Document-Management Technology and Acquisitions Workflow: A Case Study in Invoice Processing

Katharine Treptow Farrell and Janet E. Lute

Library acquisitions has moved from paper to online records for ordering and receiving, but the audit archive for invoices has remained largely paper based. Document-management technology (DMT) offers a solution to this condition. The authors survey the literature on DMT and its potential for use in the library acquisitions environment. This article considers the rationale and policy decisions that underpin the elimination of paper in favor of image files as an audit archive in library materials invoicing. A case study of the implementation of DMT to support and enhance traditional invoice processing in the acquisitions department of a large research library is included.

The paperless office has been a fiction for nearly fifty years. Indeed, quite the opposite seems to be the norm. Conventional wisdom suggests that rather than less paper, computers are responsible for more paper than ever before. E-mail alone has increased paper consumption by about 40 percent.[1] Library materials invoicing is a paper-intensive activity requiring extensive files for storing, referring, and consulting of printed documents and often results in the production of yet more paper in the process of generating payment requests from the parent institution.

Over the course of just two years, the Princeton University Library hired its first fiscal manager and migrated to its first integrated library system (ILS). The university moved to a new financial-management system, making new demands on the library in its interactions with the university financial system. The volume, as well as the complexity, of financial transactions increased, with heavy use of credit cards, online purchasing, and deposit accounts to support buying. Staff did not increase, nor is it likely to do so. Instead, new initiatives in services to users required reassignment of staff from back-room activities to public services activities. These factors taken together suggested the time was right to explore new solutions for managing the invoicing workflow.

Institutional demands for broader and simpler access to invoices, both paid and unpaid, made continued dependence on a single paper file unacceptable to collection development and technical services staff, both of whom required access to invoice documents. In addressing the need for better and more complete control over invoice documents, the university accounts payable office installed a document-management system (DMS) for use as an audit archive. This simple fact made the library initiative to implement a DMS possible, as it demonstrated that the corporate parent had already determined digital storage to be an acceptable archive medium. This deci-sion contrasts surprisingly with the findings of banking and insurance industries, which are the primary users of document-management technology (DMT). As is quite clear from the literature review, these financial institutions feel compelled for legal reasons to retain paper as the definitive archive.

The conscious decision to eliminate paper as an archive in support of audit requirements was based in equal measure on institutional need and philosophical commitment to greater efficiency; elimination or reduction of paper was not itself the goal. "Document imaging is all about digitizing paper information to improve the efficiency of downstream business processes and to ensure the proper management, control, and retention of that information."[2] In a paper-based environment, a single document might easily be handled by five or six different people, or at least five or six separate times by the same person. In a traditional library setting, opening a shipment of material, marking the order as filled, authorizing the invoice for payment, generating the check, mailing the check and the remittance copy to the seller, and finally filing the invoice copy to provide audit support are all steps that require handling the invoice. Paper can reside in only one location at a time, access is limited, the likelihood of loss either casual or catastrophic is considerable, and the cost of constant handling is great. At a time when the tools used to manage invoice processing were changing significantly and greater transparency in the financial process was increasingly demanded, a DMS seemed to make sense as the next step. "Document imaging was trendy in the 1990s, experiencing its sharpest growth when companies attempted to move away from manual, paper-based processes and 'go paperless.'"[3]

Literature Survey

A survey of library literature using the terms "document management," "digital storage," and similar rubrics revealed that virtually no applicable literature existed on the use of DMS in libraries to support technical-services activities such as invoicing and payment. Surveying the business literature with such terms as noted above revealed a very different profile. With a sixty-day limit on article publication dates, more than two thousand articles

Katharine Treptow Farrell (kfarrell@princeton.edu) is Head, Order Division and Assistant University Librarian for Technical Services and J**anet E. Lute** (jelute@princeton.edu) is Integrated Library Systems Coordinator, Princeton University Library.

were retrieved. Clearly not all were relevant, but a surprising number dealt directly with the precise issues that affect the library invoicing and payments workflow.

Prominent in the vast array of literature on DMTs were the banking and insurance industries. Both these sectors are paper-intensive and require numerous paper documents in support of transactions and audits. The following quotation from an article appropriately entitled "Paperchase" by Tyler provides a good sense of the issues confronting the insurance industry.

> Document management, in fact, is a double-entendre. I shall dismiss its interpretation by filing cabinet makers as that by which paper is filed with all due indexing for its easy retrieval. Document management that may be—and we must still by law store paper documents of various types—but we are concerned here with its meaning in an IT context. That means, essentially, shoving the paper documents through a scanner and having some software on the other end that can read all or key parts of those documents and/or respond to some manually input document reference and duly catalogue the electronic image.
>
> That's it in a nutshell—but like all nutshells, the interesting bit comes when you crack it open and sample the contents. Saving physical filing space is no longer the main aim of all this—it is speed of retrieval, removing the danger of loss and, increasingly, the ability to discover documents of relevance, the existence of which in a paper filing system would never be suspected.[4]

Most institutions are required by audit and tax law to retain copies of financial documents for up to seven years. The amount of storage space consumed is not inconsiderable even in a small operation. In a unit that handles between fifteen thousand and twenty thousand invoices annually for library materials of all types, the storage space amounted to a sizeable office in a building always cramped for space. However, as noted above, the motivation behind the move from paper to digital retention of audit documents was not space, or even the enormous amount of time invested in handling paper documents, but the security and accessibility of the document itself.

▌ Case Study

In early 2002, a small task force, including the authors and other key staff, explored and provided oversight for the implementation of a DMS for library invoicing. Working with the staff in the university treasurer's office and building on the software licensed by Management Information Systems, this staff successfully expanded the use of OnBase by Hyland Software, already in use at the treasurer's office, to library invoicing.

Princeton University Library has a primarily centralized materials-receiving and invoice-processing stream, with only the East Asian Library branch receiving material directly. This centralization should have made workflow analysis and restructuring somewhat easier. However, workflow was and continues to be one of the library's biggest challenges. Fifteen thousand invoices do not always enter such a large and complex system as Princeton's in neat groupings. Some invoices come in the first-class mail. The East Asian Library receives most of its invoices this way. Many firm orders and approval-plan invoices come with the material shipments. An increasing number of subscriptions and firm order invoices are received by electronic transfer or EDI (electronic data interchange).

In analyzing the existing paper-based workflow, it was apparent that invoices were handled numerous times, filed, unfiled, sorted, stamped, keyed into the invoicing module of the ILS, and finally filed, once a payment voucher had been issued. There were two key areas of concern, which the task force sought to address, as workflow was reimagined in an environment of digital storage for invoice documents. First, invoices arriving in advance of the firm order materials were queued in a paper file until the books were received and payment could then be approved. Second, some subscription invoices that were not clearly marked would make their way into the awaiting-receipt file and not be promptly approved for payment. In addition to managing these two critical issues, it was important to be able to improve processing time for payment of invoices and to gain better control over statement reconciliation.

OnBase, as installed for use in the library, has two basic components: the staff client, consisting of PC and scanner and the Web client.[5] Access to the Web client can be either read-write or read-only, depending on the requirements of the user. The need for two scanning stations was anticipated: one in the invoice unit in central technical services and one in the East Asian Library.[6]

The OnBase software allows keyword searching in the full text of scanned images. All the invoices submitted through East Asian Library, as well as many of the invoices handled in central technical services, are in languages that do not use the Latin alphabet. It is not possible to search these invoices by keyword. Invoices received by EDI are also not keyword-searchable as they are imported into OnBase as documents rather than being scanned using optical character recognition (OCR) technology. Therefore keyword searching, while useful, is not comprehensive.

One of the initial issues in setting up the region of the server for use by library invoicing was the determination of which indexes to establish. Five indexes have been defined at the library: vendor-invoice date, invoice number, vendor code, voucher ID as listed in the ILS, and check number. Vendor code, as stored in the ILS, was preferred to

the vendor name for reasons of efficiency and standardization in keying in the index point. The first three indexes are added at the point the document is scanned. The voucher ID is added after the invoice has been approved for payment in the local ILS. It is important to carefully consider which index points are needed and how these indexes are to be used. Original plans to add the check number to that index have been tabled as impractical. Clear links are needed between these indexes and the data elements in the local ILS so that reporting can be done efficiently. It is important to understand the implications of decisions made at this early planning stage, as it is often difficult to redefine data at a later stage.

Figure 1 illustrates the OnBase Web interface for library materials invoices. The indexes described previously are visible in the center segment. The example shows an invoice-number search with the result displayed below. The image of the document retrieved by this search can be viewed by selecting the document, as shown in figure 2.

OnBase has some basic reporting features. The reporting functionality is determined by the indexing decisions. Querying the OnBase client by key fields that are also present in the invoice record of the ILS allows for various management reports. For example, one can review scanned invoices by date or vendor to determine which invoices remain unpaid after a specified period of time.

Workflow, as previously noted, has proved to be the most problematical issue. As originally conceived, the goal at Princeton was to capture each invoice for scanning at the point of entry into the processing stream. Sellen and Harper, citing the *San Francisco Business Times,* Oct. 20, 1997, note that 3 percent of all paper documents are misfiled and almost 8 percent are eventually lost.[7] Scanning would allow for more comprehensive and secure control over the invoices. It was further imagined that most data entry or reconciliation with the local integrated library would be done from the scanned image viewed through the Web client. It was anticipated that all basic problem-solving and follow-up of vendor statements would also be done through the Web client. Realizing this workflow has been a challenge. Figure 3 illustrates the conceptual workflow.

There are two main areas of concern: ensuring that all documents are actually scanned at the point of entry into the workflow and changing the way staff work with and view the need for paper.

The flowchart shows three points of entry for invoices with the "invoice in the package" route accounting for the largest percentage. But there is overlap; invoices may come in first-class mail and then again with the package, or invoices may come in a package and by EDI. Multiple instances of the same invoice mean that care must be taken to avoid duplicate scans of the invoice and to identify as already scanned any transmission documents traveling with material to be processed.

Figure 1. OnBase Web Interface

Figure 2. Sample Scanned Document

The issue of how staff work with paper is less well defined. The significance of paper as a symbol in the workplace should not be discounted. "Paper as a symbol of the old-fashioned past is rooted in some real issues having to do with costs and interactional limitations."[8] In this case, management clearly viewed paper as a symbol of a less efficient, more expensive past, but staff appeared to have differing views. Paper is a familiar technology and does not require staff to invest in rethinking the way they work. The success of such a sweeping change in the way work is done by staff requires agreement and acceptance at all levels of the hierarchy.
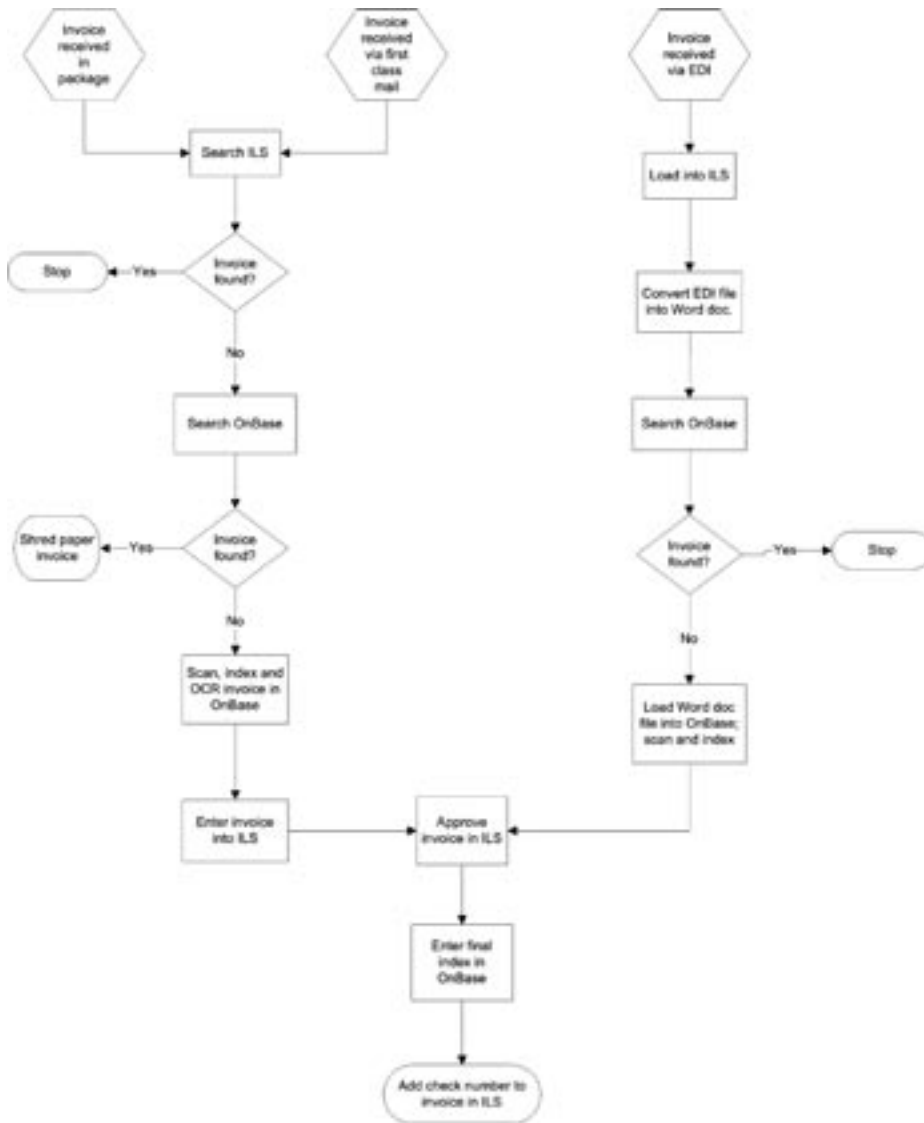
Figure 3. Invoicing Workflow

Staff resistance to certain aspects of the introduction of a DMS could be anticipated to some extent, but the avenues of this resistance were surprising. Disruption had been expected as the new workflow was introduced. In planning discussions, staff expressed their view that the new technology (scanning and indexing) was an added layer rather than a replacement for previous activities (filing and refiling). Indeed, for the first several months of the new fiscal year in which the project became active, duplicate systems were maintained as a precaution and reassurance. Staff were thus quite correct that this was additional work. What was unexpected was the resistance of staff to working with the scanned images in place of paper documents. New workstations with large monitors had been purchased expressly to facilitate working online with the document image on the Web client and the ILS client, but staff continued to prefer printing a copy of the scanned image on paper for the purpose of entering data into ILS. Additional training in Windows management has been provided to help staff manage their desktops more efficiently.

It is only when line staff can see tangible benefits to the new technology that fuller acceptance is gained, and real efficiencies are developed and even expanded. The completion of a full budget cycle has allowed the library to assess the gains resulting from the implementation of DMS. Staff have developed considerable expertise in using the scanner, which is no small feat. Invoices come in all sizes, shapes, and paper qualities. Experience has taught that not all invoices can be scanned into acceptable images, thus it has been necessary to review the original assumption that all invoices could be stored in digital form.[9] Fewer than 1 percent of the invoices are on paper that will not produce a readable image. Backup documentation of credit-card statements is the only example of paper technology being more efficient than the newer document-imaging technology.

Entering the second budget cycle of using DMT to control and make library invoices accessible, both management and staff can see clear gains. Invoice images can be shared, reproduced, reported on, and processed for payment in a more efficient and timely fashion. The average elapsed time for paying invoices has been successfully reduced to fewer than thirty days. Statements from agents can be reconciled more quickly and with less need for duplicating documents. Desktop access to the original image of an invoice has saved countless hours in

resolving questions from collection development or from agents. A recent audit by independent auditors required us to provide copies of twenty-five invoices from the previous fiscal year. Before the use of document-scanning technology, such a request would have taken staff several hours, and it is likely that at least one document would not have been located. Making use of the scanned images, staff were able to produce all the requested documents for audit inspection in fewer than thirty minutes.

Perhaps the most salutary effect of implementing DMT to track invoices has not been the gains thus far, but the potential for further changes in workflow, greater economies, and the expansion of DMT to other areas of technical-services work. Two areas ripe for exploration are managing licenses for electronic resources and managing approval-plan profiles and agreements.

The explosion in the area of licensing electronic resources over the last few years makes this an obvious application for DMS. These are legal documents; their content and wording have far-reaching implications, and the materials covered by them represent a large percentage of the library's materials budget. These are also documents to which many parties need access: acquisitions staff, collection-development staff, the institution's general counsel. The use of OnBase to support shared access and to allow annotations through the Web client should provide a secure and permanent record of decisions for each license. The intent is to scan and keep multiple versions of a license—the original, a working draft, the executed version, and subsequent revisions and addenda. The index points will be different from those for invoices. Each license will have a unique identifier, and each version will have a version number and date. There will also be indexes for the vendor, a generic title, and a link either to the ILS or to an electronic-resources management-database resource identifier. Status codes will identify the state of each version as working draft, executed, superceded, or other appropriate status.

An additional area for expansion of DMS is the control of approval-plan profiling and contract documents. Princeton manages more than thirty discrete approval plans, each with unique subject and nonsubject parameters governing the material to be supplied or treated. Many mainstream library suppliers provide access in their online system to an institutional customer's profile, although careful examination suggests that these online versions are rarely as complete as their printed counterpart. Most smaller vendors do not attempt to provide this kind of service, making the customer all the more dependant upon the printed document. Like licenses for electronic resources, approval-plan profiles and agreements are dynamic documents, changing over time to reflect the needs of the institution. Also, like licenses, these documents are frequently needed by different groups of staff and are version-critical. The adoption of

scanned, indexed images of profile documents will allow us to number, date, and annotate profiles in a shared environment so that collection-development and technical-services staff can both view and review these critical documents without resorting to filing cabinets or individual copies that then risk the introduction of conflicting versions of a profile.

▌ Summary

The following list of questions, posed by Gary Boomer of Boomer Consulting, is a useful guide in planning or implementing a document-management project.

- Do you want to implement firm standards, policies, and procedures?
- Do you want to reduce paper, supplies, filing, and retrieval costs?
- Do you want access to documents from remote locations via the Internet?
- Do you need the ability to destroy documents in accordance with firm policy?[10]

In addition to establishing procedure, professionals must determine the scope and complexity of a document-management project:

- Who will be in charge of the project?
- What is the timeline?
- What is the budget for hardware, software, implementation, and training?
- How do you plan to store the files?
- Do you plan to retrospectively scan old documents?

In planning the project, the authors established a timeline and appointed a team leader. The timeline slipped, as may be expected, but full implementation was successfully achieved within the targeted quarter. Budget was less of an issue, as the software was freely available to institutional users through the campus license. The library's cost for the two scanners was minimal. Training was accomplished with several simple instructional sessions and repeated practice. Files are stored on a remote server; no paper files are retained. Back-scanning the previous year's invoices was initially considered, but it was concluded that it was not worth the requisite staff time.

Faced with this decision again, the authors would make the same choice, but would recommend broader staff involvement at the initial planning stages, as well as the temporary addition of clerical or student staff to compensate for the additional workload at the start of the project. This is both a practical recommendation and an assurance for line-staff that there is management support where they perceive the greatest need.

The authors are pleased with the results of this initiative and are now seeing the benefits of greater control and efficiency for invoice tracking and management. The experience has emboldened them to proceed with other applications of DMT.

References and Notes

1. Samuel Greengard, "Getting Rid of the Paper Chase," *Workforce* 78, no. 11 (1999): 69.

2. "Imaging Gets a Second Look," *Information Week* (Jun. 23, 2003): 56–69.

3. Ibid.

4. Geoff Tyler, "Paperchase," *Management Services* (Aug. 2003): 24.

5. Staff workstations have a flat-screen, twenty-one-inch monitor that enables staff to have multiple windows open at the same time and still easily view documents. The Web interface for OnBase uses Internet Explorer, version 5.5 or higher, with ActiveX preferences set.

6. Specifications for scanning workstations include: 256 MB of RAM, SCSI card; Oracle ODBC drivers; OnBase client software. The scanners are Bell & Howell model 2000D FB, capable of handling thick, thin, fragile, crumpled, or duplexed originals.

7. Abigail J. Sellen and Richard H. R. Harper, *The Myth of the Paperless Office* (Cambridge, Mass.: MIT Pr., 2002).

8. Ibid., 25.

9. This limitation is partly a result of implementation decisions beyond the library's control. Use of grayscale for scanning is not an option for reasons of storage space. Had grayscale been enabled, most of the problem invoices could have been scanned.

10. Carly Lombardo, "Selling Imaging Systems—the Paperless Chase," *Accounting Technology* (July 2003): 39–42.