, *21* (1), 50–54.
10. Martin, D.; Manley, M.R. Linking compliance, risk management. *Pensions Investments* **2006**, *34* (18), 12.
11. Boles, F. *Selecting and Appraising Archives and Manuscripts*; Society of American Archivists: Chicago, 2005; 31–32.
12. Gincel, R. The awful truth about compliance. *InfoWorld* **2005**, *27* (50), 29.

BIBLIOGRAPHY

1. Allman, T.Y. Fostering a compliance culture: The role of the Sedona guidelines. *Inform. Manag. J.* **2005**, *39* (2), 54–61.
2. Kahn, R.A. Records management and compliance: Making the connection. *Inform. Manag. J.* **2004**, *18* (3), 28–35.
3. Lemieux, V.L. *Managing Risks for Records and Information*; ARMA International: Lenexa, KS, 2004.
4. Montana, J.C.; Dietel, J.E.; Martins, C.S. *Sarbanes–Oxley Act: Implications for Records Management*; ARMA International: Lenexa, KS, 2003.
5. St. Germain, R. Information security management best practice based on ISO/IEC 17799. *Inform. Manag. J.* **2005**, *39* (4), 60–66.
6. Swartz, N. Enterprise-wide records training: Key to compliance success. *Inform. Manag. J.* **2006**, *40* (5), 34–44.

15 Version Control

Jill E. Grogg and Jeff Weddle

CONTENTS

Introduction.....	203
Scope and History.....	204
Version of Record and Open Access.....	205
Peer Review	206
Appropriate Copy and OpenURL.....	207
Technological Initiatives	207
Conclusion	209
References.....	209

INTRODUCTION

For library and information science, version control is the activity of identifying and managing iterations of a document for archival and retrieval purposes. While version control can be problematic for documents and information objects of all types and formats, it is especially so for scholarly communication—particularly journal articles—distributed in electronic formats.

In all the many unforeseen challenges that electronic publishing has unleashed for the information profession, there exist two principle elements: political and technical. Version control is no exception. Many stakeholders in the information industry are shaping the version control discussion, and these discussions have resulted in recommended solutions and directions.

In 2002, a research team headed by Raymond Siemens produced a report entitled “The Credibility of Electronic Publishing” for the Humanities and Social Sciences Federation of Canada (HSSFC). This report attempted to define and assess the current state of affairs for scholarly electronic publishing as well as provide some recommendations for future action.^[1] One of the sections of the report, “Archiving and Text Fluidity/Version Control,” specifically addresses the unique situation that electronic publishing presents for text alteration. In this section, members of the HSSFC-appointed research team note:

Texts can be adapted, abstracted, translated, edited, condensed, corrected, marked up, transcribed, annotated, amended, paraphrased, transliterated, illustrated, indexed, or abridged. They can be commented upon or referenced by other texts. They can be analysed or synthesized. These modifications, and any others that might be considered, produce distinct versions, all related to the original text.^[2]

Indeed, what is important to note is that any alteration(s) to an original text produces a distinct version. Therefore, versions of a text need to be uniquely identifiable while simultaneously containing information that describes the relationships among the various versions. Version control, then, is the effort to “name the versions of the text” and to specify “the differences between the revision and the original text.”^[2] Moreover, version control endeavors to define what is meant by the original and authoritative text.

In 2005, the U.S. Government Printing Office (GPO) prepared a white paper on version control of documents within the purview of the Federal Depository Library Program. The GPO was

particularly concerned with electronically disseminated documents. Government Printing Office concerns fall into seven broad categories. These include version definition, or “unique manifestation of a publication”; version control, which includes “acquiring, cataloging, storing, preserving, and retrieving different versions of publications”; version triggers, which are changes that exceed agreed upon limits; version detection, which deals with identifying and detecting alternate versions; version identifiers, which are metadata tags; version crosswalks, which are a subset of version identifiers and provide links to all other versions; and depository library responsibilities for superseded versions. While it focuses on government documents, the GPO report is extensible in that it offers parameters that can be applied to any discussion of version control.^[3]

In April 2008, the National Information Standards Organization (NISO), in partnership with the Association of Learned and Professional Society Publishers (ALPSP), released a report that defines specific stages or iterations of a journal article. Similar to the GPO’s seven broad categories, the NISO/ALPSP Journal Article Versions (JAV) Technical Working Group identified seven terms, or stages, of an article’s development. According to members of the JAV Technical Working Group, the report is meant to offer recommended practices for the scholarly publishing community. The report summarizes the issue:

Researchers, their institutions, and journal publishers are rapidly moving on from using static, single copies of research papers that are essentially “images” of a printed document. Changes in the way we create, edit, circulate, validate, publish, find, use, and update articles are producing multiple versions whose status and provenance is often unclear. Online searching now allows multiple versions to be found but rarely makes clear the relationships between them.^[4]

Specific recommendations of the group are discussed in subsequent sections.

SCOPE AND HISTORY

Standards, rules, local practices, terminology, and definitions evolved over many years to ensure version control of printed matter. Even so, it has always been problematic to ensure that a published text is in complete agreement with all other versions of that text. Pirated or plagiarized editions, misreadings of manuscripts, and other issues have long clouded issues of text authenticity and authority. With the emergence of electronic publications, these rules, practices, and so forth face much greater challenges. Contemporary efforts to address long-standing problems of unique manuscript identification are addressed by Matthew J. Driscoll in “The MASTER Project: Defining Standards for Electronic Manuscript Catalog Records.” The international Manuscript Access through Standards for Electronic Records (MASTER) Project’s “document type definition” (DTD) standards and application of the closely associated Text Encoding Initiative’s (TEI) *Guidelines for Electronic Text Encoding and Interchange* are applied to archived medieval manuscripts. TEI guidelines allow for description along several broad and flexible categories, including its physical location; title, place of origin, and language(s); contents; physical description; history; and administrative issues such as availability and custodial history.^[5]

The shift from printed manuscripts to electronic documents greatly accelerated the appearance of document iteration. This acceleration becomes problematic when coupled with the fact that gatekeeping functions traditionally associated with scholarly and commercial publishing diminish when authors, or any other person, can, in some cases, disregard copyright law and change a text at will, store it on a server, and disseminate it via whatever scheme he or she chooses. Copyright becomes a tangled Web when one considers that the author may no longer hold copyright to his or her own work if he or she has agreed to transfer copyright to the publisher for a limited or unlimited amount of time. Therefore, he or she may not possess the legal permission to post particular versions—especially the final “published” version—of his or her own work by self-archiving on a personal Web site or in an institutional repository.

Self-archiving is usually the process of an author maintaining a document copy in electronic format. This electronic document may be made available to the author or anyone else designated by the author. The clear benefit of self-archiving is easy access and information dissemination within the scholarly community. One potentially negative consequence is that variant versions of a particular document may circulate in this way. An institutional repository is “defined to be a Web-based database (repository) of scholarly material which is institutionally defined (as opposed to a subject-based repository).”^[6]

Issues of copyright and permission to disseminate work are usually negotiated between the author and the publisher at the time of publication; many publishers have specific policies and guidelines regarding how an author may alter and/or disseminate versions of his or her work after publication. For instance, some publishers allow authors to retain copyright to their work but negotiate from the author the one-time rights, exclusive for a given amount of time (e.g., 90 days), to be the first publisher of the work, and the right to distribute the work exclusively for a given amount of time (e.g., 90 days) after publication. After this predetermined period (e.g., 90 days), the author may distribute, reprint, etc., as he or she pleases. On the other end of the spectrum, some publishers have much stricter and more limited agreements with authors wherein the author transfers copyright to the publisher in perpetuity.

Regardless of the copyright issues involved and the myriad possibilities for agreements between the publisher and the author, it remains that many different versions of any given article do and can exist. To focus on the journal article for purposes of explication, one journal article can exist in multiple electronic versions. In 2005, Sally Morris offered a partial list of possible versions of one journal article:

1. Privately circulated early draft (could be >1 iteration)
2. Version presented at public event (again, could be >1)
3. Pre-submission version(s)
4. Version as submitted to journal x (may differ when resubmitted to journal y)
5. Version amended after peer review (may go through >1 round of amendment)
6. Version as accepted by journal x
7. Accepted version, with substantive editing by journal editor and/or publisher (again, may be multiple iterations)
8. Accepted version, with substantive editing and copyediting—ready for publication
9. Publication version (as above, formatted and paginated)—proof
10. Publication version corrected and passed for publication
11. Published version, not on publisher’s site (e.g., PDF), thus potentially lacking some functionality
12. Published version (on publisher’s site, with full functionality)
13. Post-publication version with errata/addenda (maybe on publisher’s site, with functionality) or elsewhere without it^[7]

VERSION OF RECORD AND OPEN ACCESS

While it is beyond the scope of this entry to explore the complicated world of open access (OA), it is impossible to discuss version control without addressing some of the major features of OA. Broadly, OA is the movement to make scholarly literature “digital, online, free of charge, and free of most copyright and licensing restrictions.”^[8] Morris does not explicitly refer to any particular publishing models in her list; she does not overtly refer to OA or toll-based models because version control issues prevail in both worlds. Indeed, journal articles published in the more traditional toll-based scholarly literature, or journal articles published in OA journals can be versions of record.

The NISO/ALPSP JAV Technical Working Group defines version of record (VoR): “a fixed version of a journal article that has been made available by any organization that acts as a publisher

by formally and exclusively declaring the article ‘published.’”^[4] Furthermore, the notes to the VoR definition explain:

The VoR may exist in more than one location (e.g., a publisher’s Web site, an aggregator site, and one or more repositories). That is, there may be more than one *copy* of a VoR but there is only one *version* of a VoR. In Functional Requirements for Bibliographic Records (FRBR) terms, there may be more than one *manifestation* or *instance* of a VoR, but there is only one *expression* of it.^[4]

In other words, VoR transcends notions of multiple copies. Yet, the increasing prevalence OA archives and repositories have ultimately complicated the issue of version control for the scholarly community. The mere existence of such OA material (and other toll-based copies such as those that exist in aggregator databases) requires the reliance on FRBR differentiation between manifestation and expression.

Version control faces two distinct sets of challenges: political and technological. OA itself is an increasingly political movement, especially in light of the Consolidated Appropriations Act, 2008, which includes a provision directing the National Institutes of Health (NIH) to require “scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central upon acceptance for publication.”^[9] This Act was signed into law despite lobbying efforts of the Partnership for Research Integrity in Science and Medicine (<http://www.prismcoalition.org>), which is a coalition that has the support of the Professional and Scholarly Publishing Division of the Association of American Publishing.

In addition to the complexities that OA introduces, version control often resembles a Gordian knot because divergent groups have a vested interest in insuring that their voices are heard in any effort to create standardized solutions. Publishers, authors, researchers, vendors, aggregators, subscription agents, OA advocates, universities and other research centers, librarians, standards organizations—while there is some consensus of opinion among this variegated list, each group nonetheless represents a particular constituency with unique goals and interests. The JAV Technical Working Group’s report lists the following stakeholders as important variables in identifying the dimensions of an article version: author, editor, referee, publisher, librarian, reader, and funder. Other dimensions articulated by the JAV Technical Working Group are time, added value, manifestation, and siblings.^[4]

PEER REVIEW

In 2000, a working group of publishers acknowledged one outcome of the version control problem as changing the context within which peer-reviewed articles are perceived: “The peer-reviewed article will continue to play a crucial part in the certification, communication, and recording of scientific research. However, in the electronic environment it represents one point on a potential continuum of communication.”^[10] Peer review—in one form or another—has represented the bedrock of reputable scholarly output since at least the beginning of modern science in the seventeenth century, and any discussion of a potential change in its role in scholarly communication is significant.

The issue of peer review as it relates to OA and thus version control is a volatile one, with information industry stakeholders arguing many sides of this complicated issue. Some say that peer review in OA scholarly output is as rigorous and honest as it is within conventional toll-based publishing; others argue that peer review and thus quality of scholarship suffers in the face of OA. In her 2005 article, “Version Control of Journal Articles,” Morris noted the emerging problem of definitive or authoritative versions of documents versus “good enough” versions and its implications for scholars. The problem, as described by Morris, is pernicious. As listed earlier, she identifies 13 basic variants, from privately circulated early drafts, through post-publication versions with errata and addenda and argues that a reader cannot easily discern differences between these versions.^[11]

Steven Harnad addresses the worries of researchers, university administrators, and librarians in terms of version control and peer review. He notes that there is no need to be concerned about self-archiving because:

There will be self-archived preprints, revised drafts, final accepted, published drafts (postprints), updated, corrected post-postprints, peer comments, author replies, revised second editions. OAI-compliant Eprint Archives will tag each version with a unique identifier. All versions will be retrieved by a cross-archive OAI search, and the 'hits' can then be identified and compared by the user to select the most recent, official or definitive draft, exactly as if they had all been found in the same index catalog.^[12]

OAI is the Open Archives Initiative, and it “develops and promotes interoperability standards that aim to facilitate the efficient dissemination of content.”^[13]

APPROPRIATE COPY AND OPENURL

In addition to the implications for readers, Morris observed that version control is of great concern to libraries, which are faced with the multifaceted problems of tight budgets; appropriate copies based upon institutional affiliation; multiple, free online versions; and versions which are either self-archived or archived by independent institutions or other repositories. In response to some of these issues, libraries point users to their appropriate copy via link resolvers and the OpenURL framework. It is important to distinguish between the version control and the appropriate copy problem. The appropriate copy problem has two equal dimensions: 1) the multiple availability of any given resource from a variety of sources and 2) the subsequent challenge of pointing the reader to the copy of this resource appropriate for him or her, usually by virtue of his or her institutional affiliations.

If restricted to the toll-based universe, a given journal article might be available from the publisher or from a number of secondary aggregators (e.g., EBSCO, ProQuest, Gale). A researcher at a university may have access to only some of these article copies, and the library or information center wants to insure that the researcher reaches those copies which are appropriate for him or her. If expanded to include the OA universe of materials, the phrase “appropriate copy” then comes to “denote the copy of a resource to which a user has rightful access, usually by virtue of his or her institutional affiliations but perhaps also by virtue of an information object being *open access*.”^[14]

The OpenURL framework was developed to address the appropriate copy problem, and it is now a NISO standard, Z39.88, OpenURL v. 1.0. Via link resolvers, the OpenURL framework is widely used to point users to their appropriate copies. The appropriate copy problem and its solution in OpenURL/link resolvers primarily address a problem of access; version control, on the other hand encompasses issues of authority and management as well as defining the stages in the life cycle of an article. OpenURL also primarily focuses on pointing users to versions of record, which, again, can be in multiple locations. The other versions of an article—such as more ephemeral ones kept in an author’s institutional repository, for example—are currently outside the technical capability of the OpenURL/link resolvers. With version control, many technological hurdles remain, including: a standardized schema to identify the VoR or authoritative version as well as an extensible model to identify and explain the relationships among the different iterations of a document.

TECHNOLOGICAL INITIATIVES

One such effort to clearly articulate the different stages in the life of an article as well as the relationships among those stages was released in an April 2008 report by a joint partnership between NISO and ALPSP, the JAV Technical Working Group. Perhaps the most definitive work to date on the issue of version control, the recommended practices released by the JAV Technical Working Group attempts to “provide a simple, practical way of describing the versions of scholarly journal articles

that typically appear online before, during and after formal journal publication.”^[4] The working group builds upon earlier, previously mentioned efforts.

In September 2005, NISO and ALPSP charged the group with creating a plan that included:

1. Creation of use cases to identify the most common journal article life cycles.
2. Analysis of use cases to determine common life cycle stages.
3. Selection of preferred vocabulary for the most common life cycle stages.
4. Development of appropriate metadata to identify each variant version and its relationship to other versions, in particular the definitive, fully functional published version.
5. Establishment of practical systems for ensuring that the metadata is applied by authors or repository managers and publishers.^[15]

Cliff Morgan, Vice President, Planning and Development and Director, John Wiley & Sons, Ltd., was the chair of the joint NISO/ALPSP working group. In an August 2007 interview, Morgan noted: “In the digital world, multiple versions of journal articles are often available online. This can cause confusion because there is no way of identifying the various versions by either a common terminology or identification scheme.” Morgan went on to explain that the working group “will recommend terms and definitions for journal article versions and define the relationships between these versions. We’re focusing on key stages rather than every possible iteration of an article from origination to publication.”^[16]

Critical to the working group’s efforts is the notion that the group focuses on “key stages” of the version continuum rather than “every possible iteration.” The group considered a variety of attributes in order to describe different versions: “Ownership, bibliographic context, identifiers [e.g., digital object identifier (DOI)], relationships, fixity, and peer review are explicitly stated in the terms and definitions that we recommend. They can be described by the article version names and some are already covered by standard metadata elements (e.g., bibliographic reference, date, and DOI).”^[4] Of particular note is the attention paid to the version of record. Morgan elucidated the importance of this stage:

The VoR constitutes the “minutes of science”—the formally certified record of a research project. It is this version that has been peer-reviewed, edited, composed, and verified by the author; it is this version that benefits from publisher investment in managing the above process, building the journal brand, and adding functionality such as linking and e-alerts, and for which the publisher takes legal responsibility. This is also the version that is most likely to be cited by other researchers.^[16]

If the VoR is easily identifiable, then the problem of researchers questioning which version is the final, authoritative one is solved.

The seven stages identified by the working group are

1. Author’s Original
2. Submitted Manuscript Under Review
3. Accepted Manuscript
4. Proof
5. Version of Record
6. Corrected Version of Record
7. Enhanced Version of Record

Each stage is accompanied by a full definition and note. Also included in the report are graphical representations of the relationships among formal and gray literature as well as a number of “use cases” to provide a clearer picture of how the recommended terms would apply in real-world situations.

What remains to be seen is to what extent the recommendations in the report will be implemented by publishers and other stakeholders in the information industry. Version control is still

very much an issue in flux, and thus it presents interested parties with a moving target. Some publishers have already created and maintained internal practices for version control that may or may not be consistent with the group's recommended practices. If a publisher has invested significant capital in creating and maintaining an internal schema, then only time will tell if the said publisher will implement the NISO/ALPSP recommendations. Additionally, while the report explains some relationships among formal and gray literature, the group primarily concerned itself with a fairly narrow scope: journal articles. The continued evolution of the formally published journal article itself presents problems. If the electronic version with an assigned DOI is the version of record, then it begs the question about the necessity of assigning volume, issue, and page numbers for a journal article version of record. Questions of version control for other formal literature—book chapters, proceedings, etc.—remain. Moreover, an ocean of gray literature exists. With the advent of electronic publishing, gray literature and other informal publishing enjoy unprecedented accessibility. Such accessibility arguably leads to formerly “hidden” scholarship now occupying a more prominent place in our intellectual space and community of scholarship. Ultimately, many questions remain and while the group's work is essential, it is only one step toward an unknown future.

Version control is an international issue and as such, there are multiple initiatives underway to address its unique challenges. Two other programs created to study version control are the United Kingdom-based Joint Information Systems Committee Validating Repository Content project (VALREC, http://www.jisc.ac.uk/whatwedo/programmes/programme_rep_pres/tools/valrec.aspx and <http://valrec.eprints.org/>) and the London School of Economics VERSIONS Project (<http://www.lse.ac.uk/library/versions/>).

CONCLUSION

Version control remains a thorny issue with many technological and political barriers. The very nature of electronic information allows for an ease of replication and adaptation heretofore unheard of in scholarly communication. However, information industry stakeholders have recognized the issue and have formed a variety of international projects to study and recommend standardized solutions for identifying and managing iterations of a document for archival and retrieval purposes.

REFERENCES

1. Siemens, R.; Best, M.; Burk, A.; Grove-White, E.; Gue'don, J.-C.; Kerr, J.; Pope, A.; Rockwell, G.; Siemens, L. The Credibility of Electronic Publishing: A Report to the Humanities and Social Sciences Federation of Canada. 2002, <http://web.mala.bc.ca/hssc/Final/Credibility.htm> (accessed December 2007).
2. Burk, A.; Kerr, J.; Pope, A. Section 4: Archiving and text fluidity/version control. In *The Credibility of Electronic Publishing: A Report to the Humanities and Social Sciences Federation of Canada*. 2002, <http://web.mala.bc.ca/hssc/Final/Archiving.htm> (accessed December 2007).
3. United States Government Printing Office. Office of Information Dissemination, Program Development Service; Version control: Draft. 2005, http://www.access.gpo.gov/su_docs/fdlp/pubs/proceedings/05spring/cdqa/version_control_white_paper.pdf (accessed December 2007).
4. National Information Standards Organization/Association of Learned and Professional Society Publishers, Journal Article Versions Technical Working Group. Journal Article Versions (JAV): Recommendations of the NISO/ALPSP JAV Technical Working Group. April 2008. <http://www.niso.org/publications/rp/RP-8—2008.pdf> (accessed November 2008).
5. Driscoll, M.J. The MASTER project: Defining standards for electronic manuscript catalogue records. In *Care and Conservation of Manuscripts 6*, Proceedings of the Sixth International Seminar, Royal Library, Copenhagen, Denmark, October 19–20, 2000; Fellows-Jensen, G., Springborg, P., Eds.; Museum Tusculanum Press: Copenhagen, Denmark, 2002, <http://books.google.com/books?hl=en&id=iz17G5pKB1AC&dq=care+and+conservation+of+manuscripts+6&printsec=frontcover&source=web&ots=G8zvpZ8zW0&sig=oI8B6q7UHIYC0p45QkN1FYU4C7w#PPP1,M1> (accessed December 2007).
6. Bailey, C.W., Jr. Institutional repositories, Tout de suite; 2008, <http://www.digital-scholarship.org/ts/irtoutsuite.pdf> (accessed January 2008).

7. Grogg, J.E. Linking users to open access. *Searcher* **2005**, *13* (4), 52–56.
8. Suber, P. Open access overview; 2007. <http://www.earlham.edu/~peters/fos/overview.htm> (accessed January 2008).
9. U.S. Department of Health and Human Services. National Institutes of Health Public Access. <http://publicaccess.nih.gov/> (access November 2008).
10. Antelman, K. Self-archiving practice and the influence of publisher policies in the social sciences. *Learn. Publ.* **2006**, *19* (2), 85–95.
11. Morris, S. Version control of journal articles; 2005. <http://www.niso.org/workrooms/jav/Morris.pdf> (accessed November 2008).
12. Harnad, S. Open Access to peer-reviewed research through author/institution self-archiving. In *Digital Libraries*; Andrews, J., Law, D., Eds.; Ashgate: Hants, U.K., 2004; 63–98.
13. Open archives initiative. <http://www.openarchives.org/> (accessed January 2008).
14. Grogg, J.E. Introduction linking and the OpenURL. *Libr. Technol. Rep.* **2006**, *42* (1), 5–7.
15. National Information Standards Organization. NISO/ ALPSP Working Group on Versions of Journal Articles. <http://www.niso.org/workrooms/jav> (accessed November 2008).
16. *Library Connect*, Editorial Office. Five quick questions with Cliff Morgan. *Elsevier Libr. Connect Newsl.* **2007**, *5* (3), 9. <http://libraryconnect.elsevier.com/lcn/0503/lcn0503.pdf> (accessed January 2008).

21 Information Systems Failure

Chris Sauer and Gordon B. Davis

CONTENTS

Introduction.....	285
The Scale of IS Failure	286
The Evolution of IS Performance	286
Emerging Cognate Literatures	288
Approaches to Understanding IS Failure	289
Static Approaches.....	289
The Technological Perspective.....	290
The Organizational/Cultural Perspective.....	290
The Political Perspective	292
Interactionist Approaches.....	292
The Sociotechnical Perspective.....	293
The Exchange Perspective.....	293
Process Approaches.....	294
Actor-Network Theory	294
Conclusion: The Future of Failure	294
Conclusion: Summary of Lessons about Failure for Practitioners	295
Acknowledgment	296
References.....	296

INTRODUCTION

Information systems (IS) failure has been a focus for study and action for both practitioners and academics since the late 1960s.^[1–3] Most organizations have struggled to deliver systems that work and deliver benefits. Practitioners have wrestled with these difficulties in their workplace and the public has experienced systems’ shortcomings in trying to use them. Nobody seriously doubts that failure has been and continues to be a problem to be tackled.

The focus on failures does not mean that there are two classes of applications—those that succeed and those that fail. There is a continuum of results ranging from clear and utter failure to very high success. There are applications that are barely acceptable but survive, and there are applications that are mildly successful and fail.

This entry reviews only one end of the spectrum of results: applications that are failures. The meaning of “information system success” is a significant issue^[4] in itself and is not the purpose of this entry. Information system failure and information system success are not simply opposites.^[5] To achieve success, it is necessary to remove causes of failure, but doing so may not be sufficient for a successful outcome. More broadly, the study of failure is justified by the belief that, just as the study of cognitive malfunction can help us understand good mental functioning, it casts light on what it takes to achieve success. Understanding “information system failure,” therefore, represents one but only one part of what needs to be understood for the effective design, implementation, management, and exploitation of an information system.

Where once “failure” was viewed as an almost objectively identifiable state, today there is wide recognition that it is an evaluation based on an assessment of some dimension or dimensions of performance relative to goals, targets, or expectations. Failure in one dimension, say a budget overrun, may be more than compensated for by success in another dimension, say timely implementation. And, as what is regarded as performance has itself changed, so the domain of failure studies has evolved. For this reason, it is appropriate for this entry to take an historical approach.

In this entry, the study of failure is treated as one important branch of the study of IS performance. Our overall objective is to provide a rigorous basis by which to understand what may be meant by “IS failure” and a set of signposts to appropriate analytical approaches by which to develop greater insight into the phenomenon.

We start with a review of recent performance data. We examine the various changes that have affected perceptions of IS performance. We identify a number of cognate subliterations that have developed over recent years and which signal a growing maturity of analysis. We provide a framework by which to understand the historical development of academic research in this field.

THE SCALE OF IS FAILURE

Prior to 1995, most assessments of the scale of IS failure were informal in nature and anecdotally derived. In 1995, the Boston-based consulting group, Standish, published its findings under the banner headline, *Chaos!*^[6] Based on data about more than 3000 IS projects, it reported that 31% were abandoned without being completed. These it dubbed failures. Some 53% failed to meet at least one of its budget, schedule, and scope targets. These were dubbed challenged projects. The remaining 16% that hit all three targets were deemed successes. These have since become the most widely cited failure statistics.

Every 2 years since 1995, the Standish Group has released updates on these figures. Figures released for 2004 revealed 29% hit all targets, 53% were challenged in missing at least one target and 18% were classified as failures.^[7] In summary according to the Standish Group figures there has been sustained improvement but there remains continuing cause for concern. While some question the Standish survey and analysis methodology,^[8] raising questions about the accuracy of the detailed findings, in 1995 few doubted the broad conclusion that most IT projects seriously underperformed. Subsequent surveys reinforced the point (e.g., Taylor).^[9] Likewise today, more than a decade later, the improving trend they identify is commonly acknowledged.

Less clear is the extent of the improvement and therefore the extent of the outstanding problem. To measure this, it is better to use percentage variance from targets.^[10] Otherwise we cannot tell whether “challenged” projects are more or less challenged over time. Sauer et al.^[10] also argue that an absolutist definition of success by which any variance however small assigns a project to the challenged category is too strict. In most cases a small budgetary or schedule overrun is of negligible significance. Using statistical clustering techniques, they show that there is a significant gap between the majority of projects that miss their targets by just a small percentage margin and those that are clearly failing on a more substantial scale and that therefore it is worth tolerating some small margin of variance. However, even on this more optimistic assessment, at least one-third of all projects fall well short of delivering on all their targets.

THE EVOLUTION OF IS PERFORMANCE

In the previous section we saw that IS performance has changed over time and that the metrics by which performance in general and failure specifically are identified are not definitively agreed. In this section we explain how the phenomenon itself has changed and explore the implications for the measurement and identification of failure.

We start with a distinction between systems and systems projects. In the 1960s and 1970s, the system and the project were often conflated. If a system failed, say because it did not deliver the

functionality its users wanted, it was usually the case that the project had been troubled too and vice-versa. This was also a period in which computer malfunctions were common resulting in breakdowns ranging from the everyday experience of the computer being “down” to catastrophic loss. Such malfunctions have been extensively cataloged by Peter Neumann in his Inside Risks column in the Communications of the Association for Computing Machinery over many years.^[11] While more reliable hardware and standardized and re-usable software have removed many causes of such malfunctions, they remain a continuing hazard (e.g., Hodgson)^[12] but today are rarely such as to cause abandonment of a system.

With the advent of rigorous project management disciplines it has become normal to detail precise targets, in terms of budget, schedule, quality, and scope. These are essentially internal measures that in principle are within a project manager’s capacity to control. It has become apparent therefore that a project may be successful in meeting its internal targets yet not deliver beneficial business outcomes. More recently, as the role IT plays in organizations has become critical to their functioning as exemplified by the widespread adoption of e-business for front office business functions, there have been cases of projects failing to hit budget and schedule targets but nevertheless delivering highly desirable business outcomes. The distinction between the system and the project is therefore both conceptually sensible and practically relevant.

What counts as an information system and therefore what is involved in an IS project has changed over time and this has affected the measure of performance. Purpose-built systems have given way to standard packages (so-called enterprise software). Projects are therefore more about system integration than software development. This has led to two important consequences. First, it has engendered more ambitious, more complex projects that combine packaged functionality to an extent that was previously unimaginable. Second, because the software is prewritten, it has encouraged organizations to direct project managers’ efforts to implementation, business process change, and value harvesting. This latter has extended project scope considerably and turned managers’ attention to an external target, the business value achieved by investment in an information system.

This shift in emphasis from internal to external targets, from assessing performance against process to performance on outcomes, is appropriate from a commercial investment perspective. It is, however, problematic to measure because IT systems do not deliver value on their own, they do so through organizational and business process change.^[13] Even if the value itself is measurable and measured, and it often is not, the extent to which that value derives from the information system may not be possible to determine. Further problems include the question of when to measure value outcomes. Often immediately after implementation, organizational performance suffers^[14] before recovering and improving. But the longer the lag between implementation and outcome measurement, the more space there is for extraneous variables to affect the value outcome.

Thus far in this section, we have examined how IS performance has evolved in practice. We now turn to how its conceptualization has evolved academically. A significant breakthrough came with the explicit recognition that “failure” is an evaluation not an objective description. Lyytinen and Hirschheim^[15] developed a new pluralist concept of expectations failure which explicitly recognized that failure was a judgement relative to expectations and that expectations vary among stakeholders. This pluralism opened the way to the possibility that a single system or project could be a failure relative to one set of stakeholders and a success relative to others. It also triggered discussion of whether some stakeholders’ expectations might weigh more in arriving at a single overarching evaluation—this is “on balance” a failure notwithstanding some less important stakeholders viewing it as a success. Several responses have emerged. Some have tried to identify apparently unequivocal failures such as abandoned projects (e.g., Johnson^[6] and Ewusi-Mensah^[16]) though it has been argued that even an abandoned project might be viewed by some stakeholders as a success if an objective was to learn about a new technology, test customer responses, or experiment in some other way.^[17] Sauer^[17] advanced an uneasy compromise that recognized the plurality of stakeholders but recognized that a distinction should be made between supporters of a project and those

who were uninvolved or active opponents. Others have finessed the issue by focusing on specific dimensions of performance such as budget and schedule variance without using explicitly evaluative terminology.^[18–20] In doing so, by the very nature of their choices of performance dimensions to study including the recent introduction of business value as an outcome measure, they can be argued to implicitly favor the expectations of traditional stakeholders such as the owners and managers of businesses. By contrast, others have embraced pluralism and absorbed it into new theoretical approaches such as Actor-Network Theory, which we discuss below.

We can summarize this section by saying that the conceptualization of failure has evolved and matured in three ways. The first has been the distinction between systems and projects. The second has been the expansion of the legitimate scope of IS projects with consequences for what dimensions of performance are assessed. The third has been recognition of the evaluative nature of “failure.” This has surfaced disagreement as to which stakeholders’ expectations should receive priority. This is a political disagreement rather than one of facts or data. Those who believe that owners and manager as the agents of owner matter more will adopt a more traditional view of failure whereas those who believe that employees, the public, and others have just as much right to their interests and expectations being respected will take a more egalitarian and pluralist position.

EMERGING COGNATE LITERATURES

A further manifestation of the maturation of our understanding of IS failure has been the emergence of several cognate literatures. We briefly discuss in turn risk, partial abandonment, expectation management, escalation, and drift.

Risk as discussed in the extensive IS and software engineering literatures relates to failure by seeking to identify what may be conducive to failure, how to assess its probability and impact, and how to manage it. It is about the probability of failure in some shape or form. It has typically been strongest at identifying sources of risk such as starting conditions or behaviors that give rise to failure.^[21] Such studies usually result in extensive lists. The usual suspects thus identified typically include: unclear objectives, uncertain requirements, inadequate project management, too little user involvement, and shortage of top management support. The literature has been less convergent in its focus on risk assessment. While there is an in-principle agreement that risk consists of two elements (the probability of some outcome and its impact), the determination of what the actual risk amounts to has proved problematic. Some have focused on the impact of total failure as represented by abandonment,^[22] while others have looked at specific dimensions of performance such as running over schedule or budget. Some studies while recognizing that impact is an element of risk focus principally on the probability of underperformance.^[10] The management of risk, and therefore the prevention of failure, has principally consisted of attending to risk factors to prevent their occurring or mitigating their effect when they do occur.

A further set of cognate studies is that on partial abandonment.^[23,24] This recognized that a common strategy for coping with underperformance, such as projected schedule or budget overruns, or unanticipated events is to descope, which is to say abandon, some elements of a project. It advanced our thinking by showing that what might simplistically be viewed as a symptom of failure can also be an instrument for preventing failure or further failure. Failure is not a simple cut and dried function of any single indicator.

The focus on expectations failure brought into focus the idea that success and failure are a function of the targets as well as the delivery and performance of a system. That is to say that one way of preventing failure is to manage expectations closer to what is realistically achievable. Ginzberg^[25] pioneered this notion. It has received limited attention subsequently^[26] although, influenced by Danish research,^[27] the United Kingdom’s Office of Government Commerce today mandates that all projects control for “strategic optimism,” the tendency to set expectations unduly high.

A more abundant cognate literature has been that on escalating commitment to a failing course of action. It addresses the key question of why organizations do not cut their losses sooner.

Inspired by a substantial corpus of work in Organizational Behavior,^[28–32] and recognizing that many IS projects have been characterized by sponsors pouring in good money after bad to projects that would ultimately fail, researchers have explored the applicability of existing theory and sought to adapt it to the IS failure phenomenon.^[33–35] While escalation of resourcing is indeed a common feature of projects and existing research enriches our understanding of why it occurs, its practical value is limited by the question of whether we can adequately recognize failure in advance and the extent to which escalation often genuinely sustains a project that would have collapsed without it.

Finally, the “drift” literature introduces a further recognizable phenomenon associated with project performance and which may often either result in failure or substitute for it. Instigated by Ciborra,^[36] the idea is that objectives are not immutable fixtures in a project but are subject to organizational politics that often neutralizes them without radically reviewing the case for the project or system. The resultant drift leaves organizations with IS investments that do not meet their real needs but which are not subjected to the scrutiny that might lead to their redevelopment or abandonment. Empirical exploration and development of this phenomenon is in its infancy.^[37,38]

So, while risk research has worked to connect failure to management practice with a view to reducing its probability, the other cognate literatures have functioned to situate failure within a more sophisticated understanding of organizational contexts and what counts as success within them. Specifically, partial abandonment recognizes that failure on one dimension of performance, namely scope, may be sacrificed to achieve better performance on another such as budget. Expectation management extends this more realistic approach to what will count as good or bad performance by recognizing that targets in the form of stakeholder expectations are legitimately subject to influence and change. Escalation theory is an inverse of partial abandonment in that it shows why larger and larger overruns on resourcing such as budget may be promoted in the interests of achieving scope and outcome expectations, and how this can be dysfunctional. The focus on drift brings a recognition that not all systems and projects are subject to clear-cut evaluations on a continuing basis and that over time objectives and expectations may be diluted through organizational political processes without any re-evaluation occurring. Failure studies if they are to embody a realistic and sympathetic understanding of the richness and complexity of organizational life must embrace the insights that these cognate studies have brought to the table.

APPROACHES TO UNDERSTANDING IS FAILURE

As our understanding of the concept of IS failure has evolved, so too have our approaches to analyzing the causes of failure. We can distinguish three broad approaches. At one extreme, we have static analyses based on the identification of factors that predispose to failure. At the other extreme, the emphasis is on dynamic analysis and the explication of processes by which failure comes about. Bridging the two is the interactionist approach.

STATIC APPROACHES

Static, factor-based approaches were at their most popular in the early days when failure was seen more as an objective phenomenon. It was easy then to ask what caused failure and to look for factors that “determined” failure. What were seen in retrospect as common features of failures rapidly became accepted as causes which if rectified would lead to improved performance and outcomes. Research was undertaken to validate this view (see Larsen^[39] for a compendious review of factors identified in the research).

Within this class of approaches, several versions developed emphasizing different perspectives and therefore different failure factors. We identify three here: the technological perspective, the organizational/cultural perspective, and the political perspective.

The Technological Perspective

A distinguishing feature of the technological perspective has been its strong focus on the technology elements as the most important considerations for an information system (and often the only components to consider in building the system). The efficient processing of data and the efficient operating of hardware and software provide the goals around which people and procedures are to be organized. Studies in this perspective have often embodied two important assumptions. The first is that technology in all its forms (including hardware, software, and databases) is considered to be “good.” The “imperative” in technological imperative refers to the inevitability of technology to “compel” acceptance, use, and conformity to it.^[40] Any resistance to a given technology is regarded as merely an irrational, impermanent phenomenon. The second is that implementation of technology alone is sufficient to secure organizational benefits. Together these assumptions conspire to exclude other factors as relevant to success and failure.

The technological perspective therefore implies a dominant focus on formal development procedures for systems analysis and design, the purpose of which is to build or upgrade an information system. There are four major phases in the development cycle. The first phase consists of requirements determination to identify the output of the hardware–software system and the data to be input and processed. The second phase is conceptual design to create a “logical” or conceptual blueprint for the system. The third phase is the “physical” or implementation design for specific software and hardware needed in the information system. Building or upgrading the hardware–software system follows. All formal system development methodologies, as well as the tools of computer-aided software engineering that implement them, involve these four phases or variations of them. The technological perspective tends to underlie not only most development methodologies but also most textbooks on programming, systems analysis, systems design, database management systems, and telecommunications.

In the technological perspective, the diagnosis of an information system failure focuses on a hardware failure, a software failure, a data failure, or any combination of these. The chosen hardware platform may have defects in its manufacture or configuration. The software may suffer from poor design or error-prone construction. The database may harbor inconsistent, incomplete, or out-dated data. Since the technological perspective assumes that users can correctly and completely define processing requirements, a failure in the system may also arise because the requirements were specified incorrectly or incompletely. For instance, the information system could lack the capacity to support the processing load if this were not specified correctly.

In general, the remedy for a potential information system failure based on the technological perspective view is for people to conform to the development cycle procedures and to follow processes specified for operation of the system with its hardware, software, and database. The technological perspective requires that people (analysts, designers, managers, programmers, end users, etc.) do or redo the steps of the systems analysis and design correctly. Any failure by people to respond appropriately (e.g., users who do not willingly participate in the requirements determination step, users who resist using the installed system, designers who neglect to install adequate backup procedures, or designers who themselves resist using computerized tools) is seen as a failure of developers and users and not a failure of technological solutions to organizational problems.

The Organizational/Cultural Perspective

By contrast with the technological perspective, in the organizational/cultural perspective it is technology that must adapt to people and to organizations. An organization, in this view, is more than a collection of people. An organization is a structure of roles and rules in which the rules delimit the aims, resources, prerogatives, actions, and behaviors allowed to the person occupying a role. The “culture” in organizational culture refers to the organizational members’ shared knowledge or shared understanding of the structure of roles and rules. As successive generations of participants enter, pass through, and leave an organization, its structure of roles and rules continues to exist and to evolve slowly. Upon entering an organization, an individual becomes subject to the organization’s

structure of roles and rules. Different individuals, when placed in the culture, tend to display similar actions and behaviors. Thus, the culture is the ongoing realization, albeit often partial, of the formal organizational design. An information system application that reflects the desires of one person or a group of people but ignores organizational culture is unlikely to fit the way the organization works.

Preceding the installation of a new information system or the upgrading of an existing one, the organizational/cultural perspective emphasizes the existence of beliefs and assumptions shared among the organizational members, about what computers do, how technology affects jobs, why the current information system is being upgraded or replaced, and so forth. When held by people in contemporary organizations, these beliefs and assumptions operate much like myths, metaphors, and magic among “natives” in preliterate societies.^[41] In this view, an information system is but one among many artifacts crafted by organizational natives for use in their daily activities and rituals. Hence, in the organizational/cultural perspective, an information system failure is a failure of this artifact to serve and be compatible with the activities and rituals of the organizational natives.

In the organization culture of information technology, a poor fit between an information system (the hardware, software, and data) and the organizational culture is conducive to an information system failure. A framework useful for identifying a poor fit between information system and organizational culture specifies seven categories of shared beliefs and assumptions that can be violated by systems:^[42,43]

1. Shared beliefs and assumptions about innovation and action (e.g., Are original and creative behaviors “good?”)
2. Risk taking (e.g., Is the norm to be proactive and take responsibility for one’s own actions, or to be reactive and avoid responsibility?)
3. Integration (e.g., How much importance is placed on communicating and cooperating with other organizational members?)
4. Top management contact (e.g., Does top management customarily relate to subordinates by being affirming or openly critical?)
5. Autonomy in decision making (e.g., Are organizational members allowed, encouraged, expected, or trusted to make decisions on their own?)
6. Performance orientation (e.g., Are the criteria for compensation, job security, and job promotion concretely established, explicitly known, and actually used?)
7. Reward orientation (e.g., Do rewards take the form of a “carrot” or “stick?”)

In this framework, the failure potential for an information system increases with the number of categories of cultural assumptions with which the information system is incompatible, as well as the extent of the respective incompatibilities. For instance, consider an information system that makes an assumption that there is open sharing of data across the organization’s different divisions (integration) and that new users of the system will take responsibility for the decisions that they make with the help of the information system (risk taking). If the organizational members share a norm in which each division guards its data as a valued resource and if the organizational members share, as an accepted practice, the custom of “blaming the computer” for their bad decisions, there is a significant threat of failure for the new system. Any additional incompatibilities between the information system and other shared beliefs and assumptions would increase the threat of an information system failure.

To prevent or to remedy an information system failure, the organizational/cultural perspective requires that the information system be designed and adjusted to respect the constraints that the organizational culture imposes. A strong proponent of this view will maintain that an information system is an artifact crafted to serve the activities of an organization, and therefore the information system should always conform to the organization, rather than vice versa. In other words, a proponent of the organizational culture view may take the position that altering an organizational culture to fit an information system, even if possible in the long run, is a dangerous course of action.

The organizational/cultural perspective extends the technological perspective by bringing in organizational/cultural factors. It also introduces the idea of a fit between the technology and the organization/culture. However, it continues to assume that the important challenge is to establish the right conditions in terms of getting two sets of factors aligned with each other.

The Political Perspective

The political perspective extends the organizational/cultural perspective by introducing power as the real basis by which organizations work. Power is the ability of a person, group, or organizational unit to overcome resistance or opposition to achievement of its ends; politics is the exercise of this ability.^[44–46] In this view, activities that accompany the development, implementation, and use of an information system provide an arena in which a person, group, or organizational unit can exercise power, whether or not the ends for which the power is exercised bear any relevance to the purposes that the information system is ostensibly said to serve. Success or failure of the information system relative to these ostensible purposes is secondary to the achievement of the power objectives.

In a classic case study of an information system implementation,^[47] a new information system threatened to alter the power between the corporate accounting group and divisional accounting personnel. The system was designed to allow the accounting group at corporate headquarters immediate access to data belonging to the accounting group in each of the company's different divisions. In the existing procedures, each division controlled its own data (stored in thick manual ledger books), and corporate accounting could only obtain it through reports prepared by division accountants. Built into the design of the new information system was a presumption that there should be a different distribution of power relative to the division data. This design was contrary to the existing power over the data. In the technological imperative view and the organizational culture view, resistance to the proposed system could be explained as opposition to a more efficient system or opposition to changes in the culture of corporate versus division roles. In the political perspective, the efficiency considerations or role changes were secondary; the main objective of the system for the corporate accountants was political. They wanted to disrupt the existing balance of power in the organization and introduce a new balance of power based on data access that favored the corporate group.

In the political perspective, the failure potential for an information system application increases if it violates the constraints imposed by the organizational distribution of power. In some cases, an application will be allowed to succeed only if it supports the political agenda of a powerful group within the organization. IS may be used explicitly to change the distribution of power in an organization; in such cases, the design of the system is a secondary factor in success or failure. The major issue is the effectiveness of change agents in negotiating redistribution or organizational power.

To prevent or remedy an information system failure due primarily to strong political opposition, the political considerations must be surfaced and negotiated. Political considerations may be somewhat hidden; surfacing them through high-level discussions allows the effect of the system to be dealt with in the context of high-level goals of the organization. An explicit understanding of the effect of the system on the relative power of the organizational participants may allow negotiations and compromise. In the absence of such compromise, political opposition to a system may be sufficient to cause failure.

The political perspective has sometimes been operationalized in research in terms of the fit between an information system or project and the power structure. Such operationalizations fall under the factor approach. Others however have recognized that power can be renegotiated and as such the processes by which IS are treated affect their success or failure.

INTERACTIONIST APPROACHES

Interactionist approaches bridge the step from factors to processes. In accepting that factors may influence each other, they implicitly introduce a temporal element to the analysis of failure. For example, a factors analysis might identify inadequate involvement of business users in requirements specification and user resistance as critical causes of failure. The interactionist recognizes that the

user resistance may be a function of the lack of involvement. In turn, perceived user resistance may deter technologists from consulting the users. The causes of failure develop and harden over time. Implicitly, such approaches recognize that the different factor approaches embody important elements that may contribute to a comprehensive analysis of failure. We briefly describe two such approaches: the sociotechnical perspective and the exchange perspective.

The Sociotechnical Perspective

In the sociotechnical systems view, an information system is more than a technical system of hardware and software. An information system is a social system that uses information technology.^[48,49] Failure is not a failure of the information technology alone or the organization alone, because there is a strong interaction between the two. The same information technology will exhibit different properties when used in different organizational contexts. Sociotechnical theory^[50] predicts that optimizing only the technical system or optimizing only the social system will lead to information system failure. To avoid failure, neither system can be managed separately from the other.

The interactive effects between the social system and the information technology have a parallel in the chemical difference between mixtures and compounds. The properties of a mixture are simply a combined set of the respective properties of the individual elements that constitute the mixture. The reason is that the elements in a mixture do not react to one another; their respective properties remain intact. However, the properties of a chemical compound are different from the respective properties of the constituent elements. The reason is that the elements react to one another; their respective properties are replaced by the new properties of the compound entity. In this analogy, an information system is more like a compound than a mixture. Its constituent social system and information technology react to one another, so that their properties in combination are different from their properties in isolation.

In the sociotechnical view, failure potential is increased by development processes and procedures that do not integrate both social and technical considerations. In sociotechnical development, the future users of the system specify important job and work design considerations. There is a joint effort involving both users and technical personnel to consider alternative information system technical designs and select one that achieves a reasonable trade-off between job design and technical efficiency. Dominance of technical considerations increases the probability of failure due to user resistance. The sociotechnical approach recognizes the fact that user resistance can cause a technically well designed system to fail and user enthusiasm can cause a somewhat mediocre technical design to succeed. These dynamics are the basis for a process that looks for the best combination of technical and social factors.

To prevent or remedy an information system failure, the sociotechnical view specifies a socio-technical requirements and development process. Users are empowered with sufficient information to be effective participants in design discussions and decisions. Obtaining commitment by users to make the system succeed is an essential part of the process.

The Exchange Perspective

The exchange perspective^[17] develops the sociotechnical perspective in one specific direction. It emphasizes one critical basis of interaction, viz. the creation and sustaining of a balanced exchange between a project team or those who manage an information system and the key supporters of that project or system. It argues that projects and systems fail if two conditions obtain^[1]: there are insufficient resources to continue developing or operating the system, and^[2] key supporters are dissatisfied.

This perspective sees systems and projects as needing a range of different forms of support from a range of stakeholders. Support can include material resources such as funds, equipment, space, practical support such as involvement in project processes, problem-solving, and management attention, psychic support such as publicly observable commitment, as well as organizational support in terms of politics and working the organizational system. Stakeholders who may be key supporters include senior management sponsors, operational users, end customers, the project team, and the

supply chain. As supporters' levels of satisfaction may ebb and flow over time, the task of IS management and project management is one of maintaining sufficient levels of satisfaction to ensure a continuing flow of needed resources. Thus, for example, the project process can be seen as the challenge of marshalling the right resources to deliver sufficient progress and expectation of benefit that supporters will continue to provide those resources. If progress stalls, resources needed to overcome a hurdle may not be forthcoming resulting in further difficulties for the project and further damage to supporters' expectations. The project team may be able to break out of this vicious circle by actions that restore confidence such as bringing forward the delivery of some benefits or by actively managing supporters' perceptions. If on the other hand it is unable to break out of this vicious circle, the project will be set for failure. Projects can be seen as an ongoing complex interactive process where exchange is at the heart of the interactions.

PROCESS APPROACHES

Since the mid-1990s failure analyses have emerged that focus more on a detailed examination of process and the search for patterns in process. They have tended to view success and failure outcomes as not determined by factors or even by interactions but as a function of a yet more complex account of the processes by which multifarious influences affect projects and systems over time. As yet no single perspective has come to dominate the field. However, we shall give special attention to Actor-Network Theory because it has been widely employed in social studies of IS including failures, and because it encompasses ideas from numerous other perspectives.

From within the IS discipline, Newman and Robey^[51] enabled process analyses of failure through their more general articulation of the elements of process analysis, and Newman has applied it in his failure research.^[52,53] Outside the discipline, in the broad area of social studies of technology, a number of case studies have been published that have cast light on IS failure through detailed accounts of IS processes.^[54] This has then been brought back into the discipline by critical theory-oriented researchers such as Howcroft, Wilson, and Mitev adopting a social shaping of technology perspective and applying it to failures.^[55,56]

Actor-Network Theory

Actor-Network Theory has recently become a popular tool for analyzing IS failure.^[36–38,57–59] Actor Network Theory has a number of attractions. First, it is the first approach in many years to take a balanced position in relation to technology. That is, where for most researchers over the last 15–20 years it has been accepted that technology is subservient to human/organizational objectives, Actor-Network Theory places technology on a par with other stakeholders. It then views the object of study, say a project failure, through the lens of the actor-network in which it is embedded. Humans, institutions, and technologies are among the stakeholders who populate the network and propagate influences through its nodes. In this manner, the complexity of the contexts within which IS failures occur is well represented. Further, one of the core ideas of the theory is that of breakdown. The essence of this insight is that when something appears to be functioning well we treat it as a black box. Only when it begins to fail do we start to unpack it to understand what is going on. Thus, for example, we can understand “drift” in these terms.

The limitations of Actor-Network Theory are that it is still evolving a set of concepts for general use. Early analyses using it have tended to be very complex and not productive of readily practical conclusions.

CONCLUSION: THE FUTURE OF FAILURE

Ten years ago, the question arose as to whether the increased commodification of IT would lead to standard solutions to such an extent that the problem of failure might dissolve.^[60] The alternative

scenario was that the strategic benefits of IT would be such that organizations would continue to seek to differentiate themselves and would therefore continue to pursue distinctive, non-commodity solutions. The improvements in performance are noted earlier in this entry but the continued difficulties in achieving expectations suggest that elements of both scenarios have occurred. This suggests that the issues associated with understanding and preventing failure will continue to be highly relevant for the foreseeable future.

The challenge for research is to strike a balance between analyses that embody every causal constituent in a failure but where it is difficult to see the wood for the trees, and those that are simpler but which promise too much to the practitioner because they suggest that a relatively simple intervention will be sufficient to prevent failure. Having developed over the last 20–30 years a growing maturity in our understanding it is easy for research to fall into the former category. It is this trick that Actor-Network Theory or alternatives will have to master if it is to be both academically successful and practically useful.

CONCLUSION: SUMMARY OF LESSONS ABOUT FAILURE FOR PRACTITIONERS

The entry has lessons for scholars and researchers. It also has lessons for practitioners. These can be summarized as follows:

1. The evaluation of success or failure for a project is not usually a simple “success” or “failure” or identification of a single reason for the result. Success or failure depends on the criteria being used, the stakeholders being considered, and the level of expectations.
2. Organizations respond to perceptions of failure in various ways. One response is abandonment. Other possible responses include escalation and revision of expectations.
3. The ideal lessons for practitioners would be a set of simple prescriptions to prevent failure and ensure success. The analysis of the failure research provides guidance but no simple list of prescriptions. A thoughtful practitioner seeking to reduce the potential for failure should consider both static factors that predispose to failure and the interaction of factors that may lead to failure.
4. Static causes. Three types of causes that predispose to failure and therefore should be considered as part of prevention measures are
 - a. Technological. The focus in failure prevention and failure analysis is on correct and complete hardware, software, or data (or a combination of these technological elements). Today, these are the least common causes of failure.
 - b. Organization/culture. The focus is on a good fit between the system and the activities, rules, roles, rituals, and culture of an organization.
 - c. Political. The focus is on incorporating the effect of political power on design, acceptance, and use of a system.
5. Interaction causes. The focus on static causes may not be sufficient. Factors may influence each other. Two approaches to understanding and analyzing failure based on interactions are
 - a. Sociotechnical. This analysis is based on the view that an information system is a social system that uses information technology. To avoid failure, the interaction of the two systems must be considered in the designs of both.
 - b. Exchange perspective. Systems and projects need forms of support from a range of stakeholders. Critical to avoiding failure is creating and sustaining a balanced exchange between a project team and key supporters.
6. An emerging view that is perhaps not ready for widespread practitioner use is to analyze the process by which a complex set of influences affect the way projects and systems develop over time.

ACKNOWLEDGMENT

This entry adds to and develops aspects of the entry on Information Systems Failure by Gordon B. Davis, Allen S. Lee, and Kathryn Nickles in the previous edition of the Encyclopaedia of Library and Information Sciences.

REFERENCES

1. Naur, P.; Randell, B. *Software Engineering: Report on a Conference Sponsored by the NATO Science Committee*; Garmisch: Germany, October 7–11, 1968. <http://homepages.cs.ncl.ac.uk/brian.randell/NATO/nato1968.PDF> (accessed January 11, 2007).
2. Colton, K.W. Computers and police: Patterns of success and failure. *Sloan Manage. Rev.* **1972**, winter, 75–98.
3. Lucas, H.C. Jr. *Why Information Systems Fail*; Columbia University Press: New York, 1975.
4. DeLone, W.H.; McLean, E.R. Information systems success: The quest for the dependent variable. *Inform. Syst. Res.* **1992**, 3 (1), 60–95.
5. Hirschheim, R.; Smithson, S. A critical analysis of information systems evaluation. In *Information Systems Assessment: Issues and Challenges*; Bjørn-Andersen, N., Davis, G.B., Eds.; Elsevier (North-Holland): Amsterdam, 1988; 17–37.
6. Johnson, J. Chaos: The dollar drain of IT project failures. *Appl. Dev. Trends* **1995**, January, 41–47.
7. Johnson, J. *My Life is Failure*; Standish Group International: West Yarmouth, MA, 2006.
8. Jørgensen, M.; Moløkken-Østfold, K.J. How large are software cost overruns? Critical comments on the Standish Group's CHAOS reports. *Inform. Softw. Technol.* **2006**, 48 (4), 297–301.
9. Taylor, A. IT projects: Sink or swim. *Comput. Bull.* **2000**, 42, January, 24–26.
10. Sauer, C.; Gemino, A.; Reich, B. Managing projects for success: The impact of size and volatility on IT project performance. *Commun. ACM* **2007**, 50 (11), 79–84.
11. Neumann, P. CACM inside risks, 2007, <http://www.csl.sri.com/users/neumann/insiderisks.html#menu2> (accessed January 11, 2007).
12. Hodgson, M. Computer error keeps airport on summer time. *Guardian Newspaper*; Monday October 29, 2007.
13. Markus, M.L. Technochange management: Using IT to drive organizational change. *J. Inform. Technol.* **2004**, 19 (1), 4–20.
14. Ross, J.W.; Vitale, M. The ERP revolution: Surviving vs. thriving. *Inform. Syst. Front.* **2000**, 2 (2), 233–241.
15. Lyytinen, K.; Hirschheim, R. Information systems failures: A survey and classification of the empirical literature. *Oxf. Surv. Inform. Technol.* **1987**, 4, 257–309.
16. Ewusi-Mensah, K.; Przasnyski, Z.H. On information systems project abandonment: An exploratory study of organizational practices. *MIS Quart.* **1991**, 67–86.
17. Sauer, C. *Why Information Systems Fail: A Case Study Approach*; Alfred Waller: Henley-on-Thames, 1993.
18. Faraj, S.; Sproull, L. Coordinating expertise in software development teams. *Manage. Sci.* **2000**, 46 (12), 1554–1568.
19. Nidumolu, S. The effect of coordination and uncertainty on software project performance: Residual performance risk as an intervening variable. *Inform. Syst. Res.* **1995**, 6 (3), 191–219.
20. Gemino, A.; Reich, B.H.; Sauer, C. A temporal model of information technology project performance. *J. Manage. Inform. Syst.* **2008**, 24 (3), 9–44.
21. Schmidt, R.; Lyytinen, K.; Keil, M.; Cule, P. Identifying software project risks: An international delphi study. *J. Manage. Inform. Syst.* **2001**, 17 (4), 5–36.
22. Barki, H.; Rivard, S.; Talbot, J. An integrative contingency model of software project risk management. *J. Manage. Inform. Syst.* **2001**, 17 (4), 37–69.
23. Sauer, C. Partial abandonment as a strategy for avoiding failure. In *Information Systems Development: Human, Social and Organizational Aspects*; Proceedings of IFIP WG8.2 Working Conference, Noordwijkerhout, May 17–19, 1993; Avison, D.E., deGross, J., Kendall, J., Eds.; Elsevier (North Holland): Amsterdam; 143–167.
24. Martin, A.; Chan, M. Information systems project redefinition in New Zealand: Will we ever learn? *Aust. Comput. J.* **1996**, 28 (1), 27–40.
25. Ginzberg, M.J. Early diagnosis of MIS implementation failure: Promising results and unanswered questions. *Manage. Sci.* **1981**, 27 (4), 459–478.

26. Sauer, C.; Willcocks, L.P. Unreasonable expectations: NHS IT, Greek choruses and the games institutions play around mega-programmes. *J. Inform. Technol.* **2007**, *22* (3), 195–201.
27. Flyvbjerg, B.; Bruzelius, N.; Rothengatter, W. *Megaprojects and Risk: An Anatomy of Ambition*; Cambridge University Press: Cambridge, 2003.
28. Brockner, J.; Rubin, J.Z. *Entrapment in Escalating Conflicts: A Social Psychological Analysis*; Springer-Verlag: New York, 1985.
29. Brockner, J. The escalation of commitment to a failing course of action: Toward theoretical progress. *Acad. Manage. Rev.* **1992**, *17* (1), 39–61.
30. Ross, J.; Staw, B.M. Expo 86: An escalation prototype. *Admin. Sci. Quart.* **1986**, *31*, 274–297.
31. Ross, J.; Staw, B.M. Organizational escalation and exit: Lessons from the Shoreham nuclear power plant. *Acad. Manage. J.* **1993**, *36* (4), 701–732.
32. Staw, B.M.; Ross, J. Behavior in escalation situations: Antecedents, prototypes, and solutions. In *Research in Organizational Behavior*; Staw, B.M., Cummings, L.L., Eds.; JAI Press: Greenwich, CT, 1990; Vol. 9, 39–78.
33. Keil, M. Pulling the plug: Software project management and the problem of escalation. *MIS Quart.* **1995**, December, 420–447.
34. Keil, M.; Mann, J.; Rai, A. Why software projects escalate: An empirical analysis and test of four theoretical models. *MIS Quart.* **2000**, *24* (4), 631–664.
35. Drummond, H. *Escalation in Decision Making: The Tragedy of Taurus*; Oxford University Press: Oxford, 1996.
36. Ciborra, C. and Associates, *From Control to Drift: The Dynamics of Corporate Information Infrastructures*; Oxford University Press: Oxford, 2000.
37. Holmstrom, J.; Stalder, F. Drifting technologies and multipurpose networks: The case of the Swedish cashcard. *Inform. Organ.* **2001**, *11*, 187–206.
38. Elbanna, A.R. The inertia of ERP projects: Diffusion or drift. In *Organizational Dynamics of Technology Based Innovation: Diversifying the Research Agenda*, Proceedings of the IFIP TC 8 WG 8.6 International Working Conference, June 14–16, 2007; McMaster, T., Wastell, D., Ferneley, E., DeGross, J., Eds.; Manchester, U.K., 2007; 253–266.
39. Larsen, K.R.T. A taxonomy of antecedents of information systems success: Variable analysis studies. *J. Manage. Inform. Syst.* **2003**, *20* (2), 169–246.
40. Markus, M.L.; Robey, D. Information technology and organizational change: Causal structure in theory and research. *Manage. Sci.* **1988**, *34* (5), 583–598.
41. Hirschheim, R.; Newman, M. Symbolism and information systems development: Myth, metaphor and magic. *Inform. Syst. Res.* **1991**, *2* (1), 29–62.
42. Romm, T.; Pliskin, N.; Weber, Y.; Lee, A. Identifying organizational culture clash in MIS implementation. *Inform. Manage.* **1990**, *21* (2), 99–109.
43. Pliskin, N.; Romm, T.; Lee, A.; Weber, Y. Presumed versus actual organizational culture: Managerial implications for implementation of information systems. *Comput. J.* **1993**, *36* (2), 143–152.
44. Pfeffer, J. *Power in Organization*; Ballinger: Cambridge, MA, 1981.
45. Markus, M.L.; Pfeffer, J. Power and the design and implementation of accounting and control systems. *Account. Organ. Soc.* **1983**, *8* (2/3), 205–218.
46. Robey, D. Power and politics. In *Designing Organizations*; Robey, D., Irwin: Homewood: Illinois, 1991; 347–382.
47. Markus, M.L. Power, politics and MIS implementation. *Commun. ACM* **1983**, *26* (6), June, 430–444.
48. Land, F.; Hirschheim, R. Participative systems design: Rationale, tools and techniques. *J. Appl. Syst. Anal.* **1983**, *10*.
49. Davis, G.B.; Lee, A.S.; Nickles, K.R.; Chatterjee, S.; Hartung, R.; Wu, Y. Diagnosis of an information system failure. *Inform. Manage.* **1992**, *23* (5), 293–318.
50. Cherns, A. The principles of sociotechnical design. *Hum. Relat.* **1976**, *29* (8), 783–779.
51. Newman, M.; Robey, D. A social process model of user-analyst relationships. *MIS Quart.* **1992**, *16* (2), June, 249–266.
52. Newman, M.; Noble, F. User involvement as an interaction process: A case study. *Inform. Syst. Res.* **1990**, *1*, 1.
53. Pan, S.L.; Pan, G.S.C.; Newman, M.; Flynn, D. Escalation and de-escalation of commitment to information systems projects: Insights from a project evaluation model. *Eur. J. Operat. Res.* **2006**, *173* (3), 1139–1160.
54. Bloomfield, B.; Coombs, R.; Knights, D.; Littler, D., Eds. *Information Technology and Organizations: Strategies, Networks and Integration*; Oxford University Press: Oxford, 1997.

55. Mitev, N.N. More than a failure? The computerised reservation systems at French Railways. *Inform. Technol. People* **1996**, 9 (4), 8–19.
56. Howcroft, D.; Mitev, N.; Wilson, M. What we may learn from the social shaping of technology approach. In *Social Theory and Philosophy of IS*; Mingers, J., Willcocks, L., Eds.; John Wiley and Sons: Chichester, 2004; 329–371.
57. Vidgen, R.T.; Mc Master, T. Black boxes, non-human stakeholders, and the translation of IT through mediation. In *Information Technology and Changes in Organizational Work*; Proceedings of the IFIP WG 8.2 working conference on information technology and changes in organizational work. Orlikowski, W.J., Walsham, G., Jones, M.R., DeGross, J.I., Eds.; Chapman and Hall: London, 1995; 250–271.
58. Sarker, S.; Sarker, S.; Sidorova, A. Understanding business process change failure: An actor-network perspective. *J. Manage. Inform. Syst.* **2006**, 23 (1), 51–86.
59. Bartis, E.; Mitev, N.N. A multiple narrative approach to information systems failure: A successful system that failed. In *Relevant rigour, rigorous relevance*; Proceedings of 15th European Conference on Information Systems, June 7–9, 2007; St. Gallen University, Switzerland.
60. Sauer, C. Deciding the future for IS failures: Not the decision you might think. In *Re-Thinking Management Information Systems*; Galliers, R.D., Currie, W., Eds.; Oxford University Press: Oxford, 1999.