

Managing Intranet Web Pages: Issues and Techniques

Author's Note: This article was written in 1996 as a project paper for a course in database management and described Intranets the future impact of Intranets on how company's will manage data. With the current interest in Business to Business (B2B), this article references some useful articles and also whether some of the predictions made in 1996 are proven true some years later. (Article reformatted and comments added in 2001.)

Written by Harold Henke

© Copyright Harold Henke, 1996, 2001, all rights reserved

Introduction	3
Organization	4
Literature Review	4
Issues	4
Techniques	14
Four Questions from the Literature Review.....	18
How Fast is the Intranet Growing?	19
How Do Corporations Manage Intranet Web Page Content?	21
What is the Role of the Database in Managing Intranet Web Pages?	23
What is the Role of the Database Administrator/Webmaster?	25
Conclusion: Quote and Prediction.....	27
Glossary	28
References.....	30
References from World Wide Web	31
Notes	32

Introduction

This paper provides research on a phenomena of the Internet: how to manage the proliferation of “Web Pages” (also referred to as “HTML Pages”) that are being added, daily, to the Internet. And it should be noted, Web Pages (HTML Pages) are really containers of information, so the pertinent question is: How do you manage information on the Internet or Intranet? Are databases, which are the backbone of corporate information management, the solution? Especially for Intranets? Or do you just need a Browser to keep track of your Web Pages?

The reason for this proliferation is simple: Internet access is widely available and tools to create “Web Pages” are also widely available. Statistics vary about how many “home pages” are on the Internet but one quote sums up the situation: “There are already five million home pages [*as of March, 1995*] on the Web, and we anticipate that this number will rise to six or seven million home pages by September [*September, 1995*]. ‘Web fever’ has definitely hit”. (Seybold, 1995)

The proliferation of Web Pages is not limited to personal “home pages” on the Internet but also includes the Intranet, which are corporate networks. This paper focuses on the corporate Intranet because managing Web Pages, or Intranet content, is more important to corporations, than individuals who may maintain an individual home page or even a Web site. Individuals can inherently spend more time making sure their data is current, links are correct, and they can inventory their Web sites more often than a corporate Webmaster.

For instance, at Lockheed-Martin, “...the number of static HTML pages on the company’s 140,000-user Intranet has more than tripled to well over 120,000 [*HTML pages*] in the last year. Dynamic [*HTML*] pages are growing even faster.” (Moad, 1996). Additionally, experts predict that corporate networks, Intranets, are growing fast. One article stated: “Worldwide corporate spending on Intranet [*and Internet*] technologies and services more than tripled between 1994 and 1995 to reach \$12 billion and will rise to \$208 billion by [the year] 2000”. Anonymous, 1996c.

This paper investigates the growth of Web Pages with emphasis on Intranets and attempts to answer these fundamental questions:

- How fast is the Intranet growing?
- How do corporations plan to manage Intranet Web Page content?
- What is the role of the database in managing Intranet Web Pages?
- What is the role of the Database Administrator/Webmaster?

Organization

This paper is divided into these sections;

1. Introduction-presents overview of paper that describes why the paper was written. (Select Introduction)
2. Organization-lists key topics included in this paper. (Select Organization)
3. Literature Review-introduces abstract and synopsis of articles and is divided into two parts: issues and techniques. (Select Literature Review)
4. Issues-presents review of articles that describe issues such as is there a need to manage Web Pages, and if so, how should Web Pages be managed. This paper will focus more on the issues than techniques. The reason is that techniques are driven by software vendors and technology, which change rapidly, while issues, until answered, may last a long time. (Select Issues)
5. Techniques-provides review of articles that describe tools, such as database management and indexing applications, and resources needed to manage Web Pages, such as Webmasters or Database Administrators. (Select Techniques)
6. Summary-describes themes gleaned from the literature search. (Select Summary)
7. Conclusion-supplies a quote and a prediction. (Select Conclusion)
8. Glossary-provides definitions of selected terms used in this paper. (Select Glossary)
9. References-lists articles included in the literature review and articles cited in this paper. (Select References and References from WWW)
10. Notes-describes how research was conducted for this paper and the tools used to create this paper. (Select Notes)

Literature Review

Note: Editorial comments in this document are denoted by square brackets [*editorial*]. Editorial comments are used to clarify quotations from articles.

Issues

Below is a review of articles that describe issues such as is there a need to manage Web Pages, and if so, how should Web Pages be managed. The following review focuses more on issues than techniques. The reason is that techniques are driven

by software vendors and technology, which change rapidly, while issues, until answered, may last a long time.

A current example of the difference between techniques and issues is the debate over the role of Network Computers within corporations (and the home). The technology is available now as vendors are shipping Network Computers but the issue of how Network Computers fit in the corporate infrastructure is “hotly debated”. Eventually, standards will be created that define the features and functions of a Network Computer. Standards often reflect the design and purpose of a product. But before the standards are written, the features and functions of the Network Computer will change many times. Thus techniques are often quickly resolved through market forces and standards, but issues often go unresolved or linger long after standards have been created.

Anonymous, (1996, October), Webmasters: A Snapshot, PC Week, Number 42, Volume 13, Page E-8, CD-ROM, Ziff-Davis, 1996

Analysis: Article presents results from a survey conducted by Collaborative Marketing on the role of Webmasters within corporations. Webmasters stated that top five tasks are: 1) HTML development; 2) Managing E-Mail; 3) Creating Web Pages; 4) Server installations; and 5) Analyzing statistics. Webmasters also stated that the top skills they need, on a scale of 1 (little skill) to 6 (advanced skill) were: HTML (5.2); Tables (4.7); Browsers (4.7); Graphic placement (4.5); CGI Programming (3.3); UNIX (3.1); and Java (2.5).

Article states that top concerns of the Webmasters are administrative in nature. The four top concerns were: 1) Administrative issues; 2) Traffic management; 3) Content; and 4) E-mail. [*Complete survey available at www.collmktg.com.*]

Synopsis: Article does not mention how many Webmasters were surveyed nor for what types of corporations they work for, such as Fortune 500, but as purely anecdotal, the article points out a dichotomy of the Internet and Intranet. Webmasters state that their top tasks are basically creating content for Web Pages. And they list their top skills as using HTML and Browsers, which are skills related to creating content for Web Pages. Yet their top concerns are administrative issues, traffic management, and content. Therefore you might expect that Webmasters would list project management as a top skill. Or knowledge of business processes. Or a top task might have been managing Web Pages.

Again, as purely anecdotal in nature, this article seems to suggest that Webmasters are very focused on creating Web Pages, the tools needed to create Web Pages, and the hardware and software needed to support the Web Pages but management of the content and the cost of managing the content are not issues, yet.

Anonymous, (1996, May), Untangling the Web, CommunicationsWeek International, May 6, 1996, Pages 19-21, CD-ROM, CMP Media, Inc., 1996

Synopsis: This article makes a point that the World Wide Web is, at least for business, more hype than reality and that most businesses are not making money using the Web. Article states that “If organizations aren’t making money on the

Web, it's not through lack of expenditures. Worldwide corporate spending on Internet technologies and services more than tripled between 1994 and 1995 to reach \$12 billion and will rise to \$208 billion by 2000, according to INPUT, Inc., in Mountain View, California."

Article also states that Internet technologies are not widely understood and are implemented in a hodgepodge fashion in most corporations. "Web servers come in through the window, not as official policy. People waste money on equipment and preparing data in a different way. But it is a fundamental strategic decision and requires deep reflection on how organizations structure management data."

The article also quotes the following Internet statistics: "30-40 million Internet users...130 percent annual growth, 27,000 Web Servers which double every 53 days; and 10 percent of all North Americans over 16 [*age of*], have Internet access (25 million people)." [*No citation within article for these numbers.*]

Analysis: The article points out an important fact: though the Internet is growing by leaps and bounds, few corporations have figured out a business model that produces profits. Instead, corporations must pay more attention to the real important growth, which is the Intranet. And corporations are also becoming aware that the while Internet technology provides easy access to data, the data must be managed in a logical manner. "As faith in the public Web dwindles, Intranets are being seen as the means to harness the glut of data distributed across internal company networks."

The article urges corporations to analyze their data and how and to whom they want to present the data. Just having a Web site may or may not be beneficial. "And while the potential of new tools is vast, organizations will again have to reevaluate their business processes before they can justify investment in such technologies,. The Web imperative is infectious, but it must be tempered with a sober judgment of its capabilities...". In other words, Internet technology is simply another capital expense that must be justified with a return on investment and not simply something a company must have because their competitors are on the Web. Furthermore, to make data available, you have to know what data you have, what data is important, and what data should be shared. For most corporations, understanding data often leads to a better understanding of their business processes.

Anonymous, (1996, November), How to Reap Rewards of Intranet Technology, Computer Reseller News, November 11, 1996, Page 165-169, CD-ROM, CMP Media, Inc., 1996

Synopsis: This article is a summary of IBM customer study on business Intranet uses and requirements. A key point is that 16 percent of the Fortune 1000 companies have Intranets and another 50 percent are either in the consideration or planning stage, according to Forrester Research.

Analysis: A key fact pointed out in this article is that already 16 percent of Fortune 1000 companies have Intranets. This number is sure to grow given how many companies are interested and the number of vendors selling "Intranet suites'. Whether these numbers are accurate is not as important as there is much

marketing effort being made towards selling Intranet solutions, which tends to support the growth as expressed in this article.

[If you are interested in IBM's Intranet strategy, especially how that strategy incorporates Lotus's Domino product, this is an excellent article. Also provided in the article are many real-life examples of corporations using IBM Intranet technology.]

Cash, J.I., Jr. (1996, October), Mastering the Web Isn't Easy—There Are Plenty of Sources Within and Outside, Information Week, Page 114. CD-ROM, CMP Media, Inc., 1996

Analysis: Article raises the issue of inaccurate data on Web sites and makes this key point: "Some companies are apparently revisiting the early days of the PC revolution. Back then, we found people re-keying mainframe data into their PCs. Not surprisingly, we found that re-keying was error-prone. Now, we find people creating HTML pages from scratch, instead of feeding them a well-managed and maintained central sources [*called a database*]. So we shouldn't be surprised to find errors. Companies need to invest in robust database management systems, and they need to adhere to principles like 'record a single, correct occurrence of each data item [*once*].'"

Synopsis: This article encapsulates a crucial issue: how valid is the content? And also hints at issues like: where did the data come from and who owns the data. These are important issues as corporations begin to open (expose) their Intranets to customers, what are the consequences of inaccurate data. (In a related article, a "true horror story" that describes inaccurate data is retold. Select Documents on the Net.)

Comaford, Christine, (1996, September), A Lesson in Scaling the Web Wall, PC Week, Number 36, Volume 13, September 9, 1996, Page 54, CD-ROM, Ziff-Davis, 1996

Synopsis: Article states that: "Information managers tout Intranets as ideal repositories for enterprise documents and data, but the sites can quickly evolve into a complex system as content, prepared with different software, increases and limits flexibility." Article further states that another offshoot of Intranet development is that there are two dilemmas: 1) Extended development team that includes many skill sets, such as writers and programmers; 2) Extended team members use frequently incompatible development tools.

An example of past chaos now turned to order is Cisco Systems. The Cisco Web site contains 18 gigabytes of static data including PERL scripts and documents. The 18 gigabytes was created by several hundred employees with little structure, no standards, and a hodgepodge of tools. Cisco estimates that 20 percent of the content is redundant. Cisco also estimates that they need 50 to 100 Webmasters to manage the existing data. Therefore, Cisco, is looking at a tool called Compendium that manages Web sites. Cisco estimates that they will now need only four Webmasters to manage their data.

Analysis: This article points out that managing an Intranet can be very labor intensive and can spin out of control fast. And as Cisco has found out, it can

become expensive in terms of both redundant data and human resources, such as Webmasters. Cisco, like other companies, is turning to solutions that can help the company manage the Intranet and save money.

What is missing from this article is that tools cannot reshape business processes, such as why redundant data was on the Intranet nor can tools dictate what belongs on the Intranet. Secondly, the claims that one tool can reduce workload of 50 to 100 Webmasters down to 4 Webmasters seems exaggerated. Either this is greatest tool since the invention of the database or the Webmasters were not very busy. Furthermore, this article points out that the Intranet can fast increase MIS budgets if no thought is given to the total cost of managing the Intranet.

Dorshkind, Brent, (1996, August), Web-Site Costs Considered; Creative Networks Study Shows Real Expenses Involved in First Year on the Web, LAN Times, Number 17, Volume 13, August 5, 1996, Pages 37-38.

Synopsis: Article describes a report created by Creative Networks Inc. that states that establishing and administering a Web site costs more than most IS professionals suspect and that the average Web site costs \$109,000 to create and maintain in its first year. Creative Networks states that: "They do not believe companies are paying too much attention and [*the companies*] argue that technology is cost-effective, but most would-be Webmasters do not do a cost analysis. Ongoing costs do not decline because Web site must be constantly updated and managed."

Analysis: Whether an average Web site costs \$109,000 to create and maintain in its first year (which does not seem unrealistic given hardware (server), software (server and applications), and human resources (Webmaster salary), the important points made are: 1) Web sites and Intranets are perceived as being cost-efficient just by the mere presence of one; 2) Webmasters do not do cost-analysis.

This makes sense because many Web sites or Intranets are developed at department levels by "would-be" Webmasters. Neither of which are inclined to perform a thorough cost/benefit analysis. But when the costs are reviewed at an enterprise level, many Web sites might not survive a cost/benefit analysis.

Hunington-Lee, Jill, (1996, May), Intranets: Another Management Headache, Internetwork, Number 5, Volume 7, page 10, CD-ROM, Cardinal Business Media, Inc., 1996

Synopsis: Article is a follow up to a survey conducted by Internetwork on Web usage. Article states that Intranets are phoneme of grass-roots organization and require little Management Information System (MIS) support. But this grass roots effort is causing headaches for MIS managers because issues such as security and system constraints are bubbling up. This problem is summed up by: "At least one respondent complained of the 'lack of management responsibility.' Everybody needs the Web, but no one wants to own its problems."

The article sums up the problem by stating: "But Intranet management is multifaceted; it involves much more than just debugging HTML scripts. It sure would be nice if there were a comprehensive, affordable, off-the-shelf solution for

Intranet management. But at present, our leading network management vendors...are all over the map when it comes to offering solutions". Article then provides thumbnail sketches of the Intranet management software offered by Sun, IBM, and Hewlett-Packard.

Analysis: The article sums up the problem of managing information on an Intranet but also presents another problem. As stated, the technology is fairly simple and inexpensive to obtain and use, thus allowing "grass roots", at the departmental level, support for the technology. But once this technology begins to be proliferated and more servers are added at the grass roots, departmental level, who looks at the cost and resources at an enterprise level? After all, everyone wants an Intranet but who wants to bear the costs? Companies will find that their total Information Systems (IS) budgets are skyrocketing with hardware and software costs.

But the article raises another issue when the author asks for a "comprehensive, affordable, off-the-shelf solution for Intranet management." The issue is that hardware and software do not solve problems by themselves, what solves problems is changes to business processes. For instance, even with a comprehensive Intranet management solution, how do you tie in all of the Intranet users at the grass roots, departmental level, to leverage their information? How do you decide what information belongs on the Intranet? How do you decide who can create content and who should own the content? Neither hardware, software, nor vendors, can answer these questions. These questions must be answered by the users and administrators.

Korzeniowski, Paul, (1996, August), Mastering the Webmasters, InfoWorld, Number 32, Volume 18, August 5, 1996, Pages 53.

Synopsis: Though a very short article, the article makes a very salient point. "Web technology changes so quickly that training programs cannot keep pace, so Webmasters are self-sufficient and self-taught. Use of the Internet is being driven by individual departments, much the way interest in Personal Computers grew in the 1980's."

Analysis: As stated, the salient point made is that Internet (Web sites and Intranets) are being implemented at a "grass-root" or departmental level, much the same way Personal Computers were implemented (or invaded) corporations in the early 80's. And to this day, how successful the implementation of Personal Computers is still debated, as evidenced by the emergence of the Network Computer, which suggest that Personal Computers may not be the most efficient answer for all corporations. And the same debate will take place over Intranets.

Maddox, Kate, (1996, April), Identity Crisis, Masters of the Web, Webmasters are Building the Future of Computing, So Why Don't They Get More Support from IS Managers?, Information Week, April 29, 1996, Page 46. CD-ROM, CMP Publications Inc. 1996

Synopsis: Webmaster are a growing job category in corporations. InformationWeek conducted a study that listed these responses from 109 respondents: 1) 50 percent said their companies have at least one full-time

Webmaster; 2) And of those who had a full-time Webmaster, the position was indeed called Webmaster; 3) Almost 80 percent said the position of Webmaster has existed for less than one year.

The article also discusses problems that impact Webmasters. The following quotes from the Webmasters summarize their major concerns: “Its a little confusing, trying to figure out who has authority.”, “Once I say I want an Intranet, then I have brought two major things for IS into play: more users-creating potential support problems-and network performance and security. All of a sudden you have a major impact on IS”. And a very precise comment that states: “The beauty of the Web is you can do it very quickly without involving the IS group...But the drawback is, you end up building something outside the corporate structure, and you don’t take advantage of your internal resource.”

Analysis: This article, which contains comments from Webmasters, points out the entire dilemma of the growth of Web Pages and Intranets: this growth is often unplanned, happens at departmental levels, and is not leveraging corporate IS infrastructure. If the Webmasters, who’s jobs are to create content, administer Web sites, can recognize that problems exist with managing the process, then corporations will soon recognize this problem at the enterprise level. And more importantly, corporations will want to put a price tag on this growth and will want to determine if this growth is producing real value versus perceived value. What will also be interesting is the future role of the Webmasters, who for now, are charged with developing and managing Web sites but perhaps in the future they will be policy makers too.

Maddox. Kate (1995, July), Up Close, Six Webmasters, A Diverse Mix of Professionals Takes One of the Industry’s Newest Corporate Positions, Interactive Age, July 3, 1995, Page 17

Synopsis: Article presents a day in the life, more or less, of six Webmasters. Here are some pertinent quotes from the Webmasters: One Webmaster from the Smithsonian Instituted said: “ The challenge is not only to keep it [*content*], but to make it seem like one person did it...30 to 50 people [*are*] involved in creating the content.” The Smithsonian Webmaster created templates and publishing standards that simplified content creation so that “...people doing their job could just put information on the Web...Everybody has to take ownership of their own content. If it’s just one person, that person will burn out.”

Another example by a Webmaster at Quantum said that “I don’t think (management) totally understood how big this would be.” According to the article more than half of Quantum’s 9,000 employees wanted access to publish information, such as ad reprints, financial reports, and technical support documentation on the Web site.

Analysis: Like other articles in this literature review, Webmasters, perhaps more so than their managers, recognize the costs (time) of managing Web Pages. In this case, the Webmasters are reacting more to work overload than to costs or strategy concerns. And as a theme, the Webmasters are reacting by implementing controls, albeit, templates, or publishing standards. In any case, what is

happening is that Webmasters realize that unabated growth means long hours spent managing the content and soon their managers will also realize that long hours spent by Webmasters translates into higher costs. Therefore, attention will turn to processes that dictate who creates content, who owns the content, who can access the content, and what are the costs associated with the content.

Moad, Jeff, (1996, October), Overheated Web, PC Week, Number 37, Volume 13, Page 49-51, CD-ROM, Ziff-Davis, 1996

Synopsis: Article describes how Lockheed-Martin has tackled growth of their Intranet. According to Lockheed-Martin the number of static HTML (Web) pages has tripled to over 120,000 pages from an Intranet that serves approximately 140,000 users. The company is moving aggressively to manage this growth: "In the last 18 months, (the Intranet) has gone from something that was totally ad hoc to something that was a little bit more formalized and controlled to something that's now becoming much more formalized...It has to move that way because the Intranet's viewed as a way to increase productivity dramatically...We're not just letting people run all over the place anymore."

To manage this growth, Lockheed-Martin created a group called WebServe who has the mission to create policies, manage license, and provide help. A key element of this effort is to dictate security levels for Web Pages, assign owners for content, and provide standards for creating content.

Analysis: This article encapsulates the entire issue of managing Web Pages. Lockheed-Martin has recognized that Web Pages need to be managed, which includes determining who can access the data, who owns the data, and how the data is presented. Lockheed-Martin has begun looking for tools to help manage this process. Lockheed-Martin's goal is to use tools that: "can help Web-based applications that...achieve the same levels of predictable performance, reliability and availability as traditional client/server applications."

One aspect missing from this article is cost. Although the steps being taken by Lockheed-Martin should lead to increased productivity, no mention is made of cost nor anticipated cost savings. For instance, what costs are associated with the near tripling of Web Pages? Costs should include hardware, software, and human resources. Furthermore, what are the costs associated with creating the WebServe department? These are the kinds of "indirect" costs that are not fully understood nor associated with the cost of generating Web Pages. I suspect that after Lockheed-Martin has fully explored their business process, obtained the tools they need to manage the process, then they will take a hard look at the total cost of managing an Intranet and associated Web Pages.

Nash, Kim, S, (1996, September), Weaving a Web Site, Computerworld, Volume 29, Number 36, September 4, 1995, Page 53.

Synopsis: Gist of this article is that it is easy to create Web Pages and also getting easier to incorporate all kinds of information, such as sound, video, but all of these "tricks" have "heavy technological implications." These implications include hardware constraints. Article states that experienced Webmasters believe it is "normal to completely redesign a site every six months".

Analysis: I think there is an anticipation with Web sites that every time you go to the site, you will see an icon that says NEW. But the idea of turning over all of your content every six months requires much creation and management of content. It would be interesting to read a study on how often do users expect to find new or updated information on a Web site. And also it would be interesting to read a study on whether companies believe they need to re-design their Web sites often to stay competitive. The key point to this article is that Web Pages have a short shelf life and corporations need to contend with identifying information that is dated.

Paul, Laura, G, (1996, June), Getting the Best of Both Worlds, PC Week, Number 24, Volume 13, June 17, 1996, Page 63, CD ROM, Ziff-Davis, 1996

Analysis: Article details how Eli Lilly linked a document management system to their corporate Intranet. Eli Lilly chose Documentum as their document management system and built links between the Netscape Navigator graphical Browser to their Intranet called ELVIS (Eli Lilly Virtual Information System). Because Eli Lilly must maintain accurate, precise records for the Food and Drug Agency, it was necessary to use a system that could track both approved and preliminary drafts of documents. This requires the ability to check in and check out different versions of the same document.

Eli Lilly also incorporated process controls on how to use the corporate Intranet. "Users...are not automatically given access to all the information housed on the [Intranet]. For security reasons, they must first apply to a corporate data steward to gain access to certain information, demonstrating a strong business reason for needing the information...One reason for tight controls is to avoid eating up bandwidth when accessing information that is not business-critical". And while controls and processes have been implemented, the growth of the Intranet has not decreased: "In the past year, the interest in document management has snowballed because people are seeing that information is a corporate asset. "

Synopsis: Despite the cute ELVIS Intranet name, this article points out some very interesting steps Eli Lilly has taken to manage their information. Because not only the actual data, but the history of the data is important, Eli Lilly chose to use an application that maintains a history of a document. When the document is checked out, how many versions were created, who owns the document and so forth. And while they chose an application that manages much of the tracking for them, they also implemented process controls to limit access to data based on a "need to know". One interesting fact was that not only did Eli Lilly want to control access to information, they also wanted to control the amount of resources used to obtain and view information.

Not all companies need to maintain stringent, detailed, records about their information but this article points out that it is feasible to implement controls to manage information without turning the Intranet into a bureaucratic nightmare that discourages the flow of information within a corporation.

Power, Kevin, (1996, September), Webmasters Take Regs Home, Government Computer News, Number 24, Volume 15, September 23, 1996, Page 67.

Synopsis: The General Service Administration has created a watchdog group within the federal government that will assist government agencies to make sure that their Web content complies with the Privacy Act, Computer Security Act, and Freedom of Information Act. According to the article the watchdog group, called Federal Webmasters Information Management Working Group, held a meeting in August, 1996 which was attended by 30 Webmasters from 12 government agencies. The article states that the GSA set up this group to “put some control on federal Web sites.”

Analysis: Though this article is brief and describes how the GSA set up a watchdog group to provide advice to federal agencies on how to comply with specific laws, the implications for corporations and Intranets is interesting. Implied the GSA’s decision to create a watchdog group is that there is a good possibility that federal laws can be violated by allowing Web Pages to be created without any controls other than the judgment of the Webmasters or content owners.

Though the application of the above federal laws does not apply to corporations, the implied message is the same. If content is not controlled, corporations may find themselves embroiled in lawsuits. For instance, suppose an employee reads on an Internet Newsgroup a rather scathing review of a competitor product by the competitor’s customer and decides to re-post the customers’ review on the company’s Website that is available to both customers and employees. More likely than not, the person who posted the original note on the Newsgroup is probably protected by the First Amendment but a corporation could well be subject to a libel suit for publishing the note on their Website. This is another example of the need for a complete understanding of what content belongs on the Web and who owns the content.

Schein, Ester, (1996, July), ‘Twas a Fine Year for Webmasters, PC Week, Number 26, Volume 13, Pages 39-41. CD-ROM, Ziff-Davis, 1996

Synopsis: Webmaster is a fast growing job but a job that will likely lead to burnout. Article states that Webmasters must possess many skills to manage a “myriad of tasks...[that] can lead to a quick path to burnout. Moreover, the complications of content mandates coming in from all sectors of a company require the Webmaster be part juggler, part mediator, in order to balance the different requirements of the site.” Many of the Webmasters fell into these positions because they had the skills to set up a Web site, knew UNIX, and so forth and some are often considered to be “stopgaps” because they were available. One Webmaster referred to themselves as a “de factor managing editor”.

Analysis: This article describes the condition that exists that as a job with new skills is created, many of the jobs are filled with people who have some of the skills, such as how to set up a Web site. The article also presents a list of skills required to be a Webmaster. These skills include: UNIX operating system, E-mail applications, Internet applications (like FTP), and similar, technical skills. Skills not listed included project management and production control. These types of skills may become required as part of the Webmaster’s skill set and the proliferation of Web Pages continues.

Van Name, Mark, L. and Catchings, Bill, (1996, September), With the Web, What Isn't a Document?, PC Week, Number 36, Volume 13, September 9, 1996, Page N10, CD-ROM, Ziff-Davis, 1996

Synopsis: Article states that there is a document evolution underway where traditional text based documents are becoming interactive documents. Furthermore, HTML is becoming the Latin of document formats and is as important as ASCII as a universal format. Web documents are evolving fast because Web surfers, who are customers, demand updated information and want more interactive content than traditional documents. Article makes the following points on how this will affect business processes:

- 1) Documents will need data from databases, which will cause integration problems;
- 2) Integrating database technology with interactive documents will require skills not possessed by Database Administrators, such as audio and video editing.
- 3) Revision control will be a major development issue
- 4) Software tools that support revision control will need to manage both text, audio, video and many formats of information.

Analysis: Article ends with this statement: "All of these changes—and more—will come. Prepare for them now, or be prepared to be overwhelmed by them". Article does a very good job of stating two points: 1) data must be integrated with databases and 2) revision control must be able to handle all kinds of data, text (HTML files), audio, video, graphics and so forth. Without these two components, the growth of "interactive documents" will indeed overwhelm the caretakers of the data.

Techniques

Below is a review of articles that describe tools, such as database management and indexing applications, and resources needed to manage Web Pages, such as Webmasters or Database Administrators.

Anonymous, (1996, July), Text Retrieval Born Again, Computer Business Review, July 1, 1996, CD-ROM, APT Data Services, 1996

Synopsis: Article presents a survey of text retrieval vendors, which can be defined as companies that sell "search engines" used to create "full text retrieval" which enables software to look for combinations of characters to find phrases or single words. Article describes the differences between building databases of text files versus indexing text files. The business of developing text retrieval tools is booming because of the Internet. The Open Text database, which is a 10 billion word database that has indexed most of the World Wide Web absorbs almost 1 million hits per day. Article also points out that traditional database vendors, like Oracle are competing in this field to manage retrieval of text type information.

Analysis: Text retrieval tools, like databases, are tools to be considered to help manage the growth of Web Pages. But one major drawback to these text retrieval tools is that they only show you a snapshot. For instance, if you index a term, and provide its location on the Web, that term may be gone the next day. What seems

very promising is that database developers will incorporate more sophisticated text retrieval engines in their products, thus combining both the management features of databases with the search abilities of text retrieval engines. (Select Open Text to read a related article.)

Anonymous, (1996, October), Ways to Tackle Web Programming. Information Week, October 21, 1996, Page 15, CD-ROM, CMP Media, Inc., 1996

Synopsis: Article begins with: "Two new products illustrate the differences in Web application development. One addresses management of Web content, while the other features some of the discipline and coding capabilities found in enterprise development tools." The two tools are DeltaPoint Quicksite and Haht HahtSite Software. Example is given of how Quicksite was used to replace Microsoft FrontPage and HotMetal Pro because Quicksite could take advantage of existing databases. The product is described by one user as a "Its an automated page creation and maintenance tool more than it is an HTML authoring or programming tool." HahtSite on the other hand is an application server with an object-oriented language that enables programmers to "treat Web Pages as an application, rather than a set of Web Pages with links."

Another point made in this article is that there is a lack of standards on how to develop applications for the Web, as stated by the president of a small software company: "All of a sudden, with the rush to the Web, we all stepped back into the kind of application development we were doing in the '70's, as far as programming standards are concerned. Its all cut and paste from source code files all over again".

Analysis: This article raises two questions: 1). What tools should be used to develop and manage Web Pages? 2) Are standards needed? The answer to the first question is difficult as this article shows that building Web Pages can be as simple as using an editor like HotMetal Pro, to more complex using a "front-end" to access information from a database, to more complex in using object-oriented programming languages to create "application Web Pages". Also, the article alludes to the fact that one package addresses management of content, while the other addresses development. But in reality, either way, you have to create content and manage the content.

Question number two asks whether standards are needed for creating Web Pages. For now, the only real standard is ISO 8879:1986 which defined Standard Generalized Markup Language (SGML) of which HTML is a Document Type Definition (or more simply, a subset). But as for standards for how to create Web Pages, it is unlikely enough people recognize the need for standards because of the diversity, and simplicity, of creating Web Pages.

Bort, Julie, (1996, May), Documents On the Net—Document Management Technology Gets a Needed Boost From the Internet, VAR Business, May 15, 1996, Page 100

Synopsis: Article states that document management and Internet are obvious marriage partners. "The union of the two technologies makes sense on a number of fronts. First, as Web sites become larger and filled with more critical business

data, they fall prey to document management woes. That is, they need some way to track information for changes and edits, authenticate who has access and allow multiple formats.” Article also points out that users are getting used to the idea of working with document management tools because data is changed much more often, from daily to weekly. And furthermore, the article states: “The opportunity for real-time distribution of wrong information is big.”

Publishing real-time misinformation is described by a “true horror story” where a large manufacturing company received higher than normal volumes for a particular product. To meet this order, the manufacturing company needed to order additional bolts for production, which the manufacturing company’s regular vendor could not produce. To shorten the purchasing process, the manufacturing company sought out another supplier using the Internet. On the Internet, the manufacturing company published their bolt specifications and asked for bids. A company submitted a bid that was accepted. The manufacturing company took delivery of the bolts only to realize the bolts were the wrong size because the bolt specification posted on the Internet was outdated. The manufacturing company had to pay for the bolts and could not fulfill their orders.

Analysis: The article provides a real-life, “true horror story” which identifies an issue: What are the consequences for publishing inaccurate or outdated information? The article also mentions a key point, not mentioned often, in any other articles, that is “authenticate” who has access. You do not want many people to access your specifications. And in the story presented, you could say that some person simply forgot to verify the bolt specification was outdated. But perhaps someone replaced the bolt specification because they thought their version was the most accurate. Without some system or process to manage the ownership of the data, Web Pages may be only as accurate as the last person to update the material. And with data being distributed real-time, the possibilities to catching a mistake decrease.

For instance, if we are talking about a price sheet, what happens if the price sheet is updated by several people, all of whom have different prices for Widget X. Unless these people communicate with each other (maybe they are located in different cities or countries), the last person who updates the price sheet is the one who determines what price Widget X should cost. With document management defining an owner, access can be limited to an owner or at least you could track who last updated the price sheet. These are issues that will grow and will not go away quietly.

Hibbard, Justin, (1996, December), Basis Intranet System Gains Windows NT Support, Computerworld, Volume 30, Number 49, page 10.

Synopsis: Article describes a document management package developed by Information Dimensions, which is now compatible with Windows NT. The article explains why vendors are developing document management tools for the Internet and Intranet by stating: “Universal access afforded by the Intranet is the main reason users are connecting document management systems to Web servers...Suddenly, the information stored and managed in their document management system is opened up to the entire corporation.” Article also points

out that internal documents are now being shared with customers, who are given access to segments of the Intranet.

A set of statistics from the Gartner Group [*research group that focuses on technology issues*] included in the article states: “More than 95 percent of the information in enterprises is in the form of documents, and roughly 80 percent of those documents sit on desktop hard drives. An Intranet couple with a document management system is an effective way to free those documents.”

Analysis: Document management systems have an important role in managing Web content. As Web Pages continue to grow by leaps and bounds, companies will be looking for solutions to manage the content. For companies who do not have a strong infrastructure in database management, document management systems, whether stand-alone applications or as tools within an application, such as a Web site server application, will be vital applications in the future. Companies with strong backgrounds in database usage or development will opt for using databases to manage their Web content.

Seybold (1995, March), Seybold - Open Text Introduces Latitude and Web Search Index, Newsbytes, March 31, 1995, Public Relations Story

Synopsis: Article describes Open Text and how Oracle licensed the Open Text search engine for their database. Open Text can search for information on all types of media including hard drives, tape drives, Local Area Networks, and so forth. Open Text can also work with markup language, such as SGML and HTML, and provides full Boolean search capability. Article also describes how Open Text is used by Caterpillar to manage data and search for information across stacks of CD ROMs located throughout Caterpillar’s Wide Area Network. Open Text believes that search engines like Open Text have a great future because “There are already five million home pages [*as of March, 1995*] on the Web, and we anticipate that this number will rise to six or seven million home pages by September [*September, 1995*]. ‘Web fever’ has definitely hit”.

Analysis: This article points out that database developers are already integrating text retrieval engines into their databases. This points out the possibility of the database becoming the single repository for both storing and handling Web Pages. And the article also points out that text search engines will be a vital tool in managing Web Pages because of their ability to find a document, which both users and content managers will want as part of their Internet toolkit.

Strom, David, (1995, December), Managing a Web of a Mess, InfoWorld, Volume 17, Number 50, December 11, 1995, Pages 75-78.

Synopsis: Article states that managing content on the World Wide Web can be a “difficult and time consuming job.” Article suggest a method to reduce is to install a text search engine that “indexes textual information and lets visitors search for information using a fill in the blanks form.” Article mentions FrontPage from Vermeer Technologies. FrontPage is described as “authoring tool that helps track links and pages by providing a visual image of the site..and...automates the creation of forms and scripts that normally require programming.”

Analysis: This article lists a popular approach to managing Web Pages and that is tools that provide search engines, which lessen the need to build navigation aids within the Web site and also that check for outdated URLs, and provide in essence, a map of the directory structures. Where these tools often lack are determining whether the content is redundant, who owns the content, and so forth. A text search engine is valuable in searching for a topic but as administration tool, all search engines are limited by the Boolean operators supported, ability to recognize (and pair words or phrases together) and most importantly, by the user's ability to enter meaningful search terms or phrases. For this reason, jobs exist for people who are adept at searching databases by knowing what words or phrases will produce the most number of "hits". But search engines alone are not complete management tools for Web sites or Intranets.

Four Questions from the Literature Review

As stated at the beginning of this paper, there is a proliferation of Web Pages on the Internet and Intranet. The reason for this proliferation is simple: Internet access is widely available and tools to create "Web Pages" are also widely available. Statistics vary about how many "home pages" are on the Internet but this quote sums up the situation: "There are already five million home pages [*as of March, 1995*] on the Web, and we anticipate that this number will rise to six or seven million home pages by September [*September, 1995*]. 'Web fever' has definitely hit". Seybold, 1995

And further echoing (hyping, perhaps) this growth: "[*There are*] 30-40 million Internet users...130 percent annual growth, 27,000 Web Servers which double every 53 days and 10 percent of all North Americans over 15 [*age of*], have Internet access (25 million people)." Anonymous, 1996a

The proliferation of Web Pages is not limited to personal "home pages" on the Internet but also includes the Intranet, which are corporate networks. Corporate networks, Intranets, are growing fast. One article stated: "Worldwide corporate spending on Intranet [*and Internet*] technologies and services more than tripled between 1994 and 1995 to reach \$12 billion and will rise to \$208 billion by [*the year*] 2000". Anonymous, 1996c.

Another article describes the number of Fortune 1000 companies with Intranets: "16 percent of Fortune 1000 companies have Intranets and another 50 percent are either in consideration or planning stage [*to create an Intranet*]." Anonymous, 1996e

Additionally, "corporate" Web Pages are also growing fast. For instance, at Lockheed-Martin, "...the number of static HTML pages on the company's 140,000-user Intranet has more than tripled to well over 120,000 [*HTML pages*] in the last year. Dynamic [*HTML*] pages are growing even faster." (Moad, 1996).

From the literature review, these four questions emerged:

1. How fast is the Intranet growing?
2. How do corporations manage Intranet Web Page content?
3. What is the role of the database in managing Intranet Web Pages?
4. What is the role of the Database Administrator/Webmaster?

How Fast is the Intranet Growing?

Considering the number of “white papers” available from vendors on how to apply Intranet technologies to business, the number of research studies provided by research firms, and the growing number of articles in the trade press, you can surmise that Intranet technology is a growth industry within the information industry.

But where there is money, there is smoke. Many of the “findings” are based on small samples, such as a telephone survey of 170 companies representing medium to large size businesses, and often, as is usually true with industry, research corporations are hired on behalf of vendors to conduct the surveys. But, despite the hype and smoke, you cannot discount the growth of Intranet technologies. Clearly Intranet technology is a growth industry that offers benefits to any business with a Local Area Network or Wide Area Network infrastructure. (Or for that matter, any collection of computers that support TCP/IP.) The following articles give credence to the growth of Intranet technology in corporations.

Corporate networks, Intranets, are growing fast. One article stated: “Worldwide corporate spending on Intranet technologies and services more than tripled between 1994 and 1995 to reach \$12 billion and will rise to \$208 billion by [*the year*] 2000”. Anonymous, 1996c. Another article describes the number of Fortune 1000 companies with Intranets as: “16 percent of Fortune 1000 companies have Intranets and another 50 percent are either in consideration or planning stage [*to create an Intranet*].” Anonymous, 1996e.

Another couple of studies cited by Netscape Communications Corporation, in a Netscape white paper, stated that: “In fact, two-thirds of Fortune 1000 companies have an Intranet, according to Forrester Research, July, 1996...[*and*] By the turn of the century, Intranet servers will outsell Internet servers 4.6 million to 440,000, according to International Data Corporation.” Further, this same white paper states that “IDC predicts that by 1997, 80 percent of Web servers will be used for internal sites [*Intranets*]. Netscape, 1996

Echoing these figures, PC Week cites a study: “A recent study of 65 companies by Zona Research, Inc., a Redwood City, California, market researcher, found that half [*of the 65 companies*] had Intranets...And a poll of 170 decision makers at medium sized and large companies, conducted by Business Research Group, of Newton, Massachusetts, found that 23 percent have launched or plan to launch

Intranets, while 20 percent were studying them [*considering adding Intranets to their companies*].” The article further states that the Intranet may be a bigger market than the Internet: “The potential is there for the enterprise Web [*Intranet*] to become a bigger market than the Internet, [*according to*] Gartner Group, Stamford, Connecticut. The company [*Gartner Group*] predicts that more than half of all enterprises will deploy internal Webs [*Intranets*] by 1998 but cautions that it will be the end of the century before Intranets [*will be fully implemented*].” PC Week, 1996

Along these lines, Business Week also cites Zona research by stating that: “...sales of software to run Intranet servers will shoot to more than \$4 billion in 1997, from \$476 million last year [1995]. In 1998, Zona says the figure will hit \$8 billion, four times the size of the Internet sever business.” Business Week, 1996

And one last article, published by VARBusiness, a trade magazine for value added resellers (VAR), states that according to a study commissioned by VARBusiness and conducted by Intelliquest: “29 percent of VARs provide Intranet servers using their own staff ...and 54 percent of VARs add value to Intranet solutions by providing database connectivity.” [*More information is provided on the role of the database in Intranets later in this paper.*] Jacobs, 1996

From the above literature, three key points emerge about Intranet growth:

1. **Corporate spending will increase to purchase Intranet technology.** Even if the estimated growth in sales, which are in the billions of dollars, are exaggerated, the estimates suggest tremendous increases in corporate expenditures in less than three years. This growth could easily parallel the growth of the mid-80’s when personal computers began to be integrated into corporate information system infrastructures. **Note: This is where the real and profitable money has been made, getting “brick and mortar” businesses online and able to exchange data, in real-time, with their suppliers and customers.**
2. **Intranet technology will outsell Internet technology.** This is a perfectly logical assumption given that to provide Internet access to customers, a corporation merely needs a server or a local Internet Service Provider, a Home Page, and a few tools to create and manage the content. But to provide Intranet technology within a corporation, the corporation must take advantage of existing LAN or WAN infrastructure, provides tools to those who need to create content and manage content. In database terms, the Internet represents a one to many relationship whereas the Intranet represents a many to one relationship. **Note: This is at least the B2B argument. Not sure if this will pan out but connecting companies to talk to each other, securely and privately is big business.**
3. **Intranet usage is not universal, yet.** The estimates vary on how many Fortune 1000 companies have implemented Intranet technology in their information system but the estimates range from 16 to 66 percent. Regardless of the numbers, Intranet technology is not as prevalent in corporations as is Internet technology. And these numbers do not reflect

the majority of businesses which are represented by non-Fortune 1000 companies. But with companies, small to large, using Internet technology to market their products on the Internet, the next logical step is to use similar technology (Intranet) in their businesses. **Note: This would not be true in 2001; Intranets are as common as dot.com failures.**

All in all, Intranet growth will skyrocket over the next few years because: "People woke up one morning and realized they had all the pieces in place, says Paul D. Callahan [*Zona Research*]" Business Week, 1996, which means that the existing information system infrastructure, the LANs, the WANs, the servers, and so forth can be used to build the Intranet. Corporations will realize that: "The Intranet has broken down the walls within corporations, says Steve Jobs, Chief Executive Officer of NeXT Computer Inc." Business Week, 1996. This provides a tremendous vehicle to disseminate information across the corporation, whether locally or globally. With the realization that existing infrastructure can be used to build an Intranet and the ability to disseminate information easily, corporations will begin to address the issue of how to manage the Intranet Web Page content.

How Do Corporations Manage Intranet Web Page Content?

First and foremost, do corporations have a strategy for managing Web Page content? To answer this question, you must first look at how Internet and Intranet technology is being adsorbed by corporate information systems. The Internet and Intranet are growing fast in corporations partly because of a ground swell of interest at the departmental level. This ground swell makes it hard for corporate information systems to plan and "roll out" Intranets as part of a global strategy. Instead, as was pointed out earlier, the technology is mostly in place to create an Intranet, so Intranets using existing networks can be implemented at departmental levels thus bypassing traditional information system control.

These quotations describe this phoneme: "Web servers come in through the window, not as official policy. People waste money on equipment and preparing data in a different way. But it is a fundamental strategic decision and requires deep reflection on how organizations structure management data." Anonymous, (1996b)

And "In the last 18 months, (the Intranet) has gone from something that was totally ad hoc to something that was a little bit more formalized and controlled to something that's now becoming much more formalized...It has to move that way because the Intranet's viewed as a way to increase productivity dramatically...We're not just letting people run all over the place anymore." Moad, 1996

The following quotes from Webmasters, the guardians or caretakers of Web Page content, summarize their major concerns: "Its a little confusing, trying to figure out who has authority.", "Once I say I want an Intranet, then I have brought two

major things for IS into play: more users-creating potential support problems-and network performance and security. All of a sudden you have a major impact on IS". And a very precise comment that states: "The beauty of the Web is you can do it very quickly without involving the IS group...But the drawback is, you end up building something outside the corporate structure, and you don't take advantage of your internal resource." Maddox, 1996a

From these quotations, it is fair to say that Intranets are being integrated haphazardly within corporations but corporations are realizing that the growth of Web Pages, both on Internet and Intranet sites presents logistical problems. These problems can be summarized by replication of data, unmanaged growth, and inaccurate data that can cause "horror stories". The following quotations summarize these problems:

"Some companies are apparently revisiting the early days of the PC revolution. Back then, we found people re-keying mainframe data into their PCs. Not surprisingly, we found that re-keying was error-prone. Now, we find people creating HTML pages from scratch, instead of feeding them a well-managed and maintained central sources [*called a database*]. So we shouldn't be surprised to find errors. Companies need to invest in robust database management systems, and they need to adhere to principles like 'record a single, correct occurrence of each data item [*once*].'" Cash, 1996

Another article states that: "Information managers tout Intranets as ideal repositories for enterprise documents and data, but the sites can quickly evolve into a complex system as content, prepared with different software, increases and limits flexibility." Comaford, 1996

And in a case study of how one company is tackling Web Page management: "Eli Lilly also incorporated process controls on how to use the corporate Intranet. "Users...are not automatically given access to all the information housed on the [*Intranet*]. For security reasons, they must first apply to a corporate data steward to gain access to certain information, demonstrating a strong business reason for needing the information...One reason for tight controls is to avoid eating up bandwidth when accessing information that is not business-critical". And while controls and processes have been implemented, the growth of the Intranet has not decreased: "In the past year, the interest in document management has snowballed because people are seeing that information is a corporate asset. " Paul, 1996

Other articles point out that companies are carefully reviewing their content, Cisco Systems, for instance, estimates that 20 percent of their content is duplicated Comaford, 1996 and companies like Lockheed-Martin have created specific organizations, in this case called WebServe, to create policies for creating Web Page content, Moad, 1996, and others, such as the General Service Administration, have created a watchdog group to assist government agencies in providing content that does not violate laws, such as the Privacy Act. Power, 1996.

Additionally, even corporate Web sites open to the public may not be as critical as those Intranets that contain sensitive corporate data. For instance, if a Web site is used to promote a product, a mislabeled product or part number may cause confusion and disappoint the customer, but in one instance, a company created a “true horror story” Bort, 1996 because the company provided Intranet access to vendors to obtain specifications for a part but the specifications were outdated which resulted in the company ordering the wrong part from the vendor and being unable to fulfill a large order. (Of course, if the Internet is being used by customers to order products, then outdated data is as critical as any data stored on an Intranet.)

These key points can be made:

- 1. Intranet growth is occurring at departmental levels.** This growth, which is reminiscent of how personal computers were integrated into corporate information systems in the mid-80's, is in conflict with centralized information system planning. Companies are “wising up” to this fact, and are focusing on implementing and integrating Intranets into the corporate information systems organization, such as the case of Eli Lilly and Lockheed-Martin. **Note: In many companies, there has been an effort to “centralize” their Intranets because too much decentralized growth occurred. After all, with a little cable and a hub, you too can create a network.**
- 2. Corporations are creating policies and procedures to manage Web Page content.** Corporations, like Cisco Systems, Lockheed-Martin, Eli Lilly, and government agencies like the General Service Administration, are reviewing their Web Page content, creating departments to manage the growth, creating policies and procedures for creating Web Page content, and in some cases, creating organizations to regulate Web Page content. **Note: For many companies, the days of rampant web-site creation, with all kinds of different “look and feel” has been replaced with consistent “corporate” presence.**

With corporations embracing Intranet technology and recognizing the need to manage the Intranet content, tools, such as the database, are needed to provide the management of the content, and people, called Database Administrators/Webmasters are needed to use the tools.

What is the Role of the Database in Managing Intranet Web Pages?

From the literature review there is a growing recognition of the need to manage Web Pages. The literature can be summarized as stating that the most prevalent tools are document management systems, search engines, and databases. And in some instances, combinations of all three tools.

But before describing the tools, an issue that is becoming important is the lack of standards: "All of a sudden, with the rush to the Web, we all stepped back into the kind of application development we were doing in the '70's, as far as programming standards are concerned. Its all cut and paste from source code files all over again". Anonymous, 1996b

For now, the only real standard is International Standards Organization (ISO) 8879:1986 which defines Standard Generalized Markup Language (SGML) of which HTML is a Document Type Definition (or more simply, a subset). But ISO 8879: 1986 only addresses the markup of a Web Page (HTML Page) and not how to access that document. There are no standards on how a database is supposed to respond to a query generated by a CGI script.

Since product development seldom waits for standards to be defined, there are many tools available for managing Web Pages. One method is to rely on search engines that already commonly used by Internet users to search the World Wide Web. For instance, Open Text is a search engine capable of providing full text retrieval, which means Boolean searches are supported. Open Text is used by Caterpillar to manage data and search for information across stacks of CD ROMs located throughout Caterpillar's Wide Area Network. Seybold, 1995

But search engines do not provide actual document management. Search engines can be used to located documents, but a search engine cannot determine if the document is outdated. But tools, called document management systems, can determine if a document is outdated.

"The union of the two technologies makes sense on a number of fronts. First, as Web sites become larger and filled with more critical business data, they fall prey to document management woes. That is, they need some way to track information for changes and edits, authenticate who has access and allow multiple formats." This article also points out that users are getting used to the idea of working with document management tools because data is changed much more often, from daily to weekly. And furthermore, the article states: "The opportunity for real-time distribution of wrong information is big." Bort, 1996

And vendors are developing document management tools: "Universal access afforded by the Intranet is the main reason users are connecting document management systems to Web servers...Suddenly, the information stored and managed in their document management system is opened up to the entire corporation." The article also points out that internal documents are now being shared with customers, who are given access to segments of the Intranet. A set of statistics from the Gartner Group [*research group that focuses on technology issues*] included in the article states: "More than 95 percent of the information in enterprises is in the form of documents, and roughly 80 percent of those documents sit on desktop hard drives. An Intranet coupled with a document management system is an effective way to free those documents." Hibbard, 1996

In addition to document management tools, many of which are really database applications with a specific purpose, another set of tools is the combination of databases and search engines. Oracle for instance, has licensed Open Text, and

will incorporate the Open Text search engine into the database, which means in addition to performing traditional queries, you can now search text strings. Additionally many tools, such as Quicksite, Anonymous, 1996b, are touting the ability to access information from existing databases without creating CGI scripts and so forth.

These key points can be made:

1. **Text search engines will be integrated with databases.** Database vendors, like Oracle (with Open Text) and Microsoft (which recently purchased Vermeer Technologies), will integrate text search engines into their database engines to enable users to query not only data within the database but other data on the Intranet.
2. **Document management systems will become popular again.** Document management systems which have been in fashion and out of fashion, will be revitalized as a method for managing Web Pages. Document management systems have been widely used in technical publishing, book publishing, catalog publishing and with the tremendous growth of Web Pages, these tools will find new uses outside of the traditional publishing markets. But as more tools are developed that provide links to databases, and as database add text search engines, the growth the document management market may be limited.

Regardless of which tools are used to manage Web Pages, there must be a person who puts these tools to work managing the Web Page content. And this person is the Database Administrator/Webmaster.

What is the Role of the Database Administrator/Webmaster?

Many predictions have been made that the job, Webmaster, is one of the fastest growing and better paid jobs in the information industry. But a few years ago, the same was said of the Database Administrator. The reality is that many Webmasters are former Database Administrators, and if you accept the prediction that follows this section, the Webmasters may again be Database Administrators. The literature review points out that the job and tasks associated with Webmasters are still being sorted out but Webmaster job is certainly a growing employment opportunity.

According to one article, which cited an InformationWeek study that listed these responses from 109 respondents: 1) 50 percent said their companies have at least one full-time Webmaster; 2) And of those who had a full-time Webmaster, the position was indeed called Webmaster; 3) Almost 80 percent said the position of Webmaster has existed for less than one year. Maddox, 1996a. (Of course there is a downside to this growth. Cisco Systems reports that the company employs 50 to 100 Webmasters but with increased emphasis on managing the company's Web Page content, the company hopes to reduce this number to four Webmasters. Comaford, 1996.)

And as stated, the job description and tasks are being sorted out. An article presented results from a survey conducted by Collaborative Marketing on the role of Webmasters within corporations. Webmasters stated that top five tasks are: 1) HTML development; 2) Managing E-Mail; 3) Creating Web Pages; 4) Server installations; and 5) Analyzing statistics. Webmasters also stated that the top skills they need, on a scale of 1 (little skill) to 6 (advanced skill) were: HTML (5.2); Tables (4.7); Browsers (4.7); Graphic placement (4.5); CGI Programming (3.3); UNIX (3.1); and Java (2.5). Article states that top concerns of the Webmasters are administrative in nature. The four top concerns were: 1) Administrative issues; 2) Traffic management; 3) Content; and 4) E-mail. Anonymous, 1996a [*Complete survey available at www.collmktg.com.*]

Also, as cited by WebMaster magazine, Webmasters are becoming sensitive to organizational issues that include: “Role of core group: Is there a need for central group; Data quality and management: Is this the “right” information, Is it up to date? Accurate? , Who manages Web data quality”. WebMaster

From the above survey, a dichotomy is revealed. Webmasters cite administrative issues as their top concerns, yet the Webmasters listed the top skills required as mostly technical in nature, such as HTML development. It would seem that the Webmasters might have listed project management as a top skill. With the growth of Web Pages and the role of Webmaster still being defined, job burnout is an issue. Webmasters must possess many skills to manage a “myriad of tasks...[that] can lead to a quick path to burnout. Moreover, the complications of content mandates coming in from all sectors of a company require the Webmaster be part juggler, part mediator, in order to balance the different requirements of the site.” Many of the Webmasters fell into these positions because they had the skills to set up a Web site, knew UNIX, and so forth and some are often considered to be “stopgaps” because they were available. One Webmaster referred to themselves as a “de factor managing editor”. Schein, 1996.

These key points can be made:

- 1. Corporations are hiring Webmasters to manage their Web Page content.** Advertisements in publications like ComputerWorld are filled with openings for Webmasters. And many of the existing Webmasters have only had the job or title for a short while, less than two years. **Note: They still are but many corporations are also outsourcing their web-sites including content management and creation.**
- 2. Webmaster job descriptions and tasks are still being defined.** Currently, Webmasters and corporations place emphasis on technical skills such as HTML development, but this will change when project management and database skills become more crucial to the management of the Web Page content. **Note: The jobs are now well-known but evolving. Much more emphasis is now placed on proving the value of the website and less on adding singing chipmunks and flash(y) scripts.**

As corporations pay more attention to managing Web Pages, the Webmasters will become the guardians of the Web Site, which in “olden” days were called Database Administrators.

Conclusion: Quote and Prediction

A quote, from Eli Lilly: “Intranets aren’t the Holy Grail of computing...But for now, they’re hard to beat.” Business Week, 1996. And a prediction: as the Intranet grows, the database will still be the engine of choice to manage information and make that information accessible throughout the enterprise. User’s who never generated a “query” before will now generate a “search” to seek out information from the database, much like they seek information from the World Wide Web today. **Note: Well B2B was a Holy Grail and will likely become Holy Grail; the Intranet has been found and we are looking for the next best thing.**

Glossary

The following terms are used in this paper. Please note, these are the author's definitions of these terms.

Browser

In its humblest form, an application used to view HTML documents. In its grandest form, the gateway to the information age or the automobile you drive on the Information Super Highway

CGI

Common Gateway Interface. A method for creating forms to access or provide data. Often used to retrieve data or input data from a Browser to a database.

Home Page

The starting point for accessing Web Pages. A home page can be thought of as an index to a directory or the first file in a folder.

HTML

Hypertext Markup Language. Markup language used to create documents for graphical Browser display. (Markup languages are similar to formatting tags used in word processing packages. What you see what you get (WYSIWYG) word processors "hide" the fact that underneath your text are formatting (markup) tags.)

Internet

A collection of computer networks interconnected together. (Sometimes referred to as the "Information Super Highway".)

Intranet

A network that supports Internet protocols within a corporation. Used to store and provide access to Web Pages on internal Web sites.

PERL

Practical Extraction and Report Language. A programming language used to create programs to parse or filter information.

Server

A computer that supplies services, such as access to a Web Pages, to users. (Users are the clients; hence you can say the Internet is one big client/server application.)

TCP/IP

Transfer Control Protocol /Internet Protocol. A set of standards (protocol or rules) for sending and receiving data (and correcting transmission errors) from one computer to another computer. Very important set of rules, like traffic laws, for transmitting data on the Information Super Highway.

VAR

Value Added Reseller. A company that supplies additional hardware, software, or services to another company's product offering. For example, if a company wanted to install an Intranet, the company might purchase a solution from a vendor who offered both hardware and software; a VAR might be hired by the company (or sub-contracted by the vendor) to connect the company's database to the vendor's Intranet solution.

Webmaster

An administrator who manages a Web Site. (Similar to a Database Administrator.) The Webmaster's tasks may include creating content, designing the Home Page, and managing the server.

Web Page

A container of information that contains text (and links to audio, graphics, and video), marked-up with HTML tags and displayed with a graphical Browser. Also called an HTML Page.

Web Site

A collection of Web Pages that are accessed from a Home Page stored on a server.

White Paper

A report (also called a position paper) written to express a vendor's viewpoint about an issue or to define an issue. In general, white papers are written to explain a vendor's viewpoint on a given subject. [A poorly written white paper will mention the vendor's products at every turn; a well written white paper will seldom mention the vendor's products but will lead you to the conclusion the vendor's products represent the best solution. The author of this paper has written white papers in the past.]

References

The following references were included in the literature review.

- Anonymous, (1996a, May), Untangling the Web, CommunicationsWeek International, May 6, 1996, Pages 19-21, CD-ROM, CMP Media, Inc., 1996
- Anonymous, (1996b, July), Text Retrieval Born Again, Computer Business Review, July 1, 1996, CD-ROM, APT Data Services, 1996
- Anonymous, (1996c, October), Webmasters: A Snapshot, PC Week, Number 42, Volume 13, Page E-8, CD-ROM, Ziff-Davis, 1996
- Anonymous, (1996d, October), Ways to Tackle Web Programming, Information Week, October 21, 1996, Page 15, CD-ROM, CMP Media, Inc., 1996
- Anonymous, (1996e, November), How to Reap Rewards of Intranet Technology, Computer Reseller News, November 11, 1996, Page 165-169, CD-ROM, CMP Media, Inc., 1996
- Bort, Julie, (1996, May), Documents On the Net—Document Management Technology Gets a Needed Boost From the Internet, VAR Business, May 15, 1996, Page 100
- Cash, J.I., Jr. (1996, October), Mastering the Web Isn't Easy—There Are Plenty of Sources Within and Outside, Information Week, Page 114. CD-ROM, CMP Media, Inc., 1996
- Comaford, Christine, (1996, September), A Lesson in Scaling the Web Wall, PC Week, Number 36, Volume 13, September 9, 1996, Page 54, CD-ROM, Ziff-Davis, 1996
- Dorshkind, Brent, (1996, August), Web-Site Costs Considered; Creative Networks Study Shows Real Expenses Involved in First Year on the Web, LAN Times, Number 17, Volume 13, August 5, 1996, Pages 37-38.
- Hibbard, Justin, (1996, December), Basis Intranet System Gains Windows NT Support, Computerworld, Volume 30, Number 49, page 10.
- Hunington-Lee, Jill, (1996, May), Intranets: Another Management Headache, Internetwork, Number 5, Volume 7, page 10, CD-ROM, Cardinal Business Media, Inc., 1996
- Jacobs, I.G., (1996), Database Connectivity, The Web Goes Live, Internet VARs Make the Web-to-Database Connection, December 15, 1996, pages 24-32. [Not included in literature review but cited in paper.]
- Korzeniowski, Paul, (1996, August), Mastering the Webmasters, InfoWorld, Number 32, Volume 18, August 5, 1996, Pages 53.
- Moad, Jeff, (1996, October), Overheated Web, PC Week, Number 37, Volume 13, Page 49-51, CD-ROM, Ziff-Davis, 1996

Maddox, Kate, (1996a, April), Identity Crisis, Masters of the Web, Webmasters are Building the Future of Computing, So Why Don't They Get More Support from IS Managers?, Information Week, April 29, 1996, Page

46. CD-ROM, CMP Publications Inc. 1996

Maddox. Kate (1995b, July), Up Close, Six Webmasters, A Diverse Mix of Professionals Takes One of the Industry's Newest Corporate Positions, Interactive Age, July 3, 1995, Page 17

Nash, Kim, S, (1996, September), Weaving a Web Site, Computerworld, Volume 29, Number 36, September 4, 1995, Page 53.

Paul, Laura, G, (1996, June), Getting the Best of Both Worlds, PC Week, Number 24, Volume 13, June 17, 1996, Page 63, CD ROM, Ziff-Davis, 1996

Power, Kevin, (1996, September), Webmasters Take Regs Home, Government Computer News, Number 24, Volume 15, September 23, 1996, Page 67.

Schein, Ester, (1996, July), 'Twas a Fine Year for Webmasters, PC Week, Number 26, Volume 13, Pages 39-41. CD-ROM, Ziff-Davis, 1996

Seybold (1995, March), Seybold - Open Text Introduces Latitude and Web Search Index, Newsbytes, March 31, 1995, Public Relations Story

Strom, David, (1995, December), Managing a Web of a Mess, InfoWorld, Volume 17, Number 50, December 11, 1995, Pages 75-78.

Van Name, Mark, L. and Catchings, Bill, (1996, September), With the Web, What Isn't a Document?, PC Week, Number 36, Volume 13, September 9, 1996, Page N10, CD-ROM, Ziff-Davis, 1996

References from World Wide Web

The following references were not included in the literature review but were cited in this document. Each of these documents is available on the World Wide Web; you can use your Web Browser to go directly to the article by selecting the link next to the article. If you have difficulty finding an article (the links may change), try using a search engine, and search on the article title.

Anonymous (1996a), Intranets Redefine Corporate Information Systems, Netscape Communications Corporation, *Netscape*

Anonymous (1996b), Some Organizational Issues, WebMaster Magazine, 1996, *WebMaster*

Roberts, B., (1996), Internal Affairs, PC Week *PC Week*

Cortese, A., (1996), Here Comes the Internet, February 26, 1996, *Business Week*

Notes

The literature search was conducted using electronic library databases available from Nova-Southeastern's Electronic Library and other commercially available electronic library databases. These databases include both academic and business publications. Searches were conducted on-line using search tools provided by the databases. Search words included phrases, singly or in combinations, such as Intranet, Internet, Web sites, Web Pages, HTML, Webmasters, cost, issue, managing, and so forth.

All told approximately 575 abstracts were reviewed. Of these abstracts, approximately 55 articles were read in their entirety. From these 55 articles, 22 articles were included in this literature review. The articles included in the literature review were chosen based on succinct, precise comments, and also representation of a diverse sample of publications. (But for a topic such as Intranet, the number of publications with articles is generally limited to business periodical publications and much less to refereed publications.)

Further, because of the amount of information available, articles were limited to 1995 and 1996 with the majority of articles published in 1996. In traditional research papers, information published one or two years ago is quite current, but in Internet terms, six months to one year is current. (Did Intranet exist as a term three years ago?)

In addition to using on-line electronic library databases from Nova-Southeastern's Electronic Library, data was gathered using World Wide Web search engines, such as Yahoo, Lycos, and so forth. On-line sources gathered from the World Wide Web are listed under *References From the World Wide Web*.

Total word count is almost 13,000 words.

© Copyright Harold Henke, 1996, 2001, all rights reserved