Introduction

An early subtitle of this talk was “How does this help you in your work?” This question was posed by a superior in the midst of an impassioned demonstration of the World Wide Web. The question made it clear that the benefits of computerization were not universally accepted. The project I am reporting on here is one attempt to bring the benefits of computerization to conservation and to help provide an answer to the question.

While there are many computerization needs in conservation, for the Conservation Department of the Philadelphia Museum of Art, the largest and arguably most important task was to gain fuller and faster access to the treatment information stored in 67 file cabinet drawers of conservation records. These records were compiled over most of this century and contain about 20,000 records. In the course of a treatment, these documents are usually consulted singly. Groups of records, let alone the records as a whole, are almost never considered. Comparative or statistical information involving groups of objects is rarely gathered because of the difficulty of doing so with paper documents.

This paper will discuss an effort to improve access to this information—the Conservation Database Project at the Philadelphia Museum of Art. Although the term “database” technically refers to the data set and “database management system” is the correct term for the software that is used to enter and use the data, this article will follow the widespread usage of “database” to refer to either the data or the software.

Background

The Philadelphia Museum of Art has lagged somewhat behind other large museums in the United States in its computerization efforts and in this regard the Conservation Department has mirrored the museum as a whole. At the start of the project in early 1996, the department had a staff of about twenty, with nine desktop computers, two of which were used by support staff and only five were 486-class or better machines. Apart from machines dedicated to running analytical equipment, these machines were used almost entirely for word-processing. The Conservation Department was not connected to the Museum’s Novell network. There was little digital imaging work of any significance going on anywhere in the museum and none in the Conservation Department. There were scattered uses of databases throughout the Museum. The largest of relevance to conservation was a database used by the Registrar’s Office using the DOS-based database software Q&A. Database experience within the Conservation Department was limited. There were databases for basic report tracking and for cataloging the departmental library (both using Q&A for DOS). There were also databases for cataloging a collection of architectural drawings which were undergoing basic conservation and rehousing as well as an early 1990’s survey of furniture and objects on loan in the greater Philadelphia area (both using the DOS-based software DataEase).

Project Inception

With this situation in mind, a proposal was made to create an easily searchable database of all the department’s treatment records. This database would enable the conservators to both enter new records and archive and utilize existing records in a faster and more efficient manner than would be possible with the existing paper-based system. The argument was made that computer hardware and software had improved to the point that a computer-literate conservator could either manage the adoption of a proprietary database system or create a system if no suitable commercial product was found.

Goal

The proposal called for purchasing or creating a system and within the space of one year, enter at least 3000 records or the output of the department generated over the last three years.
Funding
The Project was funded by the Museum’s Computer Services Department from funds given to the Museum to assist in automation, in particular, to create a collections management system to integrate many information functions within the Museum. Because the level of experience within the Conservation Department was far behind that of the Registrar’s Office which already had a functioning DOS system, the case was made that spending on a conservation database system would bring the Conservation Department up to speed, allowing for a smoother eventual integration into a collections management system.

Past efforts
A previous effort to create a conservation records database met with limited success and was never fully designed or implemented. There were a number of factors influencing this outcome. First, the computer hardware and software at the time were not sufficiently developed to create an attractive, easy-to-use system for a group of relatively inexperienced users. Secondly, the computer services person assigned to the task never had sufficient time to devote to the project. Related to this was the fact that as a non-conservator, the computer services person had to spend a great deal of effort learning about conservation concerns and practices. The current project sought to avoid these problems by having a conservator with computer experience run the project full time.

Design Phase
FilemakerPro selection
An early decision was whether or not to use one of the existing commercial systems or to make our own. An investigation of the conservation modules in current collections management systems revealed that none were quite what we wanted, and all were expensive. It was decided that rather than work with a vendor to improve one of the existing systems, we would learn far more by doing it ourselves and would achieve a result more in keeping with our particular needs.

An investigation into similar projects around the country was undertaken. I am especially indebted to fellow conservators Robert Futernick at the Legion of Honor in San Francisco and John Burke at the Oakland Museum for their time in advising me and showing me their systems. Many of the systems investigated were Macintosh-based and used FilemakerPro as the database software. Because of its ease of use and in order to draw on the experience of others in the museum field with Filemaker, it was selected. The cross-platform capability has been useful as the department recently purchased a Macintosh to use as a digital imaging workstation. Selecting Filemaker allowed us to integrate the imaging Mac into the database system, which was a priority. A parallel development using Microsoft Access was undertaken for a short time. It became clear that, although Access was more powerful, the learning curve was too steep and long to get an acceptable result in the time allowed.

Electronic Data Entry Phase
After designing the core of the system, the task was to get data into the system. Approximately 3000 WordPerfect documents existed on disk and these were “mined” first. These reports came from the five conservation sections plus the analytical lab and were the most recent output of the department, covering most of the last five years. For each conservation record, specifics about the object, the report and the conservator were entered in separate fields. For ease of viewing, printing and comparing disparate sorts of records, the entire text of each report was entered into a large text field. The data in this large text field looks nearly identical to a word-processed report. The records were entered at a rate of 5 to 15 per hour, with an average of around 12 per hour. Copies of the WordPerfect documents which were previously scattered throughout hard drives and floppy disks were stored on a network hard drive which was dedicated to the project, consolidating the department’s data storage.
Scanning Phase
After the available electronic documents were entered, we turned to scanning documents. Two scanners were employed. The first was an Agfa Arcus II, a high-quality, high-resolution scanner fitted with a transparency adapter. The cost was split between conservation and the database project so that the Conservation Department would have a high-quality color scanner for other uses during and after the database project’s operation. This scanner was primarily used to scan images for the image database. It has subsequently been used to scan other materials such as 35 mm slides, infrared reflectography and X-ray images.

The primary document scanning machine was a Microtek E6 equipped with a 50-sheet document feeder. OmniPage Pro Optical Character Recognition (OCR) software was used to turn the scanned images into text. After editing the text, it was saved as a WordPerfect file and entered into the database in much the same manner as the electronic documents. Perhaps this phase would be more aptly named the “scanning and typing” phase as many of the early documents were handwritten. The only way to enter these documents into the system was to type them in. Fortunately, most of the early documents consisted of brief handwritten passages on printed forms. Data entry could be made relatively quickly by constructing WordPerfect templates to automatically generate the form’s printed text with only the handwritten portions needing to be typed in. There were a relatively small number of forms and most documents had only a few hundred handwritten words at most, making this task nearly as fast as scanning.

The data entry was performed primarily by two people, one with a strong computer background and one with a strong interest in conservation. As it turned out, the person with the strong computer background was involved in the electronic data entry phase where her skills were of most use and the person with an interest in conservation participated in the scanning phase where his interest and knowledge were useful in correcting scanning errors and interpreting handwriting.

Data selection
In order to get a feel for the kinds of documents and to arrive at a more accurate assessment of the entire scanning task, one or two file drawers from each conservation section were scanned first. It was decided to scan records sequentially (the records were generally organized by accession number within each section) so as to minimize handling. In order to quickly ascertain whether the records in a folder had been entered into the database, the folder for each object was stamped with the date it was looked at and the number of records entered into the database. Individual reports within the folder were marked with a pencil check mark as they were entered.

Current Status
By the one-year mark in February 1997, the project’s data entry goal had been reached. More than 3100 records had been entered, nearly exhausting the pool of electronic records. This was accomplished in roughly three months, using about 12 person-weeks of data entry time. At the time of the AIC meeting in July 1997, 3700 total records had been entered. The additional 600 records were entered using another 10 person-weeks of data entry time. The main file had grown to 22 megabytes by this time. Because of a reduction in staff time allotted to data entry, the goal of reaching 5000 by the end of June 1997, was not met. This goal was reached in September 1997.

Costs
Costs are always a consideration, even (sometimes especially) in the nonprofit world. For this project, the main costs were up front in the development stage, with the costs per record rapidly decreasing as data entry got underway. At the point where the first 3000 electronic records had been entered (nine months of development, 12 weeks of data entry) the cost per record (counting all costs for the project) was approximately $15/record. If the records in the converted and subsidiary databases (discussed later) are counted, the cost drops to $12/record. Were all the 20,000 records entered at the current rate, the final cost would be about $3–4 per record. The cost of just the data entry alone is between $1–2 per record. This is at the rate of 5 to 10 records per hour at $10/hour for data entry. Considering that the Department’s total cost per treatment record (the department’s yearly budget divided by the number of records generated per year) is close to $1000, the database cost seems negligible compared to the gain in information accessibility.
Training Phase
At the point of the presentation of this paper at AIC, the project was about to enter the training phase. There were nearly twenty users who needed to get fully conversant with the database. Much of the training occurred during the design and development phases, but during the phase dedicated to training, every conservator got four to ten hours of individual or small-group training. Although there is on-screen help specific to the use of the database, it is quite clear that users consult such help as an absolute last resort and are far more comfortable with individual training from a human being.

Selected Screens
There is no substitute for actually using an application, but the following figures will give a sense of the database’s features. Although the illustrations shown here are all gray-scale, the screens’ components are color-coded to hopefully be both pleasing and lead the user to visually group related functions or areas. Filemaker uses the term “layout” to refer to what is more commonly called a “screen.” I will use the terms interchangeably.

The opening menu screen (fig. 1) shows two main areas. On the right is a group of menus for each conservation section which take the user to a list of forms used primarily to create new records such as examinations for loan, condition report/treatment proposal, or treatment reports. Figure 2 shows the data entry screen for a furniture examination for loan. Figure 3 shows the same record as it looks when printed—nearly identical to the original WordPerfect format.

On the left side of the main menu (fig. 1) are buttons which take the user to various summary or general information screens. These are used to search the entire database in various ways, to view records across sections, to view related images, view related records in other database files,
or to enter or edit keyword entries. Figure 4 shows one such layout, here with a drop-down list for report type. Another layout (fig. 5) has a self-relation which shows the presence and number of any other records for the same object. A keyword screen was created for each conservation section so that records could be quickly searched by a consistent set of criteria. Keyword entries can be quickly selected from pick lists but usage of the list is not mandatory. A complex script (written in Filemaker's macro language) which searches the text of a report and automatically builds keyword entries has been run on nearly half the records to date. Figure 6 shows the Furniture keywords screen in development. There are four main keyword fields; object materials, object condition, treatment method and treatment materials. There are also some secondary keyword fields that varied among the conservation sections (for example, furniture ultimately had an additional field for secondary wood).

Private Conservator Use
As a former private conservator, I wanted to spend a little time discussing the applicability of databases for private conservation. I have reluctantly concluded that for most of the functions that a private conservator might use a database, the common office applications will probably provide a faster and easier solution. For moderate to large survey projects, however, the investment required to learn basic database operations is easily worth the effort. For those with the computer “bug,” the capability to easily customize an application may also be worth the effort. For those interested in constructing their own systems, Filemaker is inexpensive, has relational capability, is cross platform, and is probably the easiest to use of the major databases. In fact, one of the PMA Computer Services people who was introduced to Filemaker through this project said that in contrast to other software, “What I like about it is it’s fun to figure out.” Filemaker comes with a good set of templates for common business functions which can be customized. There are also excellent Filemaker resources online, some of which I used almost daily during the project.

Subsidiary Projects
As mentioned earlier, there were a number of subsidiary database projects that were undertaken in conjunction with or parallel to the main project. These were pursued because either particular needs arose or time was available due to delays in hiring, equipment purchasing and procuring work space which arose in the main project. The following briefly describes these projects.

Wood ID
Of particular interest to furniture conservators, a Wood ID database was constructed, instigated by a one-week wood identification session with Dr. Eastman.
Bruce Hoadley. This database kept track of his identifications and can be cross-referenced to the main records database and to another subsidiary database of common and scientific names for wood species. Thanks go to furniture conservator Joe Godla at the Getty for sending a template constructed for a similar session with Dr. Hoadley which formed the basis of PMA's database.

Analytical
An Analytical records database was created to hold the output of the department's analytical lab. The information in these records is an important addition to the data in the main records database. In various places in the main records database, the existence of any analytical reports for a particular object is flagged. The user is able to go to any of those analytical records with a click of a button.

Arms and Armor
A database was created for a two-year conservation project dealing with objects in the Museum's Kienbusch arms and armor collection. Although these records are being used apart from the main records database, the intent is to roll them into the main database at the conclusion of the project. The basis of the database was a survey of 1229 objects done with Q&A for DOS. The data was converted to Filemaker and the acronyms and abbreviations used in the survey expanded to full phrases for easier usability. One interesting aspect of this database is that both conservators and curators have access to it. This is the first time within the museum that these two parties have had access to the same live data. Another feature of note in this database is that it keeps track of location changes as the objects pass from the gallery to conservation to photography back to conservation and then finally back to the gallery. Formerly, each of these transactions would have been sent to the Registrar's Office individually for noting on index cards. Now, the entire move and location history will be stored and sent to the Registrars for entry into their databases at the end of the project.

Image Database
An Image Database was created to hold various sorts of digital images. This was done as a "proof of concept" exercise rather than to provide a large store of images. Thus, only about 100 images are currently held by the image database. The images are of before, during and after treatment photography, UV microscopy, infrared reflectography and X-ray images. There are both black and white and color images. The image files are mostly .pct files which Filemaker handles well and are of moderate size, 30-200k. In most cases, a higher resolution version of each image was created and stored as a compressed .jpg file. Although the version of Filemaker used for the
project (version 3.0) cannot handle jpeg files, the current version (4.0) does and it is hoped that the higher resolution images will ultimately be entered into the image database. During the course of the project, it was decided to split the image database into two files. The “thumbnail” image database handles relatively low resolution images which are meant simply for basic visual object identification: there is only one image per object. The main file holds all the images, including thumbnails, and can have multiple images per object. All images are entered into the main image database and only thumbnail images are exported to the thumbnail database with a one-button export process. Figure 7 shows a screenful of images in the main image database. Figure 8 shows a screen which calls up all images relating to a single object.

The images have proven, as anticipated, to be a great “hook” into using the database, and have served as a good introduction to digital imaging. There is also the hope that some of the images conservation generates will become valuable as a museum-wide resource.

DataEase Conversions
As mentioned earlier, a number of the subsidiary projects involved converting and modifying existing databases. Two of these were from the DOS-based system DataEase. One was a survey (condition reports and treatment proposals as well as some treatment reports) of furniture and objects belonging to the museum on loan to a number of park houses and other sites throughout the Delaware Valley. Because of the difficulty of using the program, the data in these databases were all but lost. Converting them to Filemaker and integrating the records into the main records database has made that information accessible again.

Another DataEase database that was converted was a partial accounting of the museum’s large store of architectural records relating to the construction of the museum itself. A conservation grant was used to assess the condition of the records and to devise improved storage. The database has fields noting the drawings’ contents, condition and location and is used by conservators, as well as buildings and operations staff. The data was corrected to some extent and a menu system was devised to make the database easier to use.

Twentieth-Century Furniture
In order to hold records generated by a conservation intern performing a survey of the museum’s 20th-century furniture, the furniture portion of the main records database was “cloned” to provide data entry forms which could be installed on a laptop.

Frames Database
An area that has been neglected in the past but is receiving much more attention in recent years is
the museum's store of frames. Only recently have frames been treated as objects separate from the pictures they hold. When the museum's frame storage was renovated, it seemed sensible to create a small subsidiary project to store information about the newly housed frames. A database modeled after a similar one at the National Gallery of Art in Washington DC was developed. Two features of note in this database are the ability to handle dimensional data in inches or centimeters (with one-button conversion) and a field for a rendering of the cross-section of the frames. One goal was to undertake a test barcoding effort with this collection, but time did not allow this aspect to get underway. Research at the time indicated that incorporating barcodes into labels on the frames could be easily done from Filemaker and would be relatively inexpensive.

**Summary**

Overall, the project was successful in using recent advances in computer technology to help make the practice of conservation a little more efficient and better organized. It was extended twice: a one-year project turned into a year-and-a-half one. As a joint venture of sorts, it showed that cooperation between departments within the museum could work. The benefits of having access to treatment information in a searchable database have been demonstrated. With only about a quarter of the records entered, however, we are just beginning to make use of this resource. Having a conservator in charge of the project did allow for more design control than would otherwise have been the case and certainly made the customization process quicker. We learned how to manage the mundane data entry part of the project in an efficient manner. Both the electronic and scanning portions entered records at acceptable rates.

Archiving the department's treatment records was the highest priority of the department's computerization tasks but there are many others. Experience with this project gives the department the capability to create other databases for data such as invoices, supplier and materials lists, MSDS's, library holdings, etc., projects which would never be implemented by the Museum's Computer Services Department.

The subsidiary projects proved to be extremely valuable and probably would not have been attempted had the project been managed by a consultant or other “outside” person.
When a full collections management system is installed, the Conservation Department should be well-placed to either integrate its system into the larger system or maintain its system alongside the collections management system if that system proves unable to perform the functions conservation needs.

Disaster Readiness
Perhaps an overlooked benefit of this computerization process is its potential role in a disaster situation. Electronic versions of records can be easily stored offsite. Copies on removable media or a laptop computer are very portable. Considering that most of the museum's most important objects have been treated in the Conservation Department, it is quite likely that any special conditions relating to an object would be noted in the treatment records and could prove valuable if readily accessible during or after a disaster.

As in all such projects, there are areas that could stand improvement. Filemaker proved to be a good choice for the database software, although it has some significant drawbacks. Because the information stored in the database consisted of sometimes lengthy text records, not a typical database information type, Filemaker's word processing tools were put to the test. While they are quite good, they cannot fully replace a word processor. For example, it is not possible to have a table within a text field and text copied from a word processor to the database loses all special formatting, such as justification, bolding, or italics. Printing was another problem area. Filemaker has particular difficulty with some Hewlett-Packard printers—which unfortunately are used throughout the department. Finding printer drivers that would work with both the software and the printers eventually solved the problem but product incompatibility was an annoying impediment. The sliding mechanism, whereby fields contract to close up any blank space, did not always work as expected when layouts grew complex. Since Filemaker does not allow fields to expand when printed, only to shrink, printing often necessitated creating an additional dedicated printing layout with fields large enough to hold the largest entry possible. A possible solution to these problems would be to have a link to the original document in its native file format. Experiments to create this capability using OLE (Object Linking and Embedding) were unsuccessful, with all too frequent crashes. Time did not permit additional testing after the department was fully populated with Windows 95 machines although it is likely that this capability could be implemented under Windows 95. In preparation for this possibility, when a document was copied from a WordPerfect file, the full pathname (where possible) of the WordPerfect file was kept in a field so that eventually a user could activate the original document in its native application from the database.

Figure 7: A “slide show” screen in the image database showing some of the information recorded for each image.
Apart from problems with the software, there were other difficulties encountered along the way. The full suite of computers necessary to make the project fully operative did not arrive until the very end of the project. Although we were able to make do prior to the new equipment's arrival, it would have been much easier to develop the database and train users had the equipment been available earlier. With the new equipment each conservation section for the first time has at least one networked, up-to-date PC with a large (17") monitor. There are also two laptop computers. This equipment means there is now a staff to PC ratio of about two. Achieving this equipment level was an early goal of the database project. Having fast computers with monitors capable of showing high resolution images was critical to the success of the project.

Although the project would not have been undertaken without the funds from the Computer Services Department, the lack of control of the funding meant that Conservation was ultimately not in charge. While wider integration into the museum's information systems will allow better cooperation with others such as registrars and curators, a major shortcoming may be that the passing of control out of Conservation will mean a slower development of enhancements and lack of customization. The decision was made to postpone additional data entry and any further development until a collections management system is installed. Hopefully, the break in continuity will not seriously affect the Conservation Department's database usage or imperil the ultimate integration of the roughly 15,000 paper records which remain to be entered into the system.

There were organizational problems outside of the control of the project which were probably exacerbated by the project's "in-between" status, that is, as a joint venture of the Computer Services and Conservation Departments. Hiring the data entry personnel was delayed for months because there was no space available. The lack of space was primarily due to the Museum's hosting of the blockbuster Cezanne show. Because the training time had to be compressed at the end of the project, scheduling busy conservators for enough time turned out to be quite difficult. During the last phase of the project, there were personnel changes at the head of both the Computer Services and Conservation Departments which orphaned the project.

There were (and are) problems inherent in the project's relatively short duration and pilot status. Long-term maintenance of the database and its eventual migration to other software or a collections management system are areas that have not been adequately addressed. Every effort was made to predict the department's long-term needs and technological upgrade path but these are inherently unknowable.

Figure 8: This screen in the image database has a self-relation allowing it to show all images for one object (i.e. those with matching accession numbers).
The author would like to thank his initial supervisor, Bob Leming, then the head of Computer Services, now the head of an Internet service provider, for his persistence in garnering management’s assent to the project and for his willingness to let an untested conservator make this foray into cyberspace.

About the Author
PETER EASTMAN worked privately as a furniture conservator for many years in Berkeley, California, before entering the Masters program in Wood Science and Technology at the University of California at Berkeley. From 1993 to 1995 he was assistant conservator of furniture at the Philadelphia Museum of Art and for a year and a half in 1996 and 1997 he worked on the Conservation Database Project. He was briefly an Assistant Scientist at PMA, and is currently a scientist in the microscopy group of the analytical research division of the Rohm and Haas Company in Bristol, PA.