Building and Managing a Digital Collection in a Small Library

by *Wayne Wilson*

The creation and management of digital library collections is a relatively new field of librarianship that nevertheless has produced a substantial literature. Because the development of digital information resources can be an expensive undertaking, it is not surprising that the institutional pioneers in digital development typically were large academic research libraries or federally funded agencies. As a result, librarians and information managers from such institutions have tended to dominate the professional discourse on digitalization. At an April 2003 conference in Los Angeles presented by the Northeast Document Conservation Center, for example, the speakers were from Harvard University, Duke University, Cornell University, UCLA, the University of California–Berkeley, Columbia University, the Research Libraries Group, the National Archives and Records Administration, and the Library of Congress — hardly a representative cross-section of American libraries.[1]

The inclination of many librarians, when encountering a body of literature about the management of digital collections produced mostly by people from elite, affluent institutions, may be to dismiss it simply as inapplicable to their situations. This, however, would be a mistake. The issues facing the manager of a small library contemplating the creation of a digital collection are remarkably similar to those faced by managers of larger institutions. The example of the Amateur Athletic Foundation of Los Angeles Sports Library (AAF), a private library that has made the Web publishing of digital documents converted from paper resources a central part of its services, is a case in point.

The AAF library holds about printed 35,000 non-circulating volumes, 90,000 photographic images, and 6,000 video volumes on sport. The library has a full-time staff of five, including a Webmaster, and an annual budget of under $900,000. The library began digitizing selected paper copy publications in late 1997 and continues to do so each year. The AAF to date has converted about 170,000 pages from paper to digital format. These reside on the AAF's Web site as 37,000 PDF files.

The AAF digital conversion project is presented here as a case study of a digital project conceived and managed by the staff of a relatively small library with a specialized collection. This paper will explain the reasons the library staff wished to digitize materials, the major decisions that had to be made, unanticipated problems, and the impact of digitization on other

library services. This is not a primer on digital technology, but rather an exploration of the policy issues facing a librarian contemplating a first-time digital project. The focus will be on the issues and choices involved in creating digital collections by converting documents from paper. There are, of course, other ways to build a digital collection. Much of this discussion will be relevant to those scenarios as well.

The Decision to Digitize

The AAF's decision to convert paper documents to digital format grew out of a desire to make selected portions of the library collection more accessible. In simple English, we wanted more people to be able to use the library's books and periodicals.

The AAF was founded with the surplus funds from the 1984 Olympic Games. It began operations in 1985. The foundation's sports library opened in 1988. The library's mission is to provide information services that enhance people's understanding of sport and its social significance. A strong public service ethos has guided library policies from the beginning. The library's collections cover all aspects of sport, but its single strongest area is a collection of materials about the Olympic Movement. There are only a handful of comparable Olympic research collections in the world.

It became clear during the library's first years of operation that the Olympic publications were among the most heavily used items in the library. The people, who sought information about the Olympic Movement, came from throughout the United States and the world. Olympic researchers included scholars, students, journalists, and professional sport administrators.

It also became clear that the Olympic collection had only limited value to people, who were unable to visit the AAF in person. Library users could request a photocopy of a portion of an official report, or they could ask a librarian to read a portion of a document over the telephone. Neither of these alternatives was a realistic substitute for users being able to read an entire publication for themselves and most library users could not travel to Los Angeles to do research.

The AAF library staff considered various ways of dealing with the problem. One possibility was to microfilm selected titles and circulate the microfilm. Another was to circulate extra paper copies of selected Olympic publications. Still another was to digitize titles, put them on CD ROMs, and circulate the CDs. While each of these suggestions had merit, none offered a particularly efficient means of reaching large numbers of researchers. Distribution over the Internet seemed like a much better way to make information available.

A second problem with the Olympic collection was that it remained unknown to many researchers. People seeking Olympic information typically thought to contact the International Olympic Committee or other national Olympic committees. It did not occur to them to contact a new, private library in Los Angeles. The World Wide Web again offered a solution because Internet search engines could make the AAF's collection known to researchers, who otherwise would have been unaware that such a resource existed.

Thus, we began with a rather vague belief that the Internet could make the AAF's Olympic resources more accessible. The next steps were to (a) identify the clientele most likely to use and benefit from digital Olympic documents, (b) select specific titles to digitize, and (c) decide how digital documents should be presented on the AAF's Web site.

Envisioning the Digital Collection

Thinking carefully at the outset and asking questions about what information sources to put on the Web site, who would the primary users be, and

what kind of user interface would deliver digital documents to these users most effectively proved to be invaluable. Many subsequent decisions flowed easily from the original vision of what the digital collection should be and how it should function. The library served a variety of constituencies, including students, professional scholars, journalists, and other serious researchers.

Selecting which Olympic information resources to put on the Web site was relatively simple. The staff knew, from several years' experience providing traditional services, what publications were in demand and were difficult for researchers to locate in other libraries. Specifically, the staff believed that two types of documents would be especially valuable: Olympic Games official reports and back issues of *Olympic Review*. Olympic Games official reports are post-Games reports published by the Olympic organizing committee in each host city that include sports results and detailed information about the organizing effort. *Olympic Review*, which began publication in 1896, is the house organ of the International Olympic Committee. A digital collection of official reports and a complete run of *Olympic Review* would constitute a historically rich collection of primary documents of great value to Olympic researchers.

The staff felt strongly that the text of the documents was only one element of their historical significance. The photographs and other graphic elements also were important. Therefore, we wanted to replicate digitally the look and feel of the actual paper publications, retaining page layouts, photographs, graphic designs, and tables in virtually their original forms.

Formulating a vision of what broadly can be called user interface was a more difficult task. One fundamental question concerned searchability. One option was to offer Web site visitors a menu of publications to read, but to provide no search engine to retrieve those documents using keywords and word strings. At the other extreme was the option of making the digital documents retrievable using a search engine that could search the entire text of each document. Between these two poles were various middle-ground solutions. Any of the choices would have provided a useful resource for researchers. In the end, however, the AAF wanted to take full advantage of digital technology and therefore chose the more expensive option of making every word in every document searchable. This decision was a trade-off in which we sacrificed money that could have been used to digitize more pages for a plan that created a more powerful retrieval capability.

Pilot Project

Having decided to exploit the advantages of the World Wide Web, the AAF undertook a pilot project in late 1997 designed to test the reliability of the conversion process, the ability of the foundation's search engine and Web server to retrieve and deliver the digital document, and user response to a digital publication. We wanted to begin with something that was a substantial information source and would have some value as a stand-alone digital resource if we decided not to continue digitizing. That document was the official report of the 1932 Los Angeles Olympic Games. The report was more than 800 pages long, and contained text, photographs, and tables.

The pilot project was completed on time and on budget. User response, judged by e-mails, telephone calls, and face-to-face conversations with users, was quite positive. Based on this input from potential users, the staff decided to move forward with more conversions of Olympic materials.

Planning the Larger Project

Planning for a larger, on-going digitization project required us to resolve a number of issues that are certain to face almost anyone planning a project. The issues included selecting a document format, deciding whether to

outsource, handling of paper documents, preservation of digital resources, metadata creation, copyright, and managing the growth of the digital collection.

Format

The question of format had two aspects. First, what type of electronic file format should be used for storing documents? Second, in the case of a periodical, should users retrieve an entire issue, or an individual article?

The file format issue was relatively easy to resolve given the staff's desire to replicate the layout of original documents. AAF chose to use PDF (Portable Document Format). PDF had already emerged as a standard format and offered several advantages. Most importantly, PDF files effectively replicated the historical documents being converted *and* did so across a wide range of platforms and browsers.

There are various types or levels of PDF, all of which have different implications for Web publishers and users. Higher-level PDF files provide greater visual clarity and more reliable searchability. They also are more expensive. In other words, the basic question was one of quantity versus quality. A lower level of PDF would reduce the per-page cost, making it possible to convert more pages. Maintaining a higher quality at a higher cost would result in fewer page conversions. We chose the latter option — quality over quantity.

As for periodicals, the staff decided that presenting individual articles rather than whole issues would provide a more direct form of access for users and would enable users with slow connections to download more rapidly. This decision would prove to be a problem following the Supreme Court's 2001 decision in *New York Times Co. v. Tasini* decision.

In-House or Outsource

Another important question was whether to convert materials in-house using library staff, or to hire an outside agency to do the conversion. We elected to outsource the job for several reasons.

The library staff simply did not have time to devote to conversion. Doing the job in-house would have meant hiring and training new staff, purchasing or leasing new equipment, licensing new software, paying annual maintenance and support fees on equipment and software, and finding space for new personnel and equipment. Having new employees with new, unfamiliar duties to perform would have entailed new administrative responsibilities. The cost of outsourcing was only marginally more than the cost of hiring new staff with full benefits, and acquiring and maintaining equipment and software. Even if there had been a significant cost difference, it is unlikely that the required space could have been created.

The AAF has used four conversion agencies during the project. At present, the foundation deals with two companies—one American and one Indian. The American company is based in California. Its advantages include physical proximity to the AAF and continuity of service based on a long-time business relationship with the AAF. The Indian company is a newer business partner that provides conversion services at a lower cost.

The workflow with both companies is essentially the same. The library sends the vendor a paper copy document. The vendor scans the document, corrects errors that occur in the scanning process, compresses and optimizes the files, creates metadata according to AAF's specifications, sends metadata to the library for quality control, and FTPs the PDF files to the Web server. The library staff checks the PDF. Once the quality of the file is verified, the Webmaster uses the site's search engine software to index each word automatically in the PDF, makes any necessary changes to the search page interface, and makes the new file available to the public.

In truth, the division of labor described here is not quite so cut and dried. Responsibility for the creation of metadata has shifted from project to project; and, within the past 18 months, a part-time AAF staff member has begun to do a limited amount of scanning.

Handling Original Paper Documents

Some institutions undertake digital projects as a form of paper conservation. Digitization of a frequently used but fragile paper document, if it does not damage the original, is a valid form of conservation. The digital version can be used by thousands of people, while the original document remains inaccessible. The AAF did not view paper conservation as an important reason for initiating the conversion project. The primary motivation was to provide access to information. That motivation was reflected in the way the library dealt with paper originals.

The less a paper document is handled during conversion, the lower the conversion costs. Removing the binding of a volume and cutting the stitching holding together the individual pages makes it possible to feed a pile of pages automatically through a scanner. This is cheaper than handling a bound volume, or the folios and signatures from an unbound volume. Fortunately, in the early stages of the project, the AAF had two or more copies of almost everything that was to be converted. We were willing to sacrifice an extra copy of a particular title to achieve lower costs. This approach left the library with at least one paper copy in the library, as well as all of the pages of disassembled volumes. After conversion, the disassembled pages were stored off-site after being converted.

As the project progressed, however, it was not always possible to find an extra copy of everything selected for digitizing. So, in some cases, a professional paper conservator removed the binding, glue, and stitching from volumes, leaving the unbound sections or signatures intact. The sections were then scanned, with the AAF paying extra for the additional handling. When the conversion was completed, the conservator rebound the volume.

There were other situations in which the AAF had to borrow periodical volumes from other libraries to complete the digitizing of a title. Because the volumes belonged to another institution, it was not possible to disassemble the original. Similarly, there were cases in which the AAF owned titles that for one reason or another made the removal of bindings and stitching unfeasible. To deal with these cases, the library purchased a top-down scanner on which paper documents can be placed face-up and scanned. This method of scanning is a comparatively benign process that involves minimal wear and tear on the original. A part-time employee handles these documents and creates TIFF or JPEG files that are then sent to the AAF's conversion vendors for further processing.

Preservation of Digital Resources

Because the longevity of digital data is unknown, the preservation of digital resources has been a constant concern. It remains an area in which we are not confident of having found an entirely reliable solution.

The AAF's approach has been to keep a CD-ROM copy of the PDF files created for the Web site. The library also retains, off-site, CD-ROM copies of the intermediate files types that pre-date the creation of the PDFs. Five times a week, a tape back-up of the Web site is made. Each month, one of those tapes is archived off-site. In other words, at the end of each year, the AAF has twelve back-up tapes that are stored with the monthly tapes from previous years. As an additional layer of redundancy, the library keeps at least one paper copy of everything that has been converted to digital format regardless of whether the original was disassembled or not.

This may sound thorough, but the library, at this point, has not devel-

oped a systematic procedure for evaluating whether archived digital data has deteriorated. To date, spot-checking of archived files has revealed no problems. The project now is six years old, however, and it is time to identify and implement a more formal method of reviewing archived data.

Finally, even if the CD ROMs and tapes prove to be reliable long-term storage media, new hardware eventually will be developed to read digital data. Therefore, the AAF, like any other developer of digital resources, has to remain alert to changing technologies.

Copyright

If preservation of digital resources has been the most disquieting part of building a Web-based digital collection, the issue of copyright has run a close second. The AAF's digital collection includes materials that clearly are in the public domain, others that probably are in the public domain, and still others that unquestionably are the intellectual properties of other parties.

When it has been obvious, or even likely, that another party owns rights to a title, the AAF has obtained written permission before undertaking a conversion. The foundation has concluded agreements with the International Olympic Committee, scholarly associations, book publishers, individual authors, and national Olympic committees that hold rights to Olympic Games official reports. Typically, in an agreement, the publisher "warrants and represents" that it has the authority to grant the AAF permission to Web publish its property. Such agreements coupled with the fact that the AAF does not charge for the use of digital documents on its site diminish, but do not entirely eliminate, the foundation's legal exposure should anyone challenge the AAF's right to Web publish an article or monograph. The Surpeme Court's decision in the *New York Times Co. v. Tasini* case, though, is a cautionary reminder that seemingly unambiguous licensing agreements do not offer complete protection.

Jonathan Tasini and other plaintiffs were freelance writers. They objected to the New York Times Company and other electronic database producers selling the writers' articles without first obtaining their permission or sharing the revenue produced from the sales. The Supreme Court ruling distinguished between "[c]opyright in each separate contribution to a collective work" and "copyright in the collective work as a whole." The Court ruled that if a database "perceptibly presents the author's contribution as part of a revision of the collective work," an express transfer of copyright from the author is not required.[2] Conversely, explicit permission of the author was required if a database vendor presented individual articles outside the context of the collective work.

This was a problem for the AAF for two reasons. First, the AAF Web site presented each author's work as a "separate contribution." That is, the Web site served individual articles rather than the entire issue in which the article appeared. Second, at the time most of the articles on the AAF site were written, neither the authors nor their publishers gave any thought to the existence of the Internet. Therefore, the publishers never requested, and the authors never granted permission to the publishers to license, the publication of their individual articles on the Internet.

Even though the AAF Web site was non-commercial, and it seemed unlikely that authors would object to their works being published there, the foundation did not want to violate the ruling. Consequently, the library has had to modify the way in which it delivers articles from copyrighted periodicals. The AAF retrospectively is implementing a new interface that still presents each article separately, but creates an accompanying bookmark showing the issue's table of contents and giving the user the ability to display instantly any and all articles that appeared in the same issue.

The bottom line regarding copyright is that the AAF went into the project with the realization that there was some risk involved. The library staff has worked closely with an attorney to adhere to the law and minimize risk, but because of the changing and unsettled nature of intellectual property law, there are no guarantees of complete protection.

Metadata

A basic issue in any digital project is how much metadata to create and who should create it. Metadata are data used to identify and describe a digital information resource. Metadata also document how a digital resource was produced. Additionally, in an online retrieval system, metadata can provide points of access to documents in the digital collection.

If the term "metadata" seems reminiscent of the more traditional words "cataloging" or "indexing," it is because all three concepts are in fact closely related. As Jessica Milstead and Susan Feldman put it, "librarians and indexers have been producing and standardizing metadata for centuries."[3]

Deciding how detailed metadata should be is akin to deciding on the depth of cataloging or indexing in a traditional library context. The trade-offs are essentially the same, too. While deep, detailed metadata require more time and money than more cursory metadata, the more thorough approach provides greater control of the digital collection.

The AAF has relied on relatively limited metadata. The metadata that the public see in the AAF's Web-based digital collection consist of the basic bibliographic information needed to enable a researcher to identify and cite a work. The metadata also provide some points of access. For example, journal title, author name, and year of publication are elements of metadata that users can search on the Web site.

In more than 99 percent of the documents on the AAF site, users see only these short metadata displays. The exception to the rule is the Olympic Games official reports. Because many users are interested in the official reports as artifacts, each PDF includes a page titled "Digital Version Notes." The primary purpose of the notes is to describe any differences between the original publication and the digital file, as well as any elements of the digital document that do not display well online. Among the notes accompanying the 1932 Lake Placid Olympic Winter Games report, for example, are the following:

> The original paper version of the 1932 Winter Games Official Report has dimensions of 8 x 11'' (20.5cm x 28cm). The words "III Olympic Winter Games — Lake Placid, 1932" are printed on the spine. The back cover is plain. The cover has a pebbled surface with an inset frame. The cover image is embossed in gilt. The book has 291 pages with one blank page of heavy bond paper at both the front and back of the book.

> Special features of the digital version:
> • The back cover and the binding are not included in the digital version.
> • Blank pages at the front and back of the book are not represented in the digital version.
> • Photos in landscape format were rotated from vertical to horizontal to make viewing easier.
> • The digital version includes a bookmark list, which functions as a hyperlinked table of contents.
>   Selecting a topic heading will take you to the corresponding section in the document.
> • The final pages of the document contain an original index and list of illustrations for the book.
>   The index and illustration entries also are hyperlinked. Selecting an entry will take you to the corresponding section in the book.[4]

The "Digital Version Notes" also provide information on file size, file format, conversion platform, conversion software, conversion date, image resolution, digital fonts, and name of the company responsible for the conversion. The library maintains similar profiles of all converted periodicals. These data, however, refer to periodical titles only. They are not attached to each separate article, and they are not displayed to the public.

The decision to create and display minimal metadata stemmed from a couple of considerations. In most cases, the information contained in the metadata also exists in the PDF document itself. Many of the converted documents contain more bibliographic information than even the most conscientious researcher would need for purposes of citation. Second, while we were aware of the theoretical arguments in favor of detailed metadata, we simply did not have the time and money to create it for more than 37,000 documents.

In terms of document retrieval on the AAF site, metadata is in one sense superfluous because most, if not all, of the access points in the metadata exist in the full digital document, too. In fact, since each word of the document is searchable, the document itself provides many more access points than does the metadata. The value of metadata in document retrieval, however, is that it facilitates more precise searches by enabling researchers to limit queries easily to particular publication dates, journal titles, and authors' names.

The AAF's conversion vendors have written and input most of the metadata during the project. The notable exception to this practice was the metadata for the *Olympic Review*, which has had a complicated publishing history. Its page layout has been complex, making it difficult to determine where one article ends and another begins. The bibliographic information in the tables of contents does not always match the data in the body of the publication itself, and the review has been published in multiple languages. For all of these reasons, the AAF staff was responsible for *Olympic Review* metadata. It was a difficult and time-consuming task.

Responding to the Unexpected

All projects encounter unexpected obstacles. For the AAF, the most striking example of this truism was the dissolution of the company that produced and supported the AAF Web site's first search engine and digital collection management software. Anyone with any experience in library automation knows that the business failure of a vendor is not entirely unexpected. Nevertheless, the library suddenly was left with unsupported software that was a central component of its digital project.

The choices we faced were (1) to do nothing and hope that no technical issue arose that we could not solve or pay someone else to solve, (2) to purchase the source code so that the AAF could make alterations to the code if necessary and so that technical problems would be easier to solve, or (3) to acquire and implement new software. We chose the third option. It was expensive, and integrating the new software with the existing user interface and PDF collection was difficult, but the end result was a functional, well-supported search engine and collection management tool that incorporated advances made since 1998.

Not all unexpected occurrences are negative, though. For example, our original vision of the digital collection was one in which Internet users would find the AAF site using commercial search engines. They then would go to the AAF's search page and retrieve their documents. What we did not anticipate is that search engines such as Google would develop the capacity to search PDFs on a Web site and send users directly to the document, bypassing, at least initially, any other part of the site. A significant percentage of the PDF usage on the AAF site now comes from search queries initi-

ated off-site.

Managing Growth

The AAF digital collection initially consisted of primary historical resources. In time, the library staff wished to add scholarly journals that analyzed and interpreted the events chronicled in the official reports and *Olympic Review*, and also covered other non-Olympic aspects of sport studies.

The first scholarly journal added to the site was the *Journal of Sport History* published by North American Society for Sport History. Since the Web publication of the *Journal of Sport History* in 1999, the library has added eight other scholarly journals and more than twenty monographs. Among the other additions are late nineteenth- and early twentieth-century sports magazines, more Olympic Games official reports and an oral history series.

During the pilot project stage, the first PDF and the search engine resided on the same server as all other foundation computer applications. When the library elected to build a larger digital collection, the AAF purchased a separate server that was co-located at an Internet service provider. When the business failure of our search engine provider prompted the AAF to license a new search engine, the foundation also acquired a more powerful server intended to accommodate both the new search engine as well as continued growth for the foreseeable future.

Collection growth required not only more storage and computing power, but also Web site redesign. As the number of digital titles grew, the original user interface became inadequate. The library recently launched its third search page interface to accommodate a collection that has grown from a single title to more than ninety titles.

Impact on Other Library Services

Identifying and measuring indirect costs associated with the digital project is a matter of some interpretation, but it is reasonable to estimate that the combined direct and indirect costs of the digital project were about 25 percent of the AAF library budget in 2003. Obviously, a percentage of such magnitude detracts from other types of library spending.

The AAF has been willing to sacrifice spending in book acquisitions and other areas because the Web publication of selected library resources has enabled the library to meet its mission more effectively than ever before. Ensuring the use of the collection is central to our mission. The digital collection has dramatically increased the number of people using library resources and the number of times those resources are used.

If the downloading of a PDF converted from paper copy is viewed as analogous to the use of a book or periodical in the library, then usage of library materials — depending on how one measures it — has increased somewhere between fifty- and seventy-fold. This increase is even more impressive when one considers that the digital documents on the Web site converted from paper amount to fewer than 2 percent of all pages in the traditional library collection.

As an aside, it is worth noting that since 1985 the AAF has produced a wide range of research reports, coaching manuals, and newsletters. Most of these were produced in Microsoft Word. The library staff has easily and inexpensively converted the Word documents to PDF. They, too, are part of the digital collection, and like PDFs converted from paper, they receive heavy usage.

The introduction of Web publishing has resulted in a more international library clientele. During the nine years of the library's pre-digital period, the staff recorded inquiries and visits from 57 countries. The Web site now attracts users from 130 to 150 nations annually.

There has been no significant change in the number of reference ques-

tions handled by the library staff since the creation of the digital collection. The digital collection generates a certain number of reference questions, but it also provides information that obviates the need to ask other questions.

The most obvious impact on reference services is that the staff has become much more efficient in responding to historical questions, especially those dealing with the Olympic Movement. In the past, assisting users with historical questions often required looking at page after page in un-indexed or partially indexed periodicals. Searches that used to take thirty minutes or more now can be done in seconds, and links to relevant documents can be e-mailed to a researcher, or downloaded if the person is in the library.

Conclusion

There is no magic to building and managing a digital collection. The same knowledge and skills that make a good manager in a traditional setting contribute to success in the digital environment, namely familiarity with the resources in the collection, an understanding of the information needs of a particular library's clientele, and sensitivity to the way that people seek information.

Creating and managing a digital collection involves making a series of decisions. Those decisions will involve trade-offs and compromises. What works for one library will not necessarily work for another. Every critical decision will be influenced by local conditions. The key to managing a successful project is anticipating the issues that will arise and comprehending the trade-offs they will entail.

References

[1] *School for Scanning: Creating, Managing and Preserving Digital Assets, Los Angeles, California, April 23-25, 2003* (Andover, MA: Northeast Document Conservation Center, 2003). The Northeast Document Conservation Center is an excellent source of information on digital issues. See, for example, Maxine K. Sitts, ed., *Handbook for Digital Projects: A Management Tool for Preservation and Access* (Andover, MA: Northeast Document Conservation Center, 2000), available at http://www.nedcc.org/digital/dman.pdf (Accessed Sept. 29, 2003.)

[2] *New York Times Co., Inc., et al. v. Tasini et al.*, 533 U. S. 483 (2001), available at http://a257.g.akamaitech.net/7/257/2422/28jun20011200/www.supremecourtus.gov/opinions/00pdf/00-201.pdf (Accessed Sept. 29, 2003.)

[3] Jessica Milstead and Susan Feldman, "Metadata: Cataloging by Any Other Name," *Online* (Jan. 1999), available at http://www.infotoday.com/online/OL1999/milstead1.html (Accessed Sept. 29, 2003.)

[4] George M. Lattimer, ed., *Official Report, III Olympic Winter Games, Lake Placid, 1932* (Lake Placid, NY: Third Olympic Winter Games Committee, 1932), available at http://209.204.62.73/6oic/OfficialReports/1932/1932w.pdf (Accessed Sept. 29, 2003.)